Local Water Structures Govern the Mixing Thermodynamics of Glycerol-Water Solutions

Debasish Das Mahanta[1,2], Dennis Robinson Brown[3], Simone Pezzotti[1], Songi Han[4,5], Gerhard Schwaab[1], M. Scott Shell[3,*], Martina Havenith[1,2,*]

[1]Lehrstuhl für Physikalische Chemie II, Ruhr-Universität Bochum, 44780 Bochum, Germany

[2]Department of Physics, Technische Universität Dortmund, 44227 Dortmund, Germany

[3]Department of Chemical Engineering, University of California, Santa Barbara, California 93106-5080, United States

[4]Department of Chemical Engineering, University of California, Santa Barbara, California 93106-5080, United States

[5]Department of Chemistry and Biochemistry, University of California, Santa Barbara, California 93106-9510, United States

*Corresponding author

Email: shell@ucsb.edu, Martina.Havenith@ruhr-uni-bochum.de

Abstract

Glycerol is a major cryoprotective agent and is widely used to promote protein stabilization. Through a combined experimental and theoretical study, we show that global thermodynamic mixing properties of glycerol and water are dictated by local solvation motifs. We identify three water populations, i.e., bulk water, bound water H-bonded to hydrophilic groups of glycerol and wrap water hydrating hydrophobic moieties. Each population provides distinct spectroscopic fingerprints in the THz/FIR spectral range, which allow to quantify their respective abundance and their partial contributions to the mixing enthalpy. We uncover a 1:1 connection between the number of bound waters and the mixing enthalpy, as deduced from experiments as well as from simulations. The balance between local hydrophobic wrap and hydrophilic bound contributions at the molecular level dictates macroscopic thermodynamics of mixing. This offers opportunities to rationally design polyol water mixtures to optimize technological applications by tuning mixing enthalpy and entropy based on spectroscopic screening.

Glycerol (1,2,3-propanetriol), a simple flexible polyol containing three hydroxyl (OH) groups, is a ubiquitously used co-solvent. For instance, it is a major cryoprotective agent (CPA),[1-4] widely used to promote protein stabilization through preferential solvation.[5-7] In cryobiology, the roles of water are strongly modulated by the specific interactions with the CPA and can be categorized into three categories: water released from the cell during freezing, water participating in intracellular ice formation and growth, and water bound to CPAs (e.g., glycerol molecules) and cytoplasmic proteins through hydrogen bonds (HBs) or electrostatic interactions that may hinder ice formation. The delicate balance of these water populations is dictated by the nature of the CPA and its mixing ratio with water. Molecular dynamics (MD) simulation studies can provide a molecular understanding of the underlying complex interplay of inter-molecular forces.[8] Among the three water populations, the one bound to CPAs and proteins is the least understood, owing to the complexity of resolving the molecular details of the CPA-water interactions.[8] Nonetheless, these interactions are key in determining the amount of water and CPA molecules in the protein first hydration shells and the balance between the three water populations in the mixture. These ultimately modulate the ability of water-CPA mixtures to preserve biological functionality by cryoprotection. The amount of water and CPA in the protein solvation shells was recently investigated by Chéron et al., who proposed that it is dictated by entropic effects related to volume exclusion and protein surface accessibility (i.e., to the size of the CPA) and by enthalpic effects due to H-Bonding.[6] Weng et al. further proposed

that the population of H-bonded water to various CPA polyols have a minor temperature dependence, but strongly vary with the nature of the polyol and its mole fraction.[8] These studies suggest that water-CPA HB interactions, which are enthalpically favored, regulate the balance between released and bound water populations, hence impacting the different cryoprotection behavior of glycerol compared to other polyols, such as ethylene glycol.

Despite intense interest and many compelling studies on the subject matter, our understanding of the roles of water and CPA in cryobiology is to date mostly based on mean field thermodynamic and kinetic theories,[8-10] and the glycerol-water mixture composition is generally adjusted heuristically to tune macroscopic properties such as viscosity. However, macroscopic and solvation properties cannot be predicted or extrapolated from CPA composition, yet, the structural, thermodynamical and solvation properties of even the simple CPA consisting of a glycerol:water mixture strongly depends on the mixing ratio and are subtly interconnected.[4, 6, 11-15] Understanding these interconnections at the molecular level is a hitherto unmet challenge for both theory and experiments that require reconciling local and macroscopic properties, and it is the missing link to rationally optimizing the fraction of glycerol in solvent mixtures used in technological applications, including cryopreservation. Moreover, owing to their simple molecular structure and complete miscibility, glycerol-water mixtures have great value as a model system to study the structural heterogeneity in binary mixtures and its connection to their thermodynamic properties. Compared to more hydrophobic model systems, such as alcohol-water mixtures, the higher number of OH moieties in the glycerol molecule (3 vs 1) offers an excellent case to investigate the impact of hydrophilic interactions on mixing thermodynamics. In glycerol-water mixtures, hydrophilic interactions are expected to outweigh contributions from hydrophobic solvation.[16]

Given its importance in applications and in fundamental solvation science, the structure and dynamics of glycerol aqueous solutions have been extensively investigated with various experimental tools, such as dielectric spectroscopy,[13, 17-21] nuclear magnetic resonance,[22] X-ray photoelectron spectroscopy,[17] infrared (IR) spectroscopy,[4, 17, 23-24] as well as with computer simulations techniques.[1, 25-26] These studies provided valuable insight on the molecular structural changes that occur with changing glycerol content. For instance, Feldman et al. reported that water-water interactions overwhelm the water-glycerol and glycerol-glycerol interactions in glycerol-water mixtures with less than 55% glycerol content.[20-21] In another study, Ahmed et al. further showed that the HB-network of water adopts three distinct states with increasing glycerol concentration: one in which bulk-like water dominates and solvates glycerol molecules at high water content, a second one containing confined water populations

and a third state dominated by bulk glycerol at high glycerol content.[17] In light of these results, it is still unknown whether and how such structural transformations of a solvent mixture dictate the mixing thermodynamics of glycerol and water, and by extension the thermodynamics of solutes and their interactions in these mixtures. Answering such fundamental questions is the basis for tailoring glycerol-water interactions to optimize cryoprotective and solvation properties, e.g., regulating the balance between released and bound water populations in cryoprotection and the composition of the solvation shells of biomolecules.

