JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Article

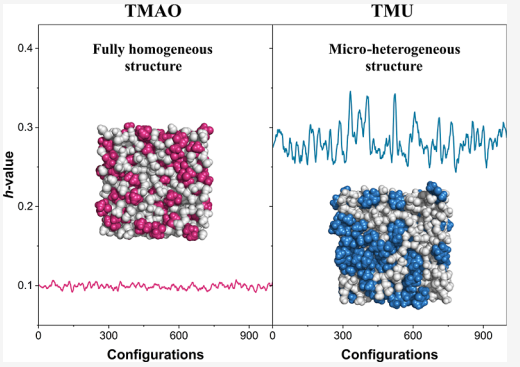
Molecular Aggregation Behavior and Microscopic Heterogeneity in Binary Osmolyte−Water Solutions

Jiwon Seo, Ravi Singh, Jonghyuk Ryu, and Jun-Ho Choi*

Cite This: https://doi.org/10.1021/acs.jcim.3c01382 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Osmolytes, small organic compounds, play a key role in modulating the protein stability in aqueous solutions, but the operating mechanism of the osmolyte remains inconclusive. Here, we attempt to clarify the mode of osmolyte action by quantitatively estimating the microheterogeneity of osmolyte−water mixtures with the aid of molecular dynamics simulation, graph theoretical analysis, and spatial distribution measurement in the four osmolyte solutions of trimethylamine-*N*-oxide (TMAO), tetramethylurea (TMU), dimethyl sulfoxide, and urea. TMAO, acting as a protecting osmolyte, tends to remain isolated with no formation of osmolyte aggregates while preferentially interacting with water, but there is a strong aggregation propensity in the denaturant TMU solution, characterized by favored hydrophobic interactions between TMU molecules. Taken together, the mechanism of osmolyte action on protein stability is proposed as a comprehensive one that encompasses the direct interactions between osmolytes and proteins and indirect interactions through the regulation of water properties in the osmolyte−water mixtures.



1. INTRODUCTION

Osmolytes are small organic molecules that affect the stability and solubility of the protein in aqueous solutions, and significant efforts are being made to shed light on their mechanism of operation.[1−4] Based on their effect on protein stability, osmolytes can be classified as protecting and destabilizing, with protecting osmolytes pushing the equilibrium of the protein folding process toward the folded state, while destabilizing osmolytes shift the equilibrium toward the unfolded state.[4] Despite numerous experimental[5−8] and computational attempts[9−14] to investigate the mode of action of osmolytes, the definitive understanding of the molecular mechanism for osmolyte-induced protein stability remains uncertain. One of the plausible hypotheses involves preferential interaction between protein and osmolyte, where protecting osmolytes exhibit a tendency to be preferentially excluded from the protein surface, whereas the destabilizing osmolytes interact favorably with the amino acid residues on the surface of the protein.[4,11,15−18] Another hypothesis emphasizes indirect interaction by modulating the water H-bond network structure in the presence of an osmolyte that can either lead to weakening or enhancement of the water−water H-bond interaction.[5,10,12,13,19] Interestingly, several research groups suggest that both direct and indirect mechanisms are involved in elucidating changes in protein stability and functionality.[12,20]

On the other hand, several spectroscopic measurements[21−27] and computational simulations[28−35] have supported evidence for microscopic heterogeneity, indicating a nonuniform distribution of component molecules in various aqueous mixtures. This spatial inhomogeneity of dissolved molecules in liquid water is determined by the complicated contribution of intermolecular interactions of components such as solute−solute, solvent−solvent, and solute−solvent. Using molecular dynamics (MD) simulation and the Voronoi polyhedra method, the self-association behavior in urea molecules was observed with an average size of 3−4 molecules.[36] This urea self-aggregation appears to be due to the relatively weak H-bond interaction between urea and water compared to that between water molecules.[37] Stirmemann et al. conducted MD simulation studies in concentrated aqueous solutions of amphiphiles, showing that tetramethylurea (TMU) molecules are aggregated at low concentrations, while self-association of trimethylamine-*N*-oxide (TMAO) is not observed even at high concentrations.[38] Here, it was also revealed that the water dynamics is significantly retarded in concentrated TMAO solutions, whereas the translational and reorientational motions of water are not largely affected near TMU aggregates. The distinct aggregation patterns of TMU and TMAO, which are mainly generated depending on the interaction of osmolyte with water, were also reported in previous studies.[31,39−44] In the case of dimethyl

Received: August 30, 2023
Revised: November 7, 2023
Accepted: November 7, 2023

sulfoxide (DMSO)−water mixtures, the aggregation-induced microscopic heterogeneity was also observed using two-dimensional infrared (2D IR) spectroscopy[45] and MD simulation analysis,[30,33,34] where Oh et al. reported the fastest picosecond H-bond dynamics in a certain DMSO concentration range of 20−50 mol % due to the existence of confined water in the binary mixture.[45]

Meanwhile, by employing the combination of MD simulation and graph theory, the aggregation behavior of solute and water molecules has been examined with the calculation of network properties of molecular aggregates in a wide range of binary mixtures of water including salts,[46−49] osmolytes,[9,13,50] alcohols,[51−56] and so on.[57−59] Sundar et al. conducted a detailed analysis, calculating the radial distribution function (RDF) and graph theoretical properties of the H-bond network of solution in various osmolyte−water mixtures.[9] The results indicate that the local solvation structure and overall H-bond network properties in urea and glycerol solutions are similar regardless of the nature of the osmolyte, whether it is a denaturant or stabilizer. Lee et al. performed graph theoretical analysis in aqueous osmolyte solutions by constructing separate adjacency matrices for osmolytes and water and showed that the morphological structure of aggregates composed of protecting osmolytes such as sorbitol and trimethylglycine is distinct, in comparison to the destabilizing osmolyte urea.[13] They also showed that the water H-bond structure is influenced differently according to their aggregation pattern.

In various concentrated aqueous solutions, a solvation-induced molecular aggregation pathway has recently been proposed in use of a combination of MD simulation and graph theoretical analysis, which reveals formation of distinct molecular aggregates, categorized as either self-associated or spatially extended, depending on their interactions with water.[32,46,52,55,58] For example, the primary alcohols of methanol and ethanol, which have a short carbon chain, form spatially extended alcohol aggregates while retaining the interactions with water, but the *n*-butanol prefers to form self-associated aggregates to avoid interaction with water due to their relatively long carbon chain. These two types of alcohol aggregates possess distinct morphologies in their respective graphs and have varying effects on the water H-bond structure and phase behavior in the binary alcohol−water mixtures. That is, self-association of alcohol aggregates and their growth facilitate separation from liquid water in *n*-butanol−water mixtures, exhibiting a miscibility gap in the phase diagram. Contrastively, the expansion of spatially extended alcohol aggregates with preferred interaction with water can be compatible with the water H-bond network and explains the full miscibility of methanol and ethanol at all concentrations.[52] This bifurcating pattern of molecular aggregation pathways was shown in various aqueous mixtures comprising ions,[46,48,49] alcohols,[52,55,56] osmolyte,[13] and methane,[58] and the relationship with the physical properties of solution such as solubility and liquid−liquid phase separation was well established.

Furthermore, the measurement of spatial inhomogeneity and analysis of network properties of molecular aggregates enable them to demonstrate the effect of molecular aggregation behavior on the spatial distribution of components and, consequently, on the phase behavior in binary mixtures.[32,58] In the case of methanol−water mixtures, the development of spatially extended methanol aggregates was found to result in a uniform distribution of component molecules, leading to the formation of a homogeneous solution. Conversely, the growth of self-associated dichloromethane (DCM) aggregates induces heterogeneous distribution of water and DCM molecules, eventually bringing about liquid−liquid phase separation in a given composite system. The degree of microscopic heterogeneity in the cosolvent−water mixtures was quantitatively estimated with the *h*-value, which is the descriptor of measuring spatial inhomogeneity, representing that the *h*-value corresponds to 0 and 1 for a perfectly uniform distribution and an extremely localized case, respectively.[32]

Here, we arrange four osmolyte−water mixture systems at the two representative concentrations of 5 and 10 m, including two destabilizing osmolytes, urea[60] and TMU,[61,62] a protecting osmolyte, TMAO,[60] and DMSO as a cryoprotectant.[63−65] Although all four osmolyte molecules are completely soluble in liquid water within the concentration range to form macroscopically homogeneous solutions, the presence of microscopic heterogeneity in the binary mixtures is anticipated due to varying interactions between water and osmolytes. MD simulation and relevant analysis of four osmolyte solutions were systematically performed, and the network properties of osmolyte and water aggregates were examined based on graph theoretical analysis. Moreover, the spatially inhomogeneous distribution of osmolyte and water molecules was quantitatively described by calculating the *h*-value from the corresponding MD trajectories of the aqueous mixture. Taken together with the measurements of spatial inhomogeneity and analysis of network properties of the given molecular aggregates, it addresses how the osmolyte aggregation pattern is differentiated by the molecular property of each osmolyte and how morphologically distinct osmolyte aggregates influence water H-bond structure and, furthermore, microheterogeneity in a given binary mixture system.