Terahertz (THz) absorption spectroscopy has the potential to address this challenge, since it directly probes the collective motions of water and that of the solvation shells, which allow characterizing the complexity of the water-glycerol network in combination with MD simulations.[17, 27-30] In a recent study of hydrated alcohols, we identified spectroscopic fingerprints in the THz/FIR spectral range that directly probe the distinct structural motifs of HB-network around alcohol solutes.[27] This study showed that water populations close to the hydrophilic and hydrophobic moieties of the alcohol provide two distinct THz/FIR fingerprints that are well distinguished from that of bulk water. These two populations are illustrated in figure 1 for the case of hydrated glycerol. On the one hand, water molecules around hydrophobic (CH_{1-4}) groups are not H-bonded to the solute and form a water-water HB-network wrapped around the hydrophobic surface. These "wrap" water molecules provide their characteristic THz-signature in the 100-250 cm^{-1} (3-7.5 THz) spectral region of HB-stretching modes, due to the slightly weaker water-water HBs formed within the wrap as compared to bulk water, which causes a red-shift of the inter-molecular stretch.[31] The more the water network is perturbed to accommodate the solute, the more this fingerprint is evident in the THz/FIR spectrum. Therefore, the wrap population provides a measure of how good/bad a solute molecule fits in the water network, and it is expected to disappear for solutes that do not perturb the water network (such as urea).[32] On the other hand, "bound" water molecules H-bonded to the hydrophilic (OH) moieties have their imprint in the 350-600 cm^{-1} (10-18 THz) frequency range of water's librational mode, originating from steric constraints in water rotational motions induced by the proximity to and direct H-bonding with the solutes, which causes a blue shift.[29] This population informs about the strength and number of attractive interactions that stabilize the solvation of a solute in water.[29] Using THz-calorimetry to relate spectroscopic signatures to thermodynamic quantities, combined with MD simulations to independently quantify the structural aspects of the distinct local contributions of the two identified populations to solvation entropy and enthalpy, we showed that the wrap population carries an entropic cost for (small) alcohol hydration, compensated by an enthalpic gain from the bound water

population.[27, 33] This concept of spectroscopically distinct hydrophilic and hydrophobic hydration populations is more general and can also be applied to water hydrating other solutes and biomolecules.[31, 34-36]

In this present work, we exploit these two THz-fingerprints in a combined experimental-theoretical approach to reveal that solvation thermodynamics of glycerol-water mixtures can be traced back to subtle changes occurring at the molecular level for specific HB motifs in the glycerol-water network. This result establishes a direct link between macroscopic/thermodynamic and microscopic/structural changes on a molecular level occurring as a function of the glycerol:water mixing ratio, paving the way to a rational tuning of solvation properties of these ubiquitously used binary solvents. In particular, analysis of the THz spectroscopic measurements allows us to probe local hydration populations without incorporating any molecular label. Theoretical THz/FIR spectra calculated from DFT-MD simulations are introduced to confirm the assignment of the experimentally identified populations. Classical MD simulations were used to characterize the associated structural motifs and to quantify the impact of changes in the water structure on the global mixing enthalpy in water-glycerol solutions. Our study maps a direct connection between mixing enthalpy and the local hydration motifs associated with the bound water population across the whole miscibility range of glycerol-water solutions.

Figure 2a shows room temperature (20°C) spectra of absorption coefficients ($\alpha_{solution}(\nu)$) of glycerol-water mixtures in the 30-600 cm^{-1} THz/FIR spectral range as a function of glycerol mole fraction (X_{gly}). From $\alpha_{solution}(\nu)$, we calculate the frequency dependent molar extinction coefficient $\varepsilon_{solution}(\nu)$ (following equation 2) as shown in figure 2b. As discussed in section 1 of the supporting information (SI), the estimated error due to signal-to-noise is on the order of 1% of the obtained $\varepsilon_{solution}(\nu)$ spectra. The extinction coefficient of pure glycerol and pure water at three different temperatures (5°C, 20°C, and 40°C) are shown in figure S2 of the SI, for comparison. The THz spectral range directly probes specific glycerol modes (at ~140, 310, 415, 490, 560 cm^{-1}), bulk water modes, as well as the specific hydration water fingerprints.[37] Specifically, we are interested in the THz fingerprints of wrap water molecules hydrating the hydrophobic groups without being H-bonded to the solute molecule (in the HB-stretching ~100-250 cm^{-1} region) and of bound water molecules H-bonded to the glycerol -OH hydrophilic groups (in the ~350-600 cm^{-1} region of water librations). All these contributions from glycerol, bulk and hydration water are convoluted in the spectra of figure 2.

To disentangle the spectroscopic signatures of the solvation shells, we calculate the effective molar extinction coefficient ($\varepsilon_{hydration}(\nu)$). $\varepsilon_{hydration}$ is obtained by subtracting the mole fraction-scaled extinction spectrum of bulk water (ε_{water}) and of pure bulk glycerol ($\varepsilon_{glycerol}$) from the extinction spectrum of each mixture (see equation 3). As per definition $\varepsilon_{hydration}$ gives exactly zero for the ideal case where the mixture is the sum of the two individual bulk components. Therefore, $\varepsilon_{hydration}$ spectra shown in figure 3a report on the direct interactions between water and glycerol (and associated structural perturbations, if any).

Focusing first on the ~100-250 cm^{-1} region that contains the signatures of the wrap population, we notice that this part of the spectra does not show any clear trend with X_{gly}. This indicates minor perturbation of the network, i.e., glycerol molecules fit well within the water network[18] (and vice versa at high X_{gly}). This is opposite to what observed for alcohol solutes, which showed a prominent wrap water band.[29]

In the ~350-600 cm^{-1} region (i.e., the onset of the water libration band), $\varepsilon_{hydration}$ increases remarkably with frequency and the steepness of such increase shows a clear and non-monotonic dependence with X_{gly}. An $\varepsilon_{hydration}$ increase at the onset of the libration band is typically observed for bound water molecules interacting with polar solutes, and the steepness of such increase was reported in a previous study[29] to inform on the number of bound water molecules, i.e., the number of water-solute HB in the system. Therefore, such feature is assigned to the formation of HBs between glycerol and bound water molecules. In order to quantitatively confirm this assignment, in figure 3b we present the theoretical THz spectra of bound water obtained from DFT-MD simulations as a function of the number of bound waters in the system (from 1 to 5). As described in section 2 of the SI, the spectrum of a bound water molecule interacting with an alcohol -OH has been calculated from a previously performed DFT-MD simulation, taking into account all the self- and cross- correlation terms involving the selected bound water molecule and averaging over all the bound waters in the model system. The theoretical spectra (figure 3b) demonstrate that each bound water molecule causes an increase in $\varepsilon_{hydration}$ in the librations frequency range and the steepness becomes larger with increasing number of bound water molecules. Specifically, as shown from the DFT-MD analysis, the bound water spectrum has a different line shape and a blue shifted peak position with respect to the spectrum of bulk water; both effects contribute the steepness (i.e., slope) of the intensity rise in the librational frequency range. Their combined effect, as measured by the slope, scales almost linearly with the number of bound water molecules, as demonstrated by the inset of figure 3b. In the inset, the theoretical slope is calculated by linearly fitting the THz intensity in

the 400-500 cm^{-1} range, and we carefully checked that the linear relationship between the slope and the number of bound waters holds true when freely varying the extremes of the fit within the frequency range were the intensity rise (i.e., in the 400-600 cm^{-1}). Based on these results, in the following we use the slope to quantify the variations in the number of bound water molecules (N_{bound}) in the glycerol-water mixtures directly from the measured $\varepsilon_{hydration}$ spectra.[29, 36]

To this end, we linearly fit the $\varepsilon_{hydration}$ spectra in the frequency range from 340 cm^{-1} to 475 cm^{-1} (blue shaded region in figure 3a), where 340 cm^{-1} is the turning point at which $\varepsilon_{hydration}$ starts to rise in the experiments. The experimentally determined slopes are shown in figure 3c as a function of X_{gly}, for three different temperatures. For low X_{gly} (<0.4), the slope systematically increases with increasing glycerol content in the mixture, which signifies an increase in the effective population of bound water molecules. The slope attains its maximum value at X_{gly}=0.4, indicating that all binding sites of glycerol molecules in the system are saturated. Further increasing glycerol content beyond this point leads to a decrease in the number of bound waters (quantified by the slope) since the number of water molecules in the system is reduced and glycerol-water interactions are progressively replaced by glycerol-glycerol interactions. These results are consistent with previous studies[17-18] that proposed a progressive structural transition from a state where bulk-like water dominates and solvates glycerol molecules at high water content (where N_{bound} is hence expected to increase monotonically with X_{gly}), to a state containing confined water populations, where there are not enough bound water molecules to solvate all glycerol molecules (where N_{bound} must decrease with X_{gly}).

To further characterize these structural transformations, we performed classical MD simulations and evaluated the relative mole fractions of wrap and bound water populations, as well as of bulk-like water not involved in the hydration of glycerol solutes (denoted hereafter non-shell water) as a function of X_{gly}. We identify these three populations in the simulations based on structural criteria, i.e., on their proximity with respect to OH and hydrophobic glycerol moieties and their H-bonding properties (see method section and references[27, 29] for details). The simulated variations in the three populations with X_{gly} are shown in figure 3d, as obtained by direct counting of water molecules belonging to each identified population. Most notably, the bound water population follows most closely the trend observed experimentally for the slope, i.e., first increasing with X_{gly} for low glycerol content, saturating at around X_{gly}=(0.3-0.4), and then decreasing with further increasing the glycerol mole fraction. The simulated wrap water population shows a similar non-monotonic behavior but reaches saturation for much

lower glycerol content (X_{gly}=0.15). This is easily understood by considering that wrap water interacts much more weakly than bound water with glycerol (since it is not H-bonded), and therefore it is the easiest part of the hydration layer that is stripped off as soon as the glycerol molecules in the system become too many to be fully hydrated. The small number of wrap water above X_{gly}>0.15 and its almost bulk-like character in glycerol-water mixtures below X_{gly}=0.15 (see figure 4) prevents an unambiguous identification of this population in the THz-spectra. The non-shell water population shows a straightforward monotonic decrease with X_{gly} since increasing numbers of water molecules are involved in the solvation of glycerol molecules as their concentration increases. To further corroborate our interpretation of the molecular arrangement of glycerol-water mixtures into three water populations, i.e., bulk, bound and wrap populations, we performed Raman measurements in the OH-stretching (2500-3700 cm^{-1}) frequency range. The results are presented in section 3 of the SI. We analyzed these spectra with the multivariate curve resolution (MCR) technique pioneered by Ben Amotz and co-workers.[38] The MCR technique is extremely powerful to reveal solute and solvation spectroscopic fingerprints. The method requires the choice of a reference point, where both glycerol and water have no or negligible intensity. While MCR works well for the high-frequency range, when all contributions are positive, and a reference point with zero intensity can be well defined, our THz decomposition approach is specialized for the low frequency range (where a reference point with zero intensity cannot be unambiguously defined) and allows to deal with non-ideal contributions of mixtures, which can be by construction positive or negative. Therefore, both the frequency range and the analysis method chosen for the Raman spectra are complementary to our approach for THz spectroscopy. As detailed in the SI, the results from the MCR analysis of the Raman spectra are consistent with the X_{gly} dependence observed for the bulk, wrap and bound water populations in the THz measurements and simulations. Moreover, Raman and THz spectra are found to provide complementary information: Raman spectroscopy provides a signature of the wrap water population, which is predicted to be present in the simulation but missing in the THz spectra (as discussed above), while the slope of the librational band in the THz/FIR region provides an unambiguous spectroscopic signature of bound water.