2. MOLECULAR DYNAMICS SIMULATION ANALYSIS AND RESULTS

2.1. Molecular System and MD Simulation Detail. Four osmolyte molecules, DMSO, TMAO, TMU, and urea (see left side in Figure 1) were chosen to arrange osmolyte−water mixture systems with two representative concentrations of 5 and 10 m at 273 K. Although all the osmolyte molecules are completely dissolved in liquid water, forming a single liquid phase within their respective concentration ranges, their aggregation behaviors were reported to be distinctly different. It was found that DMSO and urea make self-associated molecular aggregates under ambient conditions,[13,30,33,34,36] and the amphiphilic compound TMU forms hydrophobic aggregates even at a mole fraction of 0.005.[31] In contrast, TMAO molecules do not exhibit a noticeable aggregation behavior, even at a high concentration of approximately 8 m,[38] while maintaining a preferential interaction with water.[43]

To create the osmolyte solutions of 5 and 10 m, 90 and 180 osmolyte molecules, respectively, were added to the entire box containing 1000 water molecules. A single-component system of pure water or pure osmolyte, each containing 1000 corresponding molecules, was arranged to obtain the distance criteria to define water−water and osmolyte−osmolyte interactions, which is required in the description of molecular aggregates in a given graph representation. Each osmolyte−water system was constructed using Packmol,[66] and the general AMBER force field (GAFF) in the Antechamber module of the AMBER 20 package was utilized to generate the force field of osmolyte molecules.[67] The particle-mesh Ewald method was employed to take account of the long-range electrostatic interaction, and a cutoff distance of the nonbonding interaction is 10 Å.[68] Prior to

https://doi.org/10.1021/acs.jcim.3c01382
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

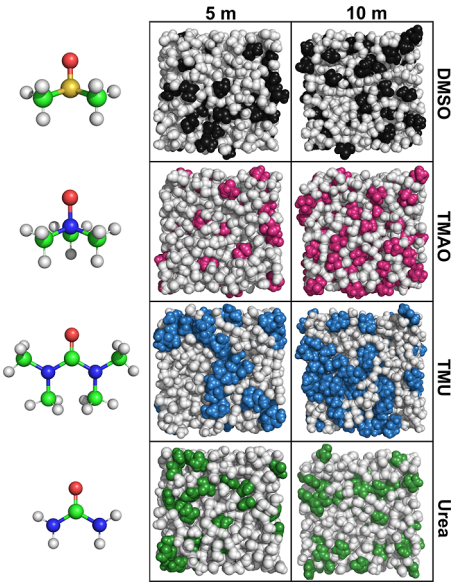


Figure 1. Representative snapshot configurations of each osmolyte/water solution at 5 and 10 m at 273 K. The water molecules are represented by bright-gray spheres, and DMSO, TMAO, TMU, and urea molecules are colored black, pink, blue, and green, respectively. Four osmolyte molecules are fully dissolved in liquid water, macroscopically forming a homogeneous solution, but TMU molecules exhibit an existence of large osmolyte aggregates even at a lower concentration of 5 m which represents microscopic heterogeneity in the spatial distribution of components.

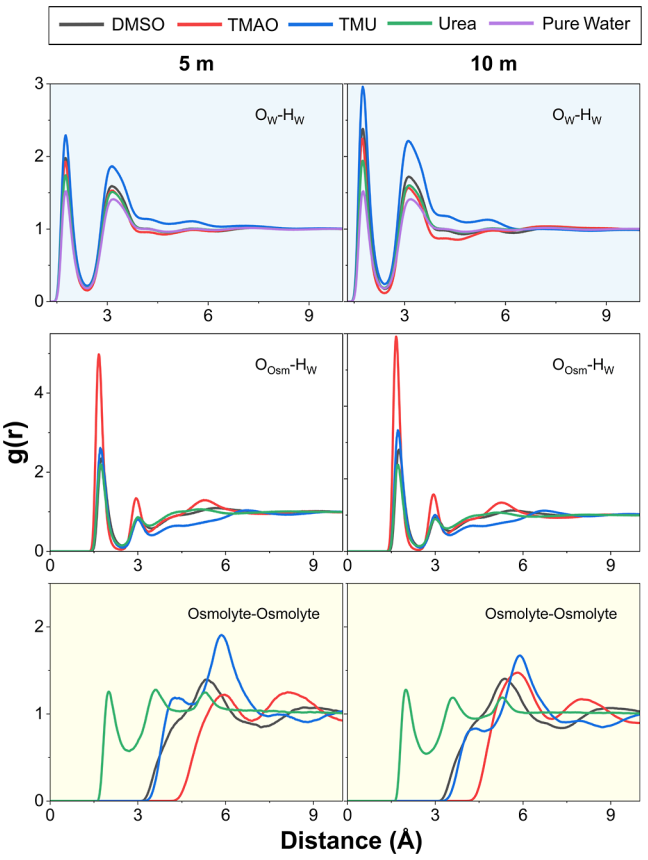


Figure 2. Radial distribution functions of each osmolyte solution at 5 and 10 m. O and H atoms in water (O_W−H_W), the O atom in osmolyte, and the H atom in water (O_{Osm}−H_W) are chosen to investigate the H-bond interaction of water−water and osmolyte−water, respectively. For the osmolyte−osmolyte RDFs of DMSO, TMAO, and TMU, the central heavy atom of each osmolyte is chosen, that is, S, N, and C atom, respectively, while O and H atoms in urea are selected to investigate intermolecular H-bond interaction between urea molecules.

the equilibrium MD simulation, the osmolyte−water mixture systems were energy minimized by the steepest descent and conjugate gradient methods. Subsequently, 2 ns equilibrations in the *NPT* ensemble were performed at a constant pressure of 1 atm and a constant temperature of 273 K, to adjust the density of the two-component and single-component systems. An additional 2 ns NVT simulations were carried out at 273 K with a 1 fs time step to equilibrate the binary and neat systems. Finally, for the pure water and osmolyte and aqueous osmolyte solutions, 100 ns NVT MD trajectories were produced with 10 ps intervals, and saved 10 000 configurations were used to extract structural information on osmolyte and water aggregates and to construct adjacency matrices for analyzing their network properties.

2.2. Osmolyte Aggregation Behavior and Radial Distribution Function Analysis. For each of the four systems, namely, pure osmolyte and osmolyte−water mixture, the RDF was calculated to obtain local structural information on osmolyte and water aggregates. The corresponding RDFs were plotted in Figure S1 for the Supporting Information and Figure 2, respectively. The RDF of pure urea in Figure S1 was obtained with two atoms, the O atom (−CO) and H atom ($-NH_2$), for the description of the H-bond interaction between neighboring urea molecules, and the site−site RDFs of the other three osmolytes were calculated by taking the central atom of each compound, S for DMSO (S_{Osm}-S_{Osm}), N for TMAO (N_{Osm}-N_{Osm}), and C for TMU (C_{Osm}-C_{Osm}), to determine the distance criteria for the formation of osmolyte aggregates. The RDFs of DMSO and TMU in their pure osmolyte systems show strong first and weak second peaks which are characteristic of a liquid state, while three prominent peaks appear in urea and TMAO, reflecting formation of more regularly packed molecular aggregates similar to a crystalline structure.[48]

In the case of osmolyte−water mixtures shown in Figure 2, all of the g(r) values converge to 1 at long-range distance region, which means that water and osmolyte molecules are randomly distributed and form a single-phase liquid in both concentrations, 5 and 10 m. The peak positions of the first maximum and first minimum in the RDFs exhibit no significant change with respect to the concentration of osmolyte molecules. As the osmolyte−water solution is formed at the two representative concentrations, the peak intensity of the osmolyte−osmolyte RDF in Figure 2 tends to remarkably decrease to around 1 in comparison with that of pure osmolyte in Figure S1, except for the relatively intense first peak of TMU, which represents a strong hydrophobic interaction between the TMU molecules. When a solution consisting of osmolyte and water is formed, the spatial distribution of component molecules is macroscopically uniform. However, from a microscopic perspective, the distribution is no longer uniform and varies with the molecular aggregation pattern. More specifically, as self-association of osmolyte occurs to reduce the interaction with water, component molecules are no longer evenly distributed, representing the microheterogeneity in the solution. The water−water RDF analysis of TMU−water mixtures exhibits an increase in the intensity of the first and second peaks, reflecting the enhanced H-bonding interaction between water molecules near osmolyte molecules.[31,69] Additionally, in the

https://doi.org/10.1021/acs.jcim.3c01382
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