Further, we leverage MD simulations to probe shifts in the underlying structure of hydration water around glycerol with increased glycerol concentration. It is well-known that changes in water tetrahedrality impact the solvation thermodynamics in aqueous solutions.[37, 39-44] While other order parameters can probe water tetrahedrality,[45] we compute the three-body angle distribution[40, 43-44] by cataloguing angles between each water molecule and its two nearest

neighbors (excluding neighbors further than 3.4 Å from the central water molecule). To better visualize the changes in the distributions, we examine the differential three-body angle distribution $P_i(X_{gly}) - P_{pure}$ where, $P_i(X_{gly})$ is the distribution from population i at X_{gly} and P_{pure} is the reference distribution from pure water at 18°C. As shown in figure 4a, increments in the probability density at θ=109.5° and θ=64° correspond to enhancements in the tetrahedral and icosahedral (simple fluid-like) structure of water relative to pure water, respectively. The bound and wrap water populations exhibit distinct behavior for the entire range of compositions from X_{gly}=0 to 0.5. For small X_{gly}, the bound waters display more icosahedral character compared to pure water, suggesting that glycerol-water HBs disrupt the water-water HB-network for bound waters. On the other hand, the wrap waters show enhanced tetrahedrality and decreased icosahedrality compared to pure water for the entire concentration range. This enhancement of tetrahedrality near the hydrophobic CH_x groups is consistent with previous theoretical,[46] computational,[47-48] and experimental[49] findings, showing that the presence of small hydrophobic solutes (<1 nm in radius) enhances the structure of the surrounding hydration waters. By integrating over the tetrahedral region ($100° < \theta < 120°$) of the three-body angle distributions, we estimate the population of tetrahedrally-coordinated waters p_{tet} for bound, wrap, and the system-averaged (all water molecules in the simulation box). While wrap and system-averaged waters show higher p_{tet} than pure water, bound waters show lower p_{tet} relative to pure water for X_{gly}< 0.3 and higher p_{tet} for X_{gly}> 0.3 (figure 4b).

These structural changes as revealed by MD simulations and spectroscopic fingerprints are expected to directly influence the thermodynamic properties of glycerol-water mixtures. Deriving changes in entropy and enthalpy from spectral changes is the working principle of THz-calorimetry.[27, 29] This allowed us to determine that bound water molecules favorably contribute to the mixing enthalpy (ΔH_{bound}<0) when water is mixed with organic compounds containing at least one hydrophilic functional group.[29] However, for primary alcohols, the global thermodynamic properties were found dominated by the wrap water population (i.e., by the hydrophobic solvation mechanism).[27, 29] We hereafter demonstrate that this picture is reversed for glycerol-water mixtures.

In figure 5a, we plot the spectroscopically measured slope (for all the mixtures in the full X_{gly} range), i.e., the fingerprint of bound water molecules, against mixing enthalpy (ΔH_{mixing}) data from the literature.[12] Remarkably, the figure shows a nearly linear relationship, i.e., the larger the slope, the more negative (i.e., favorable) ΔH_{mixing} is. Similar plots with other available datasets[12, 50-51] of ΔH_{mixing} are reported in figure S6 of the supporting information and provide similar trends. To better investigate the correspondence between slope and ΔH_{mixing}, in

figure 5b, minus the slope (we multiply the slope by -1 because bound water molecules enthalpically stabilize the mixture) is displayed as a function of X_{gly} and compared with several ΔH_{mixing} experimental datasets[12, 50-51] as well as the theoretical mixing enthalpy values from the present MD simulations. The scatter of the experimental ΔH_{mixing} datasets gives an idea of the related confidence interval. Within this interval, we observe a very good agreement between simulated and measured ΔH_{mixing} values. Their trend with X_{gly} can be quantitatively predicted from the trend of the THz slope, i.e., from the bound water population H-bonded to glycerol molecules. This signifies that the changes in mixing enthalpy with X_{gly} can be quantitatively explained by considering the enthalpic term due to the variations in the bound water population, only. More specifically, the decrement of ΔH_{mixing} with X_{gly} (for $X_{gly}<0.4$) is ascribed to an increase in the number of bound water molecules that enthalpically stabilize the mixture, while the further increase of ΔH_{mixing} (for $X_{gly}>0.4$) is caused by the progressive decrease in the number of bound water molecules. This result has two important implications. First, bound water molecules enthalpically stabilize glycerol-water mixtures by the HB they form with glycerol –OH groups. Second, and most important, this local contribution from bound water controls the mixing enthalpy. As a consequence, the global trends in mixing enthalpy of heterogeneous glycerol-water binary mixtures can be understood by only considering the changes in one local structural hydration motif, the bound water population, without invoking contributions from more complex collective structural transformations occurring in the mixtures. This allows us to establish a powerful, well-defined structural and spectroscopic marker to track the changes in the thermodynamic properties of ubiquitously used glycerol-water solvents.

We now focus on the implications of our findings for cryopreservation. Cryopreservation traditionally requires the addition of organic solvents, glycerol being a gold-standard, to mitigate the damage caused by ice formation and growth, which would otherwise lead to cell death.[52] From the spectroscopic and simulations results of the present work, we can identify a glycerol concentration window with optimal structural and thermodynamic conditions to fulfill such cryoprotective role. This corresponds to a range of X_{gly} from 0.15 to 0.4. In this range, glycerol-water direct H-bonding and mixing enthalpy are maximized due to bound water being the dominant population (reaching the maximum at $X_{gly}=0.4$, figure 5), while the bulk-like water population is close to zero (figure 3d). The formation of such a glycerol-water network was proposed in a previous study to promote protein stabilization via preferential solvation.[4] In contrast, at $X_{gly}<0.15$, bulk water and wrap water (which has a bulk-like

character) are predominant, which is expected to lead to unwanted ice crystallization and growth upon cooling, as confirmed by previous experiments.[4] This offers a new light to explain the empirical observations that X_{gly}=0.15 is sufficient to achieve cryoprotection of many macromolecules.[53] We propose that this is due to X_{gly}=0.15 being the lowest glycerol concentration at which bound water is the dominant water population, while the "bulk-like" water population drops to almost zero. For X_{gly}>0.4, glycerol-water HBs are progressively replaced by glycerol-glycerol direct bonding as the dominant interactions, as marked spectroscopically by the decrease of the bound-water fingerprint, and thermodynamically by the decrease of the mixing enthalpy. This causes a progressive transition toward a glycerol "bulk-like" state. In such conditions, the cytotoxic effects of glycerol will manifest, possibly hindering the viability of too concentrated mixtures for specific applications.[52] Following the procedure introduced here, THz spectroscopy can be now used to systematically study the partial contributions of local water populations to the thermodynamics of binary mixtures and optimize their composition to fulfill a specific role (cryoprotection in our case).