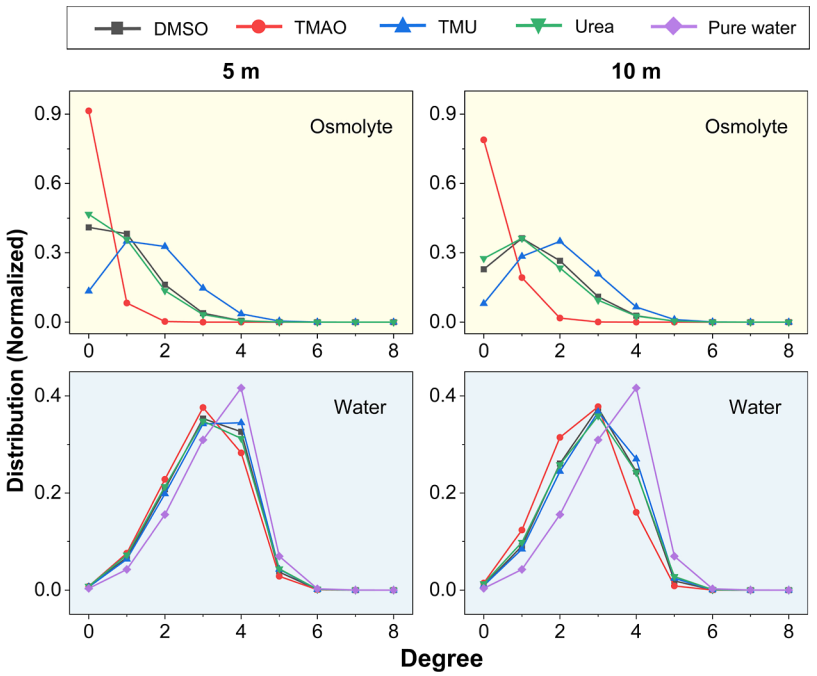


Figure 3. Normalized degree distribution of osmolyte and water molecules at 5 and 10 m with a constant temperature, 273 K. The top panel is the degree distribution of osmolyte and the bottom panel is that of water. TMAO has a maximum population at 0 in both concentrations, and TMU exhibits a peak shape with its position of 2 at 10 m. No significant changes in the H-bond network of water are observed against the type of osmolyte except for a slight decrease in water degree in TMAO−water mixtures.

RDF of the TMU solution, the long-range tail is seen in a distance region of 3 to 6 Å, indicating incomplete mixing of the components caused by the formation of TMU aggregates.

In the case of TMAO−water mixture, the intense first peak of osmolyte−water RDF in Figure 2 implies that the amphiphilic compound is almost completely hydrated due to the strong interaction of hydrophilic *N*-oxide group of TMAO with water molecules.[38] This distinct aggregation behavior of osmolyte molecules is also evidently displayed in Figure 1 for the four aqueous osmolyte mixtures. The three osmolyte molecules of DMSO, TMAO, and urea appear to be dispersed with the presence of small aggregates, whereas large osmolyte aggregates in TMU−water mixtures emerge even in the formation of macroscopically homogeneous solution.

As discussed in the previous work, the existence of molecular aggregates with different topological structures provokes an alteration of the water structure and ultimately a change in phase behavior in the cosolvent−water mixtures.[32] For instance, the *n*-butanol molecule at lower temperatures prefers the formation of self-associated alcohol aggregates by minimizing the interaction with water, and their development eventually generates a separation into two liquid phases, that is, water-rich and alcohol-rich. As the temperature increases, the self-associated butanol aggregates change into spatially extended H-bond aggregates that interact substantially with water. The expansion of the network-like butanol aggregates allows for coexistence with the water H-bond network, leading to the formation of a single-phase solution.[32,55]

Although the spatial inhomogeneity and aggregation behavior in a two-phase system of cosolvent−water mixtures were systematically studied, the issue of molecular aggregation pathway and its effect on microheterogeneity needs to be addressed in a one-phase system such as osmolyte−water solutions. In order to quantify the network properties of osmolyte and water aggregates, aggregate size distribution and graph theoretical analysis are to be performed. In addition, the spatial distribution of osmolyte and water, depending on molecular aggregation behavior, will also be quantitatively described using the *h*-value. The first minimum position for urea and water and the first maximum position for the other three osmolytes are used as a criterion to define the H-bond interaction[13,32,52] and osmolyte−osmolyte hydrophobic interaction,[32,58] respectively.

3. GRAPH THEORETICAL ANALYSIS AND AGGREGATE SIZE DISTRIBUTION OF OSMOLYTE SOLUTIONS

The graph theoretical representation of molecular aggregates provides important information on the morphological structure that depends on the composition, temperature, and pressure of binary liquid mixtures.[9,13,32,54,58] Recently, the graph theoretical analysis including degree distribution,[9,30,32,46,70] eigenvalue spectrum,[13,32,49,71] and Laplacian spectrum,[54,71] was widely used to investigate the network properties of solute and water aggregates in various aqueous mixture systems. In order to obtain graph theoretical properties of molecular aggregates, the adjacency matrix of the corresponding aggregates should be constructed from the graph representation, where the adjacency matrix of osmolyte and water aggregates is constructed separately. Graph $G{=}{=}G(V, E)$ consists of vertices V, representing each osmolyte molecule and edges E connecting two vertices, which corresponds to the water−water (osmolyte−osmolyte) interaction in water (osmolyte) aggregate. The edge in this work is considered an undirected one and the degree d_j is defined as the number of neighboring vertices around the jth vertex.[72] The adjacency matrix, A, is symmetric with the size of $N \times N$, where N is the number of osmolyte or water molecules in binary mixture systems. The values of adjacency matrix elements, A_{ij}, are determined according to the connectivity between vertices.[73]

https://doi.org/10.1021/acs.jcim.3c01382
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

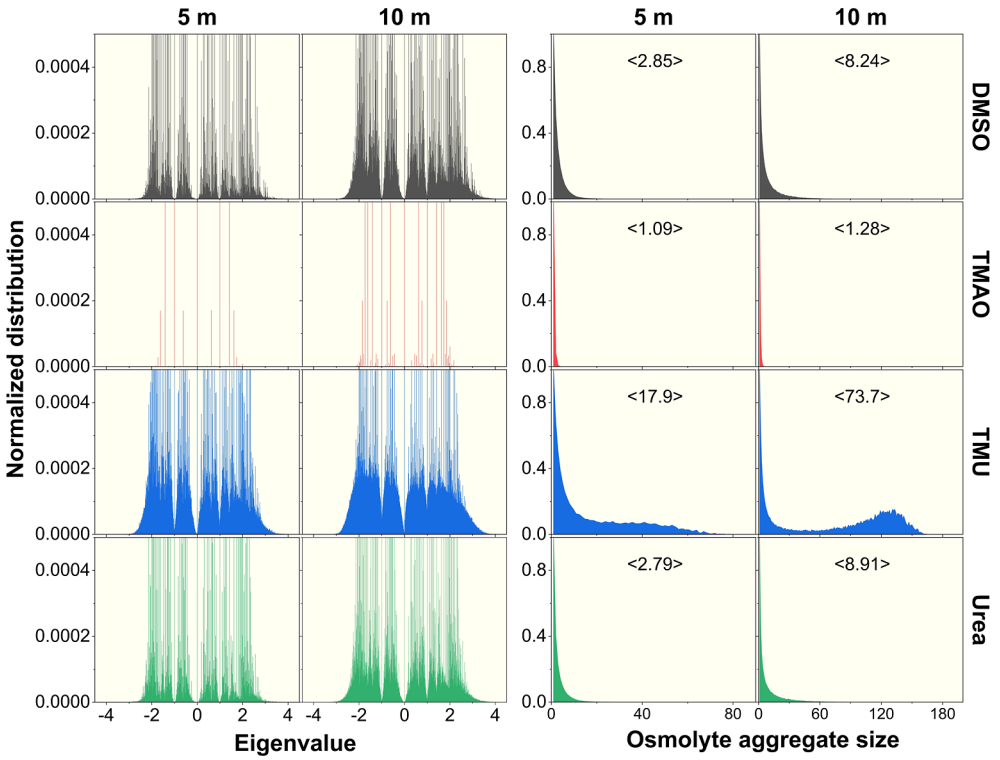


Figure 4. Normalized distribution of eigenvalue and aggregate size of osmolyte molecules for 5 and 10 m binary aqueous system at 273 K. TMU molecule exhibits a formation of large hydrophobic aggregates, while most of TMAO molecules exist as an isolated one with no osmolyte−osmolyte interaction.

A_{ij}=1 when i and j are connected by an edge

A_{ij}=0 when i and j are not connected

A_{ij}=0 when i equals to j

Here, the size of the adjacency matrix for water H-bond aggregates of pure water and osmolyte−water solution is given as 1000 × 1000, while that of osmolyte aggregates in aqueous solution corresponds to 90 × 90 and 180 × 180 for 5 and 10 m, respectively. Furthermore, by diagonalizing the adjacency matrix of the given aggregates, which was constructed with 10 000 atomic configurations extracted from respective MD trajectories, the eigenvalue spectrum is obtained in the four aqueous mixtures.

3.1. Graph Theoretical Analysis and Aggregate Size Distribution of Osmolyte Aggregates. In the construction of the adjacency matrix of osmolyte and water aggregates, the osmolyte−osmolyte distance criteria correspond to 5.40, 4.96, 6.10, and 2.85 Å, for DMSO, TMAO, TMU, and urea, respectively, and that of water−water H-bond interaction is 2.42 Å. The osmolyte and water degree distributions in binary aqueous solutions are plotted at the two concentrations of 5 and 10 m in Figure 3. The pattern of degree distribution at 10 m is overall similar to that of 5 m, except for a slight increase in average degree due to the addition of osmolyte molecules. However, the degree distribution behavior is found to be dependent upon the nature of the osmolyte, that is, TMAO exhibits a monotonous decay pattern, indicating that most of the TMAO molecules are completely solvated by water molecules without the aggregation of osmolyte.[38,42,43] In the case of TMU molecules, a peak appears at degree 2, resulting from the significant intermolecular interaction between TMU molecules,[31,39] and the other two osmotic molecules, urea and DMSO, exhibit an intermediate tendency between that of TMU and TMAO.