In summary the addition of glycerol to aqueous solutions goes beyond its well-known impact on water's macroscopic viscosity. It alters the local molecular structures of hydration water and thereby dictates the mixing thermodynamics. The basic understanding provided here on such structure-thermodynamics relations offers a basis for a rational optimization of glycerol-water mixtures solvation properties by tailoring the mixture composition. Depending on the application, a goldilocks regime may be the minimum glycerol concentration in which all bulk water is transformed into H-bonded water populations, while minimizing destabilizing effects of glycerol on the protein, or a much higher glycerol concentration in which water-water HB interactions are minimized, and rather replaced by water-glycerol interactions. Our study also highlights the importance of explicit inclusion of glycerol in simulation studies of proteins for the purpose of comparing computed conformational dynamics and interactions to experimental studies that may require glycerol as solution constituents, e.g., in the study of Double Electron Electron Resonance (DEER).[54] By combining THz spectroscopy with ab initio and classical MD simulations and mixing enthalpy estimations, we characterized, at the molecular level, how the structural transformations occurring in glycerol-water mixtures dictate their thermodynamic properties. We built our interpretation on the "wrap" and "bound" hydration water spectroscopic populations, i.e., well-defined THz markers that allow the direct monitoring of the changes in the hydrophobic and hydrophilic solvation mechanisms as a function of mixtures composition, respectively. We propose that the mixing enthalpy between water and glycerol -

a macroscopic property - is dominated by local contributions from the (hydrophilic) hydration water populations directly bound to glycerol. For glycerol, the bound water and not the wrap water population dictates mixing enthalpy since glycerol fits nicely into the water network and vice versa. By contrast, for primary alcohols, a dominating volume exclusion term identified by the wrap population was found, as is typical for hydrophobic molecules. The general picture emerging by combining these results is that the balance between local hydrophobic/wrap and hydrophilic/bound contributions dictates global mixing thermodynamics. The balance depends crucially on the size of the molecules that are mixed with water as well as on the number and distributions on polar groups.

Our results offer exciting opportunities to tune in future studies the enthalpy and entropy of polyol-water mixtures - ubiquitously used co-solvents in biological and technological processes - by rationally modulating the nature of the polyol. For example, from the present results we can anticipate that the balance between the released water that is lost by the cell during freezing and the bound water population H-bonded to CPA (glycerol) and biomolecules - a key parameter for cryoprotection – can be tailored by experimentally exploiting the bound water THz fingerprint and making direct connection to mixing enthalpy. Specifically, the mixing ratio and type of polyol can be tuned to find conditions in which the bound water THz fingerprint is maximized in the spectrum, and hence maximize the number of bound waters and minimize that of released free water during freezing that, presumably, will more readily crystalize. Moreover, we expect that increasing the size of the polyol with respect to glycerol will be an efficient strategy to manipulate mixing enthalpy, by shifting the balance between bound and wrap hydration populations toward the latter, due to an increased volume exclusion effect. Such a shift in the bound/wrap balance could be used for instance to promote the solubility of more hydrophobic biomolecules (i.e., containing large hydrophobic surface patches exposed to the solvent), for which one important unfavorable thermodynamic contribution to solvation is the enthalpic penalty required to break strong solvent-solvent HBs (i.e., bound water-polyol interactions in the mixtures). Based on the present results, mixture compositions where such penalty is minimized could be efficiently designed with a high-throughput THz screening, by "simply" finding the mixture composition for which the THz/FIR fingerprint of bound waters is minimized. The proposed Ansatz is general and can be transferred even to complex mixtures, since the spectroscopic fingerprints of the structural motifs around hydrophobic and hydrophilic hydrated moieties are well resolved in the THz/FIR range between 50-600 cm^{-1}, and less obscured by intramolecular modes then in the infrared.

Methods

Experimental section: Glycerol was purchased from Sigma-Aldrich and was used without further purification. All the aqueous binary mixtures were prepared using de-ionized ultrapure water. The temperature-controlled densities of all the mixtures were measured by Anton-Paar DMA 4500M density meter to calculate the concentrations of water and glycerol in the solutions ($c_{water,solution}$ and $c_{glycerol,solution}$). Far infrared (FIR/THz) spectra of glycerol aqueous solutions with various mole fractions were recorded in the frequency range from 30 cm^{-1} to 600 cm^{-1} (0.9 THz – 18 THz) at three different temperatures (5^0C, 20^0C and 40^0C) using Bruker (USA) Vertex 80v Fourier transform infrared (FTIR) spectrometer. A mercury lamp was used as the light source of the broadband THz radiation and an Infrared Laboratories liquid-helium-cooled silicon bolometer (HDL-5, Infrared Laboratories, USA) was employed as an external detector. A Mylar-multilayer beam splitter was utilized to generate the interferogram. All the THz/FIR experiments were measured using a temperature-controlled (T±0.2 K) liquid transmission cell with diamond windows (0.5 mm thick) and a Kapton spacer with a nominal thickness of 26±0.2 μm. The interference pattern of the empty cell was used to determine the exact sample thickness (d) and was used for further data analysis. During the measurements, the sample compartment was constantly purged with dry nitrogen to minimize the absorption arising from the residual water vapor. Further details about the spectrometer can be found elsewhere.[33, 36, 55] Each spectrum was measured with a resolution of 2 cm^{-1}. Each spectrum presented here is the average of 64 individual measurements. All the data were validated with repeated experiments.

Experimental data analysis protocol: The frequency-dependent absorption coefficient α(ν) of glycerol-water solutions are determined using the Lambert–Beer law,

$$\Delta\alpha(\nu) = \alpha_{solution} - \alpha_{water} = -\frac{1}{d} ln \left(\frac{I_{solution}(\nu)}{I_{water}(\nu)}\right) \quad (1)$$

where d is the thickness of the sample, $I_{solution}$ and I_{water} are the transmitted intensities through the samples and pure water, respectively. Water is used as a reference to minimize "etalon effects", i.e., standing waves. From the absorption coefficient we calculate the average molar extinction coefficients,

$$\varepsilon_{solution}(\nu) = \frac{\alpha_{solution}(\nu)}{c_{water,solution} + c_{glycerol,solution}} \quad (2)$$

The molar effective extinction is deduced by subtracting the partial contributions from bulk water and bulk glycerol, respectively.

$$\varepsilon_{hydration}(\nu) = \varepsilon_{solution}(\nu) - X_{gly}\varepsilon_{glycerol}(\nu) - (1 - X_{gly})\varepsilon_{water}(\nu) \quad (3)$$

where $\varepsilon_{glycerol}$ and ε_{water} are the molar extinction coefficient of pure glycerol and pure water, respectively. $\varepsilon_{hydration}(\nu)$ bears information solely about the inter-molecular interactions that are changed due to the mixing of water and glycerol (i.e., water-water, water-glycerol, glycerol-glycerol interactions in the solvation shells). The glycerol content in the mixtures is quantified in terms of mole fraction:

$$X_{gly} = \frac{n_{gly}}{n_{gly}+n_w} \quad (4)$$

where $n_{gly} and\ n_w$ are the number of moles of glycerol and water, respectively, in the solutions.

MD Simulation details: We model glycerol-water mixtures using Bliek-Chelli parameters for glycerol[26, 56] and the OPC 4-site water model.[57] This combination of models reliably reproduces the equilibrium thermophysical properties of glycerol-water under standard temperature and pressure conditions.[25, 56, 58] Coulombic interactions are modelled using the particle-mesh Ewald summation scheme (PME).[59] We simulate glycerol-water mixtures with X_{gly} = 0, 0.012, 0.033, 0.05, 0.075, 0.10, 0.15, 0.20, 0.30, 0.40, 0.50, and 1.0 at 18°C and 1-bar using the GPU-optimized OpenMM package.[60] All simulations are conducted in the NPT ensemble, with temperature and pressure held constant using a Langevin thermostat[60] and Monte Carlo barostat,[60] respectively. We generate initial configurations containing 729 glycerol molecules and vary the numbers of water molecules using Packmol software.[61] We first minimize the energy of the initial configurations and equilibrate in the NPT ensemble for times ranging from 100 to 500 ns depending on the glycerol concentration. We conduct production runs for 100 ns with system coordinates saved every 100 ps. Following the simulations, we analyze the production run trajectories using the pytraj Python library[62] and in-house python code. For each simulation-computed quantity, we estimate the standard error of the mean using block averaging.

We define HBs via the widely used geometric criteria of Luzar and Chandler,[63] namely, O-H distance and $\angle O - OH$ angular cut-off values of 3.5 Å and 120° degrees, respectively. Tracking the waters participating in water-glycerol H-bonds, we directly quantify the mole fraction of bound waters (X_{bound}). We also calculate the mole fraction of wrap waters (X_{wrap}) by defining them as waters lying within the second hydration shell of glycerol molecules (~4.2 Å, as defined by the second minima of the 2D radial distribution function between glycerol heavy atoms and water oxygens), but not H-bonded to the glycerol molecule. Non-shell waters

are simply all waters residing more than 4.2 Å from glycerol heavy atoms. We compute concentration of non-shell waters on a molar basis via $X_{non-shell} = (1 - X_{bound} - X_{wrap} - X_{gly})$. Using the simulation-averaged molar energy $< E >$ and molar volume $< V >$ for each simulated glycerol-water mixture, we also estimate the molar energy of mixing and molar volume of mixing,

$$\Delta E_{mixing} =< E > -X_{gly} < E_{glycerol} > -\left(1 - X_{gly}\right) < E_{water} > \quad (5)$$

$$\Delta V_{mixing} =< V > -X_{gly} < V_{glycerol} > -\left(1 - X_{gly}\right) < V_{water} > \quad (6)$$

where, $< Y_{water} >$ and $< Y_{glycerol} >$ correspond to the average molar value of Y for pure water and glycerol, respectively. The molar enthalpy of mixing ΔH_{mixing} hence stems from the relation

$$\Delta H_{mixing} = \Delta E_{mixing} + P\Delta V_{mixing} \quad (7)$$

where P is system pressure (1 bar).

Acknowledgements

M.H., G.S., S.P. and D.D.M. acknowledge financial support by European Research Council (ERC) Advanced Grant 695437 THz-Calorimetry. This study is funded by the Deutsche Forschungsgemeinschaft (DFG, German Re- search Foundation) under Germany's Excellence Strat- egy—EXC2033–390677874—RESOLV. We thank the Mercator Research Center Ruhr (MERCUR) for funding. M.S.S. and D.R.B. acknowledge support from the Center for Materials for Water and Energy Systems, an Energy Frontier Research Center funded by the US Department of Energy, Office of Science, Basic Energy Sciences under Award DE-SC0019272. Use was made of computational facilities purchased with funds from the National Science Foundation (OAC-1925717) and administered by the Center for Scientific Computing (CSC). The CSC is supported by the California NanoSystems Institute and the Materials Research Science and Engineering Center (MRSEC; NSF DMR 1720256) at UC Santa Barbara. S.H. acknowledges support from the Deutsche Forschungsgemeinschaft (DFG) im Rahmen der Exzellenzstrategie des Bundes und der Länder – EXC 2033 – Projektnummer 390677874 – RESOLVDFG-EXC-2033 project #390677874.