The distinct osmolyte aggregation behavior in the aqueous mixtures is also shown in the osmolyte aggregate size distribution and eigenvalue spectrum in Figure 4. As the concentration increases in the osmolyte mixtures, the average aggregate size tends to increase with the addition of osmolyte molecules, but the values of the average size at a given concentration clearly depend upon the type of osmolyte. The compound TMU self-associates readily and forms large osmolyte aggregates with an average size greater than 10 at both concentrations of 5 and 10 m, while the average size of TMAO is ∼1.3 even at the high concentration of 10 m, meaning nearly full hydration of the solute molecule. The aggregation propensity of DMSO and urea is moderate, with an average aggregate size of 8−9 at the 10 m concentration. As shown in the eigenvalue spectrum of osmolyte aggregates in Figure 4, the spectrum pattern reflects the osmolyte aggregation behavior in the binary mixtures. In the extreme case where all the osmolyte molecules are perfectly isolated ignoring osmolyte−osmolyte interaction, the eigenvalue should be 0, and for the aggregate in which two osmolyte molecules associate, the eigenvalue is assigned as ±1.[13,46,55] Discrete peaks appear with the formation of variously sized small aggregates, and the eigenvalue spectrum becomes broad and continuous as huge osmolyte aggregates are formed. As expected, the eigenvalue spectrum of TMU is comparatively wide and continuous in the existence of large osmolyte aggregates, but the peaks of TMAO appear discrete, with strong peak intensities at 0 and ±1, representing a high population of monomers and dimers. The spectral densities of TMU and TMAO are the highest and the lowest among the four osmolyte molecules, respectively, and those of DMSO and urea

https://doi.org/10.1021/acs.jcim.3c01382
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

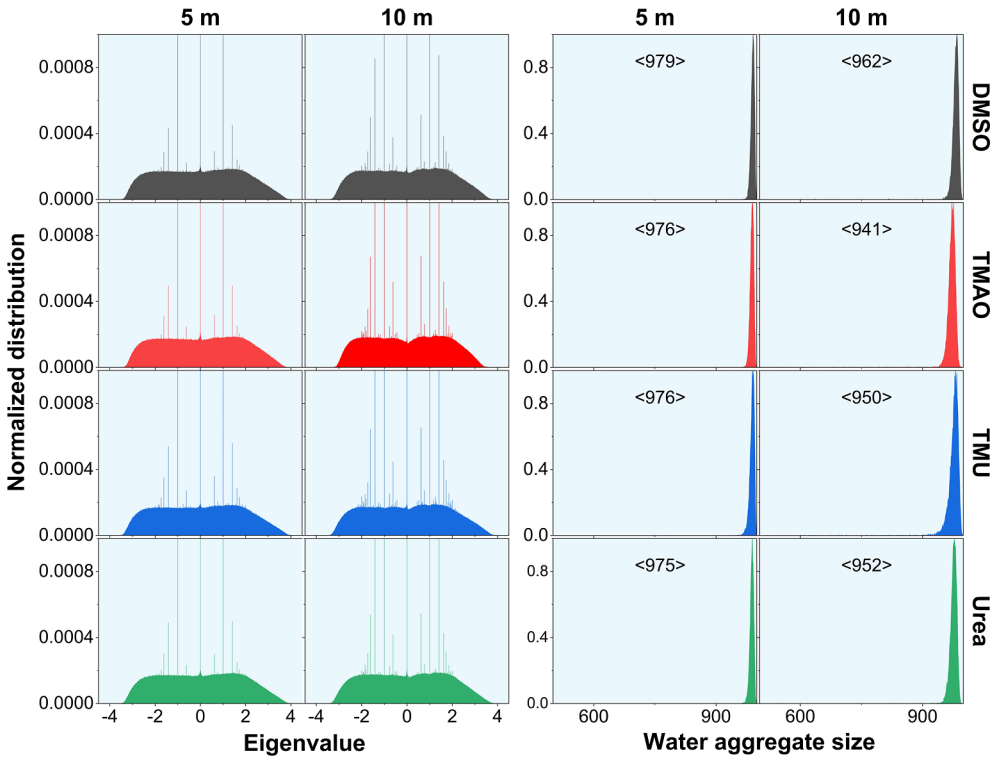


Figure 5. Normalized distribution of eigenvalue and aggregate size of water for 5 and 10 m binary aqueous system at 273 K. The water H-bond networks are almost maintained even at highly concentrated solutions. Most of the 1000 water molecules participate in the formation of a percolating water H-bond network.

are intermediate between those of TMU and TMAO osmolyte molecules. Again, this characteristic molecular aggregation is determined by the subtle balance of molecular interactions between osmolyte and water molecules, between water molecules, and between osmolyte molecules. The TMAO molecule with highly polar *N*-oxide exhibits preference in interaction with water, but the TMU molecule possessing four methyl groups prefers the formation of hydrophobic aggregates in the aqueous mixtures. The other two osmolytes of urea and DMSO are presumed to have generally an equal preference for the two interactions of osmolyte−water and osmolyte−osmolyte, resulting in the emergence of moderately sized aggregates. This distinct aggregation behavior does not appear in a pure osmolyte system, that is, the eigenvalue spectrum and aggregate size distribution are similar across all four osmolytes, forming a huge osmolyte aggregate, as shown in Figures S2 and S3, respectively.

Before closing this section, it is to be addressed that these graph theoretical quantities are not directly connected to the free energy of a given molecular system, despite the critical use in a description of the network properties of molecular aggregates. In recent several cases, such attempts to understand thermodynamic properties such as entropy and free energy in the use of graph theory were made.[74,75] The topological analysis in peptide systems was applied to probe the free energy surfaces of β-hairpin and three-stranded β-sheet and revealed existence of a highly heterogeneous denatured state ensemble including fluctuating non-native secondary structure.[74] Oliveira et al. employed the graph theoretical analysis such as degree distribution and average path length (L) to investigate the temperature effect on the hydrogen bond network of liquid water and found the coexistence of three H-bond components with distinct L at temperature ranges between 273 and 373 K.[75] They also estimated the standard enthalpy and entropy of equilibrium, considering structural transitions between these three sorts of networks.

3.2. Graph Theoretical Analysis and Aggregate Size Distribution of Water Aggregates. Despite the distinct osmolyte aggregation pattern in the binary mixtures, the network properties of water are not noticeably different with respect to the osmolyte type, as plotted in Figures 3 and 5. The peak of the degree distribution of water at 5 m is located between 3 and 4, and an increase in osmolyte concentration to 10 m causes a further down-shift of peak to 3. This denotes that the water H-bond network is interrupted in the presence of osmolyte aggregates, given that the pure water mainly has four water−water H-bonds, as plotted in the bottom panel of Figure 3. However, the extent of disruption in the water H-bond network is not severe at both 5 and 10 m concentrations, and there is no remarkable difference in the degree distribution against the nature of osmolyte, except for the slightly down-shifted pattern of TMAO in comparison with that of the other three osmolytes. The eigenvalue spectrum of water in Figure 5 appears to be wide and continuous at all four binary mixtures with two concentrations of 5 and 10 m, which implies the preservation of a huge water H-bond network regardless of the existence of sizable osmolyte aggregates. This robustness of the global water network is also confirmed in the water aggregate size distribution analysis by demonstrating that almost all water molecules participate in forming a giant H-bond aggregate.

Although the overall water H-bond structure does not differ significantly according to the type of osmolyte molecule, the local water structure in close proximity to the osmolyte appears to depend on the characteristic of the osmolyte. In use of O_{Osm}−H_W RDF (see middle panel of Figure 2), the water molecules adjacent to osmolyte are designated and their degree

https://doi.org/10.1021/acs.jcim.3c01382
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

distribution is displayed in Figure 6 at 5 and 10 m concentrations. The *y*-axis scale represents the ratio between

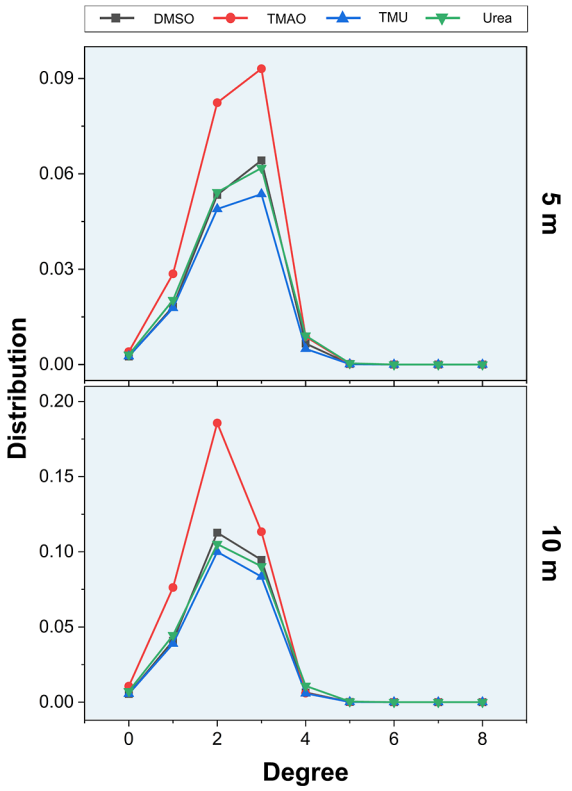


Figure 6. Degree distribution of water at the osmolyte/water interface for 5 and 10 m at 273 K. The water H-bond networks in the interfacial region show clear differences according to the nature of the osmolyte molecule. TMAO solutions have the highest number of water molecules that are adjacent to osmolyte molecules, while there are a relatively small number of water molecules at the osmolyte/water interface of TMU solutions.