Figures

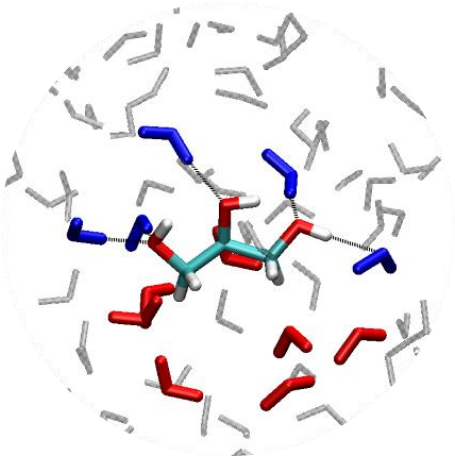

Figure 1. Schematic representation of the hydration shell of a glycerol molecule with bound (blue colored) and wrap (red colored) water molecules. The non-shell (bulk-like) water molecules are indicated with gray color.

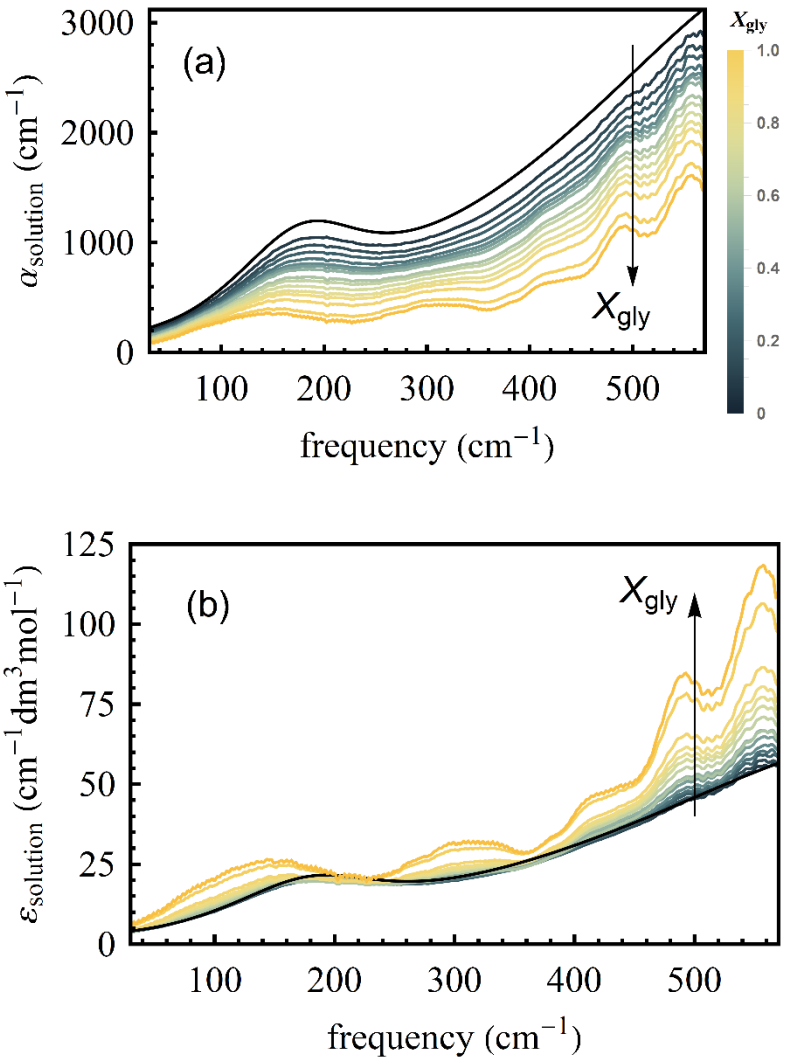

Figure 2. The frequency dependent (a) absorption coefficient and (b) average molar extinction coefficient of glycerol-water binary mixtures with various mole fraction of glycerol (X_{gly}). Arrows indicate increasing X_{gly}.

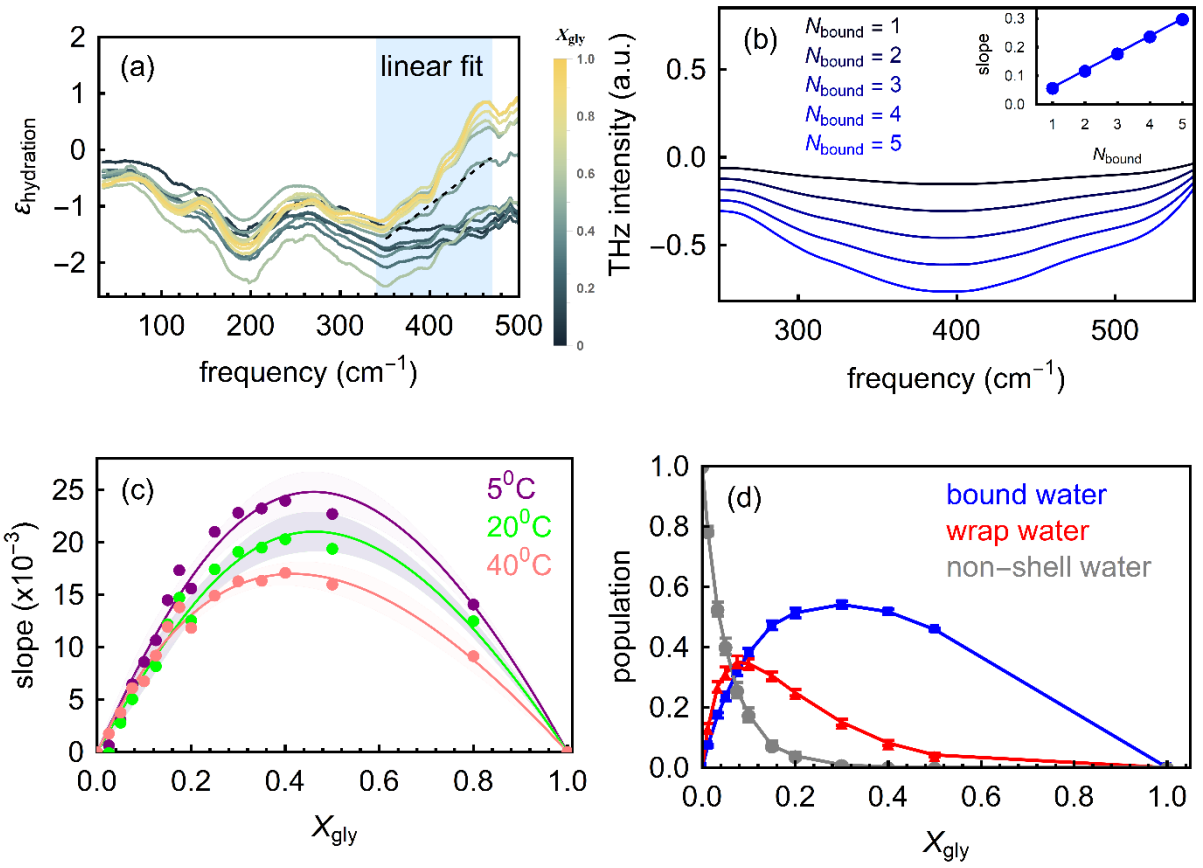

Figure 3. (a) Molar effective extinction coefficient ($\varepsilon_{hydration}$) at 20°C as calculated by subtracting the bulk water and bulk glycerol spectra from the glycerol-water mixtures spectra (equation 4). $\varepsilon_{hydration}$ increases with frequency for all the solutions in the 350-450 cm^{-1} spectral region (gray shaded area). The dotted black line is the linear fit in that region for glycerol mole fraction; X_{gly}=0.8, as an example. The slope obtained from such fit is then used for interpreting the spectral trends. (b) Theoretical THz spectrum of bound waters as a function of the number of bound water (N_{bound}) molecules. The inset shows that the theoretically calculated slope depends linearly on N_{bound}. (c) The slopes derived from experimental $\varepsilon_{hydration}$ spectra are plotted for all X_{gly} at three different temperatures (5°C, 20°C and 40°C, respectively). The lines are guide to the eye and the error bars are indicated by the shaded regions. (d) The populations of various types of water in glycerol-water mixtures as predicted by MD simulations.

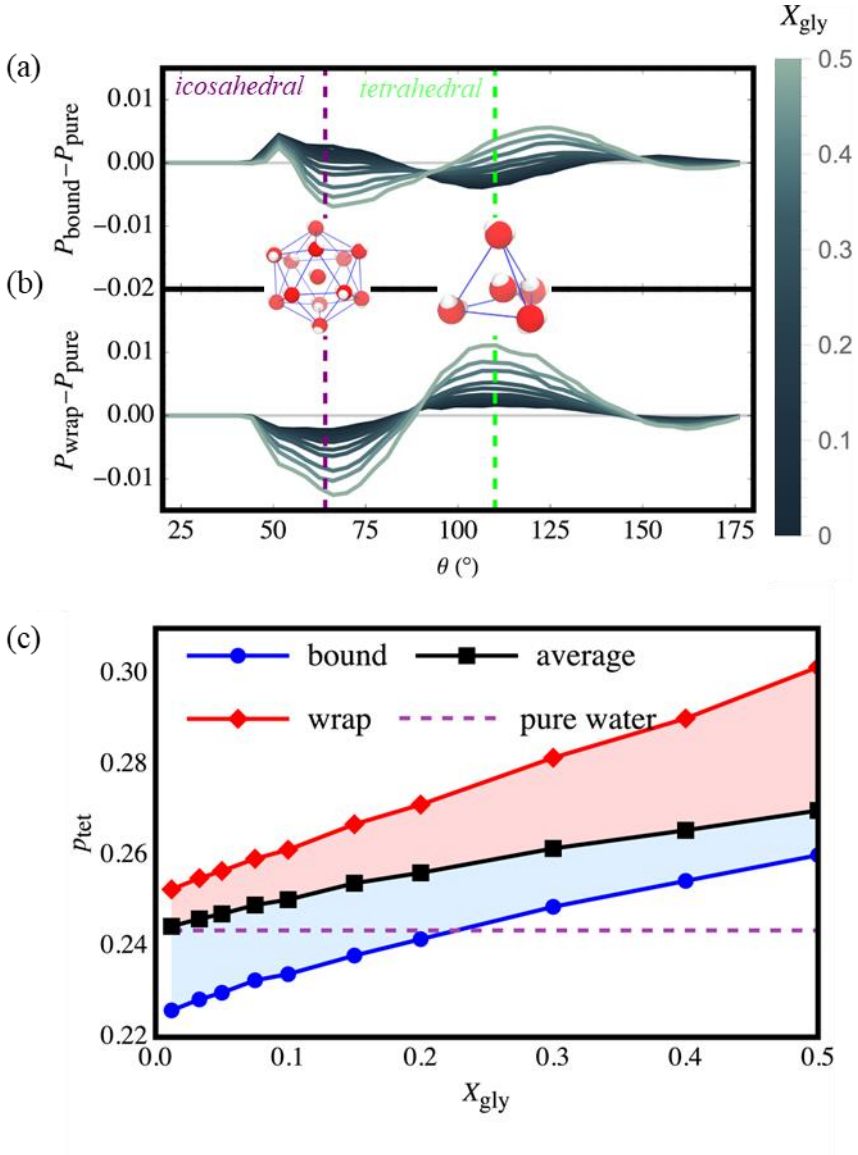

Figure 4. The three-body angle distributions for the (a) bound and (b) wrap waters exhibit a monotonic increase in the population of tetrahedrally-coordinated waters (θ=109.5°) and a monotonic decrease in the population of icosahedrally-coordinated waters (θ=64°) relative to pure water at 18°C. (a) At low concentrations, the bound waters are more icosahedral and less tetrahedral than pure water. (b) Wrap waters show higher tetrahedral populations and lower icosahedral populations than pure water for the entire concentration range. (c) Here, we depict the population of tetrahedrally-coordinated waters p_{tet} for wrap (red diamonds), sample-averaged (black squares), and bound (blue circles) as a function of glycerol mole faction. For reference, we indicate the pure water p_{tet} with a horizontal dashed line.