the number of water molecules near the osmolyte and that of total water molecules. The peak position of specified water, which is adjacent to the osmolyte molecule, is red-shifted against the water degree distribution of the water−osmolyte mixture in Figure 3, and the peak height varies with respect to the nature of the osmolyte. It is manifested that local water structures surrounding osmolyte aggregates are affected by the solute aggregation behavior. The strongest peak intensity of TMAO in Figure 6 is associated with the preferential osmolyte−water interaction and TMU has the lowest intensity meaning less contact with water due to the favored osmolyte−osmolyte interaction.[38] Despite demonstrating how the osmolyte aggregation behavior promotes nonuniform distribution of constituent molecules by modulating local water structure near osmolyte aggregates, the attempts to measure quantitatively the spatial distribution of component molecules and to elucidate how the microheterogeneity of the mixture contributes to understanding osmolyte operating mechanism are not yet made.

4. SPATIAL DISTRIBUTION ANALYSIS AND MICROHETEROGENEITY IN BINARY OSMOLYTE SOLUTIONS

Several recent experimental[23,45,76] and computational studies[32,34,38,69] revealed the inhomogeneous distribution of solute molecules in aqueous binary mixtures and its effect on water structure and dynamics. Among the theoretical attempts to analyze the spatial distribution of constituent molecules, the *h*-value as a theoretical index developed to calculate the nonuniform distribution of nodes in 2D space,[77] has been successfully applied to three-dimensional (3D) metal alloy and binary liquid systems.[32,78,79] By estimating the spatial distribution of the cosolvent−water mixture with the use of the *h*-value, the effect of the cosolvent aggregation behavior on the spatial distribution of the aqueous mixture was examined.[32] For instance, methanol, which is completely soluble in liquid water, forms spatially extended H-bond aggregates in an aqueous alcohol solution with a small *h*-value close to 0, but water-insoluble DCM tends to form self-associated aggregates separated from water, exhibiting a relatively large *h*-value. This bifurcating aggregation pathway accompanies the distinct spatial distribution of constituent molecules and ultimately produces distinguishable phase behavior in binary solutions. Here, we will calculate the *h*-value to describe the spatially heterogeneous distribution of osmolyte and water molecules and demonstrate how their distinct aggregation behavior causes the microheterogeneity in the binary osmolyte solutions.

4.1. Calculation of *h*-Value as a Measurement of Spatial Inhomogeneity. The *h*-value in the 3D space is defined as follows[32,58,79]

$$h = \sum_{j=1}^{r} \omega^{1-j} \hat{h}(2^{3j}) = \frac{1}{2n} \sum_{j=1}^{r} \omega^{1-j} \max \sum_{i=1}^{2^{3j}} \left| m_i - \frac{n}{2^{3j}} \right| \quad (1)$$

The detailed procedure for obtaining the *h*-value was already given in the reference,[32,77] and a brief description is provided to help understand. The weighting factor ω of 8.887 is given so that the *h*-value has a value between 0 and 1, and the term $\hat{h}(2^{3j})$, written as $\frac{1}{2n}\sum_{i=1}^{2^{3j}} \left| m_i - \frac{n}{2^{3j}} \right|$, which quantifies the nonuniform spatial distribution of nodes by dividing the simulation box into sub-boxes (grids) of the same size. Suppose that each edge of a simulation box is divided into 2, 4, 8, ... 2^r segments, which produces 2^3, 4^3, 8^3, ... 2^{3r} sub-boxes, where the *r* is selected so that each sub-box includes at most one node. *n* is the total number of nodes in a whole simulation box, and m_i is the number of nodes in the *i*-th sub-box. The term $\frac{n}{2^{3j}}$ corresponds to the expectation value when all the nodes are uniformly distributed over all the sub-boxes. Thus, for the case of the ideal uniform distribution, the term $\hat{h}(2^{3j})$ should have a minimum value of zero because the value of m_i is identical to that of the term $\frac{n}{2^{3j}}$. In the opposite case, if all the nodes are positioned at the same point, the term $\hat{h}(2^{3j})$ becomes $1-\frac{1}{2^{3j}}$, when *j* goes to infinity, the term of $\hat{h}(2^{3j})$ has a convergence value of 1. In consideration of the weighting factor ω, the *h*-value with a value between 0 (the extreme delocalization) and 1 (the extreme localization) can be used to quantitatively describe the microscopic heterogeneity of components in various mixture systems.

To address the linear operation issue in the calculation of the *h*-value, which depends on the choice of the origin of the periodic box imposed in MD simulation,[32] the randomly selected 100 sites in the periodic box are given as the origin, and the maximum among the 100 values of $\hat{h}(2^{3j})$ obtained from each corresponding origin is taken when determining the *h*-value. For the four osmolyte molecules of DMSO, TMAO,

https://doi.org/10.1021/acs.jcim.3c01382
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

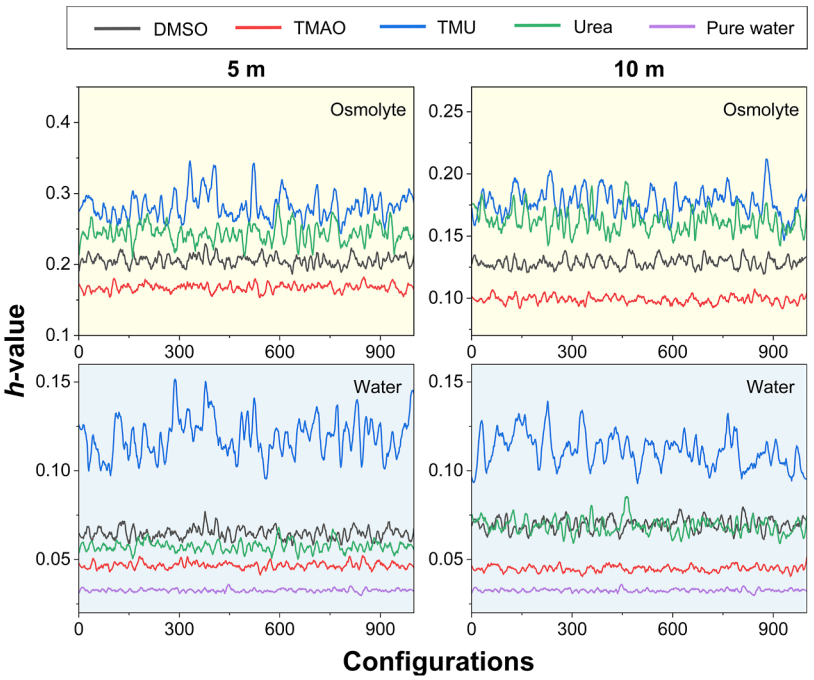


Figure 7. Time-varying *h*-value of osmolyte−water solutions at 273 K. Among four osmolyte−water binary aqueous systems, TMU solutions have relatively high *h*-values, indicating the existence of localized TMU hydrophobic aggregates. In contrast, TMAO solutions have the lowest *h*-value at both concentrations because most of the TMAO molecules do not make aggregates and are widely spread as monomers.

TMU, and urea, the central atom of each osmolyte is considered to represent a node, which corresponds to S, N, C, and C atoms, respectively. The node for the water molecule is described by taking the oxygen atom and time-varying *h*-values of each component in pure liquid water, and the binary mixtures are calculated with 1000 osmolyte−water configurations obtained from the corresponding MD trajectories.

4.2. Measurement of *h*-Value in Binary Aqueous Solutions. The time-fluctuating *h*-value and average *h*-value of each component in pure water and osmolyte−water solutions are denoted in Figures 7 and S4, respectively. As expected, in pure liquid water, the water molecules are uniformly distributed throughout the entire box, which generates a very small *h*-value close to 0. However, for the four osmolyte solutions, the time-dependent *h*-values clearly discriminate the microheterogeneity of each mixture, even though all four mixtures appear to be homogeneous solutions from a macroscopic perspective. For example, the spatial inhomogeneity of TMU aggregates is evident with the high *h*-values, while the *h*-values of the rarely aggregated TMAO molecules are very small, indicating a uniform distribution of the osmolyte molecules. This tendency for microheterogeneity in an osmolyte solution can be inferred from the O_{osm}−H_w RDF analysis in Figure 2. That is, the TMAO has a strong first peak intensity meaning significant interaction with water, which causes even mixing between the two components, while TMU shows weak peak intensity at the distance of 3−6 Å with a tendency to avoid interaction with water, forming large hydrophobic aggregates.

In the case of the other two osmolytes, DMSO and urea, an intermediate propensity between TMU and TMAO in the spatial distribution is observed with the existence of small osmolyte aggregates. Similar to the spatial distribution pattern of osmolyte aggregates, the *h*-value of water aggregates in TMU and TMAO mixtures has the largest and the smallest values, among the four osmolyte solutions, respectively. However, it is puzzling to understand that urea has significantly larger *h*-values than those of DMSO, as shown in the upper panel of Figure 7, considering that urea and DMSO are largely similar in the water−osmolyte interaction, osmolyte aggregate size, and network properties of osmolyte aggregates such as degree distribution and eigenvalue spectrum. This issue on spatial inhomogeneity measurement of osmolyte can be resolved by reflecting the volume of the corresponding aggregates as well as the network properties of the osmolyte aggregates.[58] More specifically, the occupied volumes of the two osmolyte aggregates are distinct, despite their similar aggregate sizes, which are 3 and 8 for DMSO and 3 and 9 for urea, for the concentrations of 5 and 10 m, respectively (see right panel of Figure 4). The occupied volume of urea aggregates is smaller than that of DMSO aggregates because urea forms tightly packed aggregates with strong urea−urea H-bond interaction, while DMSO aggregates possess loosely packed structure due to weak intermolecular interactions of DMSO−DMSO.

Assuming that the aggregation patterns of the two osmolytes are similar to each other, as the occupied volume of aggregates is increased, the *h*-value is decreased with evenly spreading across the entire box. This is why the urea, which has a smaller aggregate volume, has a relatively large *h*-value in comparison with that of DMSO. This volume effect on the spatial distribution of osmolyte exhibits the opposite behavior in the estimation of the *h*-value of water, that is, when the occupied volume of osmolyte aggregates is increased in osmolyte solutions, the spatial distribution of water molecules is relatively localized with higher *h*-value. As shown in the bottom panel of Figures 7 and S4, the *h*-value of water aggregates in the DMSO solution is slightly larger than that of the urea solution at the 5 m concentration. On the whole, the results obtained from the *h*-value calculation are consistent with the experimental observation that a significant microscopic heterogeneity appears in TMU solutions,[38,39,44] but no noticeable aggregation is found

https://doi.org/10.1021/acs.jcim.3c01382
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

in the TMAO−water mixtures.[38,40,41] More specifically, some calorimetric measurements and dynamic light scattering experiments in aqueous TMU solution revealed the presence of large TMU aggregates of several hundred nanometers in a lower mole fraction of 0.068.[80]

4.3. Microheterogeneity and Operating Mechanism of Osmolyte on Protein Stability. There are two hypotheses proposed to describe the operating mechanism of osmolytes: the direct mechanism, which focuses on the direct interaction between osmolyte and protein,[11,18,81] and the indirect mechanism, wherein the osmolyte regulates protein stability by changing the water properties such as H-bond network structure.[5,10,12,13,19] In this work, we attempt to elucidate how the osmolytes work to modulate the stability of proteins based on the microheterogeneity of the osmolyte mixtures obtained from the spatial inhomogeneity analysis. The two representative osmolytes of TMAO and TMU are amphiphiles containing polar and nonpolar groups and correspond to protecting and destabilizing osmolytes, respectively. The protecting osmolyte, TMAO, with low *h*-value exhibits little aggregation behavior while interacting preferentially with water, therefore, it is expected that there is seldom direct interaction of osmolyte with protein. Meanwhile, the destabilizing osmolyte TMU with a high *h*-value prefers the osmolyte−osmolyte interaction and generates noticeable microheterogeneity by forming osmolyte aggregates. This sort of intermolecular interaction between the amphiphiles is expected to enable the osmolyte−protein interaction because protein also contains hydrophilic and hydrophobic groups similar to the amphiphilic compound TMU. This interpretation of the operating mechanism inferred from the concept of microheterogeneity of the mixture system supports the direct mechanism explaining how osmolyte acts on protein stability. As mentioned before, the interaction between osmolyte and water also contributes to determining the spatial distribution of components in aqueous osmolyte solution. That is, the strong osmolyte−water interaction provokes the spreading of components into the entire space, but the heterogeneous distribution of components mainly originates from the self-association of osmolyte. In addition, it is evidently shown in Figure 6 that the local water structure around the osmolyte aggregates is altered with the type of osmolyte aggregates. In consideration of the interplay between osmolyte and water to describe the microheterogeneity in the binary mixture, the contribution of the second hypothesis is assessed to be essential for explaining the mode of osmolyte action to control the equilibrium between the unfolded and native state of the protein. In summary, the two hypotheses work in combination, complementing each other rather than being exclusive explanations.

5. CONCLUSIONS

The aggregation behavior of the osmolyte and its effect on water H-bond structures in the four osmolyte−water mixtures were examined with MD simulation and graph theoretical analysis. The protecting osmolyte, TMAO, does not aggregate with a favored interaction with water, whereas the denaturant TMU exhibits a strong propensity to promote molecular aggregation caused by the hydrophobic interaction between them. In the case of urea and DMSO, they form variously sized osmolyte aggregates in binary aqueous solutions, although the size is not as large as that of TMU. The distinct osmolyte aggregates do have little discriminatory effect on global water structure, but the local structure of water near the osmolyte varies significantly with respect to the aggregation type. The increase in self-association between osmolytes facilitates the microheterogeneity of osmolyte−water mixture, which was shown by calculating the time-varying *h*-value from corresponding MD trajectories in the solution system. Among the binary mixtures of four osmolyte molecules, the destabilizing agent, TMU, has the largest *h*-value with remarkable spatial inhomogeneity of components and the smallest *h*-value appears in the aqueous solution of protecting osmolyte, TMAO, indicating uniform distribution of constituent molecules. Based upon graph theoretical analysis and evaluation of microheterogeneity of osmolyte−water mixtures, it is proposed that both hypotheses of direct mechanism, emphasizing the direct interaction between osmolyte and protein, and indirect mechanism, through modulation of water structure, contribute to the stability of the protein in the presence of osmolyte.

■ ASSOCIATED CONTENT

Data Availability Statement

The four osmolyte molecular structures are available in the Supporting Information. MD simulations are performed with AMBER 20 and the representative snapshots are generated with PyMOL, an open-source molecular visualization software.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.3c01382.

Figures of RDF, eigenvalue spectrum, aggregate size distribution in pure osmolytes, and average *h*-value of osmolyte and water molecules in osmolyte−water solutions (PDF)

Osmolyte molecular structure of DMSO (PDB)

Osmolyte molecular structure of TMAO (PDB)

Osmolyte molecular structure of TMU (PDB)

Osmolyte molecular structure of urea (PDB)

■ AUTHOR INFORMATION

Corresponding Author

Jun-Ho Choi − *Department of Chemistry, Gwangju Institute of Science and Technology (GIST), Gwangju 61005, Republic of Korea;* orcid.org/0000-0001-5237-5566; Email: junhochoi@gist.ac.kr

Authors

Jiwon Seo − *Department of Chemistry, Gwangju Institute of Science and Technology (GIST), Gwangju 61005, Republic of Korea;* orcid.org/0000-0003-0085-9774

Ravi Singh − *Department of Chemistry, Gwangju Institute of Science and Technology (GIST), Gwangju 61005, Republic of Korea*

Jonghyuk Ryu − *Department of Chemistry, Gwangju Institute of Science and Technology (GIST), Gwangju 61005, Republic of Korea;* orcid.org/0009-0008-7868-2693

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.3c01382

Author Contributions

J.-H.C.: conceptualization, supervision, resources. J.S.: carried out MD simulations with spatial inhomogeneity analysis and graph theoretical analysis, visualization of data. J.S., R.S., J.R., and J.-H.C.: interpretation of the data, validation of results, contribution to the scientific discussions, manuscript writing, and proofreading of the manuscript.

https://doi.org/10.1021/acs.jcim.3c01382
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MEST) (grant nos. 2018R1D1A1B07042015 and 2023R1A2C1002559) and the Technology Innovation Program Development of Core Technologies in Carbon Neutral Industries (RS-2023-00262743, development of GWP 150 or lower alternative gas and process technology for chemical vapor deposition chamber cleaning process for display TFT gate insulator film) funded by the Ministry of Trade, Industry and Energy (MOTIE, Korea).

■ REFERENCES

(1) Zhang, Y.; Cremer, P. S. Chemistry of Hofmeister anions and osmolytes. *Annu. Rev. Phys. Chem.* **2010**, *61*, 63−83.
(2) Towey, J. J.; Soper, A. K.; Dougan, L. What happens to the structure of water in cryoprotectant solutions? *Faraday Discuss.* **2013**, *167*, 159−176.
(3) Stasiulewicz, M.; Panuszko, A.; Bruździak, P.; Stangret, J. Mechanism of Osmolyte Stabilization-Destabilization of Proteins: Experimental Evidence. *J. Phys. Chem. B* **2022**, *126*, 2990−2999.
(4) Street, T. O.; Bolen, D. W.; Rose, G. D. A molecular mechanism for osmolyte-induced protein stability. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 13997−14002.
(5) Bruździak, P.; Panuszko, A.; Stangret, J. Influence of osmolytes on protein and water structure: a step to understanding the mechanism of protein stabilization. *J. Phys. Chem. B* **2013**, *117*, 11502−11508.
(6) Ma, J.; Pazos, I. M.; Gai, F. Microscopic insights into the protein-stabilizing effect of trimethylamine N-oxide (TMAO). *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 8476−8481.
(7) Senske, M.; Törk, L.; Born, B.; Havenith, M.; Herrmann, C.; Ebbinghaus, S. Protein stabilization by macromolecular crowding through enthalpy rather than entropy. *J. Am. Chem. Soc.* **2014**, *136*, 9036−9041.
(8) Pace, C. N.; Tanford, C. Thermodynamics of the unfolding of β-lactoglobulin A in aqueous urea solutions between 5 and 55. *Biochemistry* **1968**, *7*, 198−208.
(9) Sundar, S.; Sandilya, A. A.; Priya, M. H. Unraveling the influence of osmolytes on water hydrogen-bond network: From local structure to graph theory analysis. *J. Chem. Inf. Model.* **2021**, *61*, 3927−3944.
(10) Wei, H.; Fan, Y.; Gao, Y. Q. Effects of urea, tetramethyl urea, and trimethylamine N-oxide on aqueous solution structure and solvation of protein backbones: a molecular dynamics simulation study. *J. Phys. Chem. B* **2010**, *114*, 557−568.
(11) Rodriguez-Ropero, F.; van der Vegt, N. F. Direct osmolyte-macromolecule interactions confer entropic stability to folded states. *J. Phys. Chem. B* **2014**, *118*, 7327−7334.
(12) Bennion, B. J.; Daggett, V. The molecular basis for the chemical denaturation of proteins by urea. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 5142−5147.
(13) Lee, H.; Choi, J.-H.; Verma, P. K.; Cho, M. Spectral graph analyses of water hydrogen-bonding network and osmolyte aggregate structures in osmolyte-water solutions. *J. Phys. Chem. B* **2015**, *119*, 14402−14412.
(14) Adamczak, B.; Wieczór, M.; Kogut, M.; Stangret, J.; Czub, J. Molecular basis of the osmolyte effect on protein stability: a lesson from the mechanical unfolding of lysozyme. *Biochem. J.* **2016**, *473*, 3705−3724.
(15) Canchi, D. R.; Jayasimha, P.; Rau, D. C.; Makhatadze, G. I.; Garcia, A. E. Molecular mechanism for the preferential exclusion of TMAO from protein surfaces. *J. Phys. Chem. B* **2012**, *116*, 12095−12104.
(16) Lin, T. Y.; Timasheff, S. N. On the role of surface tension in the stabilization of globular proteins. *Protein Sci.* **1996**, *5*, 372−381.
(17) Bruździak, P.; Adamczak, B.; Kaczkowska, E.; Czub, J.; Stangret, J. Are stabilizing osmolytes preferentially excluded from the protein surface? FTIR and MD studies. *Phys. Chem. Chem. Phys.* **2015**, *17*, 23155−23164.
(18) Auton, M.; Bolen, D. W.; Rösgen, J. Structural thermodynamics of protein preferential solvation: osmolyte solvation of proteins, aminoacids, and peptides. *Proteins: Struct., Funct., Bioinf.* **2008**, *73*, 802−813.
(19) Bennion, B. J.; Daggett, V. Counteraction of urea-induced protein denaturation by trimethylamine N-oxide: a chemical chaperone at atomic resolution. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 6433−6438.
(20) Wei, H.; Yang, L.; Gao, Y. Q. Mutation of charged residues to neutral ones accelerates urea denaturation of HP-35. *J. Phys. Chem. B* **2010**, *114*, 11820−11826.
(21) Wakisaka, A.; Matsuura, K. Microheterogeneity of ethanol-water binary mixtures observed at the cluster level. *J. Mol. Liq.* **2006**, *129*, 25−32.
(22) Tomza, P.; Czarnecki, M. A. Microheterogeneity in binary mixtures of propyl alcohols with water: NIR spectroscopic, two-dimensional correlation and multivariate curve resolution study. *J. Mol. Liq.* **2015**, *209*, 115−120.
(23) Chakraborty, S.; Pyne, P.; Mitra, R. K.; Das Mahanta, D. Hydrogen bond structure and associated dynamics in micro-heterogeneous and in phase separated alcohol-water binary mixtures: A THz spectroscopic investigation. *J. Mol. Liq.* **2023**, *382*, 121998.
(24) Takamuku, T.; Tabata, M.; Yamaguchi, A.; Nishimoto, J.; Kumamoto, M.; Wakita, H.; Yamaguchi, T. Liquid structure of acetonitrile- water mixtures by X-ray diffraction and infrared spectroscopy. *J. Phys. Chem. B* **1998**, *102*, 8880−8888.
(25) Egashira, K.; Nishi, N. Low-frequency Raman spectroscopy of ethanol- water binary solution: evidence for self-association of solute and solvent molecules. *J. Phys. Chem. B* **1998**, *102*, 4054−4057.
(26) Shin, D. N.; Wijnen, J. W.; Engberts, J. B.; Wakisaka, A. On the origin of microheterogeneity: a mass spectrometric study of dimethyl sulfoxide- water binary mixture. *J. Phys. Chem. B* **2001**, *105*, 6759−6762.
(27) Dixit, S.; Crain, J.; Poon, W.; Finney, J. L.; Soper, A. K. Molecular segregation observed in a concentrated alcohol-water solution. *Nature* **2002**, *416*, 829−832.
(28) Gupta, R.; Patey, G. Association and microheterogeneity in aqueous 2-butoxyethanol solutions. *J. Phys. Chem. B* **2011**, *115*, 15323−15331.
(29) Jalili, S.; Akhavan, M. Molecular dynamics simulation study of association in trifluoroethanol/water mixtures. *J. Comput. Chem.* **2010**, *31*, 286−294.
(30) Zhang, N.; Li, W.; Chen, C.; Zuo, J. Molecular dynamics simulation of aggregation in dimethyl sulfoxide-water binary mixture. *Comput. Theor. Chem.* **2013**, *1017*, 126−135.
(31) Gupta, R.; Patey, G. Structure and aggregation in model tetramethylurea solutions. *J. Chem. Phys.* **2014**, *141*, 064502.
(32) Seo, J.; Choi, S.; Singh, R.; Choi, J.-H. Spatial inhomogeneity and molecular aggregation behavior in aqueous binary liquid mixtures. *J. Mol. Liq.* **2023**, *369*, 120949.
(33) Luzar, A.; Chandler, D. Structure and hydrogen bond dynamics of water-dimethyl sulfoxide mixtures by computer simulations. *J. Chem. Phys.* **1993**, *98*, 8160−8173.
(34) Borin, I. A.; Skaf, M. S. Molecular association between water and dimethyl sulfoxide in solution: A molecular dynamics simulation study. *J. Chem. Phys.* **1999**, *110*, 6412−6420.
(35) Požar, M.; Lovrinčević, B.; Zoranić, L.; Primorać, T.; Sokolić, F.; Perera, A. Micro-heterogeneity versus clustering in binary mixtures of ethanol with water or alkanes. *Phys. Chem. Chem. Phys.* **2016**, *18*, 23971−23979.
(36) Idrissi, A.; Damay, P.; Yukichi, K.; Jedlovszky, P. Self-association of urea in aqueous solutions: A Voronoi polyhedron analysis study. *J. Chem. Phys.* **2008**, *129*, 164512.

https://doi.org/10.1021/acs.jcim.3c01382
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

(37) Stumpe, M. C.; Grubmüller, H. Aqueous urea solutions: structure, energetics, and urea aggregation. *J. Phys. Chem. B* **2007**, *111*, 6220−6228.
(38) Stirnemann, G.; Sterpone, F.; Laage, D. Dynamics of water in concentrated solutions of amphiphiles: key roles of local structure and aggregation. *J. Phys. Chem. B* **2011**, *115*, 3254−3262.
(39) Almásy, L.; Len, A.; Székely, N. K.; Pleštil, J. Solute aggregation in dilute aqueous solutions of tetramethylurea. *Fluid Phase Equilib.* **2007**, *257*, 114−119.
(40) Freda, M.; Onori, G.; Santucci, A. Hydrophobic hydration and hydrophobic interaction in aqueous solutions of tert-butyl alcohol and trimethylamine-N-oxide: a correlation with the effect of these two solutes on the micellization process. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4979−4984.
(41) Freda, M.; Onori, G.; Santucci, A. Infrared study of the hydrophobic hydration and hydrophobic interactions in aqueous solutions of tert-butyl alcohol and trimethylamine-N-oxide. *J. Phys. Chem. B* **2001**, *105*, 12714−12718.
(42) Fornili, A.; Civera, M.; Sironi, M.; Fornili, S. L. Molecular dynamics simulation of aqueous solutions of trimethylamine-N-oxide and tert-butyl alcohol. *Phys. Chem. Chem. Phys.* **2003**, *5*, 4905−4910.
(43) Paul, S.; Patey, G. Why tert-butyl alcohol associates in aqueous solution but trimethylamine-N-oxide does not. *J. Phys. Chem. B* **2006**, *110*, 10514−10518.
(44) Holz, M.; Grunder, R.; Sacco, A.; Meleleo, A. Nuclear magnetic resonance study of self-association of small hydrophobic solutes in water: salt effects and the lyotropic series. *J. Chem. Soc., Faraday Trans.* **1993**, *89*, 1215−1222.
(45) Oh, K.-I.; You, X.; Flanagan, J. C.; Baiz, C. R. Liquid-liquid phase separation produces fast H-bond dynamics in DMSO-water mixtures. *J. Phys. Chem. Lett.* **2020**, *11*, 1903−1908.
(46) Choi, J.-H.; Lee, H.; Choi, H. R.; Cho, M. Graph theory and ion and molecular aggregation in aqueous solutions. *Annu. Rev. Phys. Chem.* **2018**, *69*, 125−149.
(47) Han, K.; Venable, R. M.; Bryant, A.-M.; Legacy, C. J.; Shen, R.; Li, H.; Roux, B.; Gericke, A.; Pastor, R. W. Graph-theoretic analysis of monomethyl phosphate clustering in ionic solutions. *J. Phys. Chem. B* **2018**, *122*, 1484−1494.
(48) Choi, J.-H.; Cho, M. Ion aggregation in high salt solutions. IV. Graph-theoretical analyses of ion aggregate structure and water hydrogen bonding network. *J. Chem. Phys.* **2015**, *143*, 104110.
(49) Choi, J.-H.; Cho, M. Ion aggregation in high salt solutions. II. Spectral graph analysis of water hydrogen-bonding network and ion aggregate structures. *J. Chem. Phys.* **2014**, *141*, 154502.
(50) Anand, D. V.; Meng, Z.; Xia, K.; Mu, Y. Weighted persistent homology for osmolyte molecular aggregation and hydrogen-bonding network analysis. *Sci. Rep.* **2020**, *10*, 9685.
(51) Bakó, I.; Bencsura, Á.; Hermannson, K.; Bálint, S.; Grósz, T.; Chihaia, V.; Oláh, J. Hydrogen bond network topology in liquid water and methanol: a graph theory approach. *Phys. Chem. Chem. Phys.* **2013**, *15*, 15163−15171.
(52) Choi, S.; Parameswaran, S.; Choi, J.-H. Understanding alcohol aggregates and the water hydrogen bond network towards miscibility in alcohol solutions: graph theoretical analysis. *Phys. Chem. Chem. Phys.* **2020**, *22*, 17181−17195.
(53) Kartha, T. R.; Madhurima, V. Hydrogen-bonded networks in alcohol-acetone binary mixtures: molecular dynamics study. *J. Mol. Model.* **2022**, *28*, 382.
(54) Pothoczki, S.; Pethes, I.; Pusztai, L.; Temleitner, L.; Ohara, K.; Bakó, I. Properties of Hydrogen-Bonded Networks in Ethanol-Water Liquid Mixtures as a Function of Temperature: Diffraction Experiments and Computer Simulations. *J. Phys. Chem. B* **2021**, *125*, 6272−6279.
(55) Parameswaran, S.; Choi, S.; Choi, J.-H. Temperature effects on alcohol aggregation phenomena and phase behavior in n-butanol aqueous solution. *J. Mol. Liq.* **2022**, *347*, 118339.
(56) Choi, S.; Parameswaran, S.; Choi, J.-H. Effects of molecular shape on alcohol aggregation and water hydrogen bond network behavior in butanol isomer solutions. *Phys. Chem. Chem. Phys.* **2021**, *23*, 12976−12987.
(57) Bakó, I.; Csókás, D.; Pothoczki, S. Molecular aggregation in liquid water: Laplace spectra and spectral clustering of H-bonded network. *J. Mol. Liq.* **2021**, *327*, 114802.
(58) Choi, S.; Seo, J.; Singh, R.; Choi, J.-H. Pressure dependence of molecular aggregation and phase behavior in methane-water mixtures. *J. Mol. Liq.* **2023**, *383*, 122123.
(59) Vishveshwara, S.; Brinda, K.; Kannan, N. Protein structure: insights from graph theory. *J. Theor. Comput. Chem.* **2002**, *01*, 187−211.
(60) Canchi, D. R.; García, A. E. Cosolvent effects on protein stability. *Annu. Rev. Phys. Chem.* **2013**, *64*, 273−293.
(61) Herskovits, T. T. Nonaqueous solutions of DNA; denaturation by urea and its methyl derivatives. *Biochemistry* **1963**, *2*, 335−340.
(62) Pace, C. N.; Marshall, H. F. A comparison of the effectiveness of protein denaturants for β-lactoglobulin and ribonuclease. *Arch. Biochem. Biophys.* **1980**, *199*, 270−276.
(63) Ashwood Smith, M. Radioprotective and cryoprotective properties of dimethyl sulfoxide in cellular systems. *Ann. N.Y. Acad. Sci.* **1967**, *141*, 45−62.
(64) Rajeshwara, A.; Prakash, V. Structural stability of lipase from wheat germ. *Int. J. Pept. Protein Res.* **1994**, *44*, 435−440.
(65) Sterling, H. J.; Prell, J. S.; Cassou, C. A.; Williams, E. R. Protein conformation and supercharging with DMSO from aqueous solution. *J. Am. Soc. Mass Spectrom.* **2011**, *22*, 1178.
(66) Martínez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.* **2009**, *30*, 2157−2164.
(67) Case, D. A.; Belfon, K.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, III, T. E.; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Giambasu, G.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Harris, R.; Izadi, S.; Izmailov, S. A.; Kasavajhala, K.; Kovalenko, A.; Krasny, R.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Man, V.; Merz, K. M.; Miao, Y.; Mikhailovskii, O.; Monard, G.; Nguyen, H.; Onufriev, A.; Pan, F.; Pantano, S.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Skrynnikov, N. R.; Smith, J.; Swails, J.; Walker, R. C.; Wang, J.; Wilson, L.; Wolf, R. M.; Wu, X.; Xiong, Y.; Xue, Y.; York, D. M.; Kollman, P. A. *AMBER 2020*; University of California: San Francisco, 2020.
(68) York, D. M.; Darden, T. A.; Pedersen, L. G. The effect of long range electrostatic interactions in simulations of macromolecular crystals: A comparison of the Ewald and truncated list methods. *J. Chem. Phys.* **1993**, *99*, 8345−8348.
(69) Indra, S.; Biswas, R. Is dynamic heterogeneity of water in presence of a protein denaturing agent different from that in presence of a protein stabilizer? A molecular dynamics simulation study. *J. Chem. Sci.* **2016**, *128*, 1943−1954.
(70) Bakó, I.; Megyes, T.; Bálint, S.; Grósz, T.; Chihaia, V. Water-methanol mixtures: topology of hydrogen bonded network. *Phys. Chem. Chem. Phys.* **2008**, *10*, 5004−5011.
(71) Mikalčiūtė, A.; Vilčiauskas, L. Insights into the hydrogen bond network topology of phosphoric acid and water systems. *Phys. Chem. Chem. Phys.* **2021**, *23*, 6213−6224.
(72) Biggs, N.; Biggs, N. L.; Norman, B. *Algebraic Graph Theory*; Cambridge University Press, 1993.
(73) Godsil, C.; Royle, G. F. *Algebraic Graph Theory*; Springer Science & Business Media, 2001.
(74) Caflisch, A. Network and graph analyses of folding free energy surfaces. *Curr. Opin. Struct. Biol.* **2006**, *16*, 71−78.
(75) de Oliveira, P. M.; de Souza, J. I.; da Silva, J. A.; Longo, R. L. Temperature Dependence of Hydrogen Bond Networks of Liquid Water: Thermodynamic Properties and Structural Heterogeneity from Topological Descriptors. *J. Phys. Chem. B* **2023**, *127*, 2250−2257.
(76) Yang, B.; Cao, X.; Wang, C.; Wang, S.; Sun, C. Investigation of hydrogen bonding in Water/DMSO binary mixtures by Raman spectroscopy. *Spectrochim. Acta, Part A* **2020**, *228*, 117704.
(77) Schilcher, U.; Gyarmati, M.; Bettstetter, C.; Chung, Y. W.; Kim, Y. H. Measuring Inhomogeneity in Spatial Distributions. *VTC Spring 2008-IEEE Vehicular Technology Conference*; IEEE, 2008; pp 2690−2694.

https://doi.org/10.1021/acs.jcim.3c01382
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

(78) Kim, H.-K.; Ahn, J.-P.; Lee, B.-J.; Park, K.-W.; Lee, J.-C. Role of atomic-scale chemical heterogeneities in improving the plasticity of Cu-Zr-Ag bulk amorphous alloys. *Acta Mater.* **2018**, *157*, 209−217.
(79) Kim, H.-K.; Lee, M.; Lee, K.-R.; Lee, J.-C. How can a minor element added to a binary amorphous alloy simultaneously improve the plasticity and glass-forming ability? *Acta Mater.* **2013**, *61*, 6597−6608.
(80) Kustov, A. V.; Smirnova, N. L. Thermodynamics of TMU-TMU interaction in water, ethylene glycol and formamide-From pair solvophobic interaction to cluster formation. *J. Mol. Liq.* **2022**, *358*, 119185.
(81) Stumpe, M. C.; Grubmüller, H. Interaction of urea with amino acids: implications for urea-induced protein denaturation. *J. Am. Chem. Soc.* **2007**, *129*, 16126−16131.

https://doi.org/10.1021/acs.jcim.3c01382
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX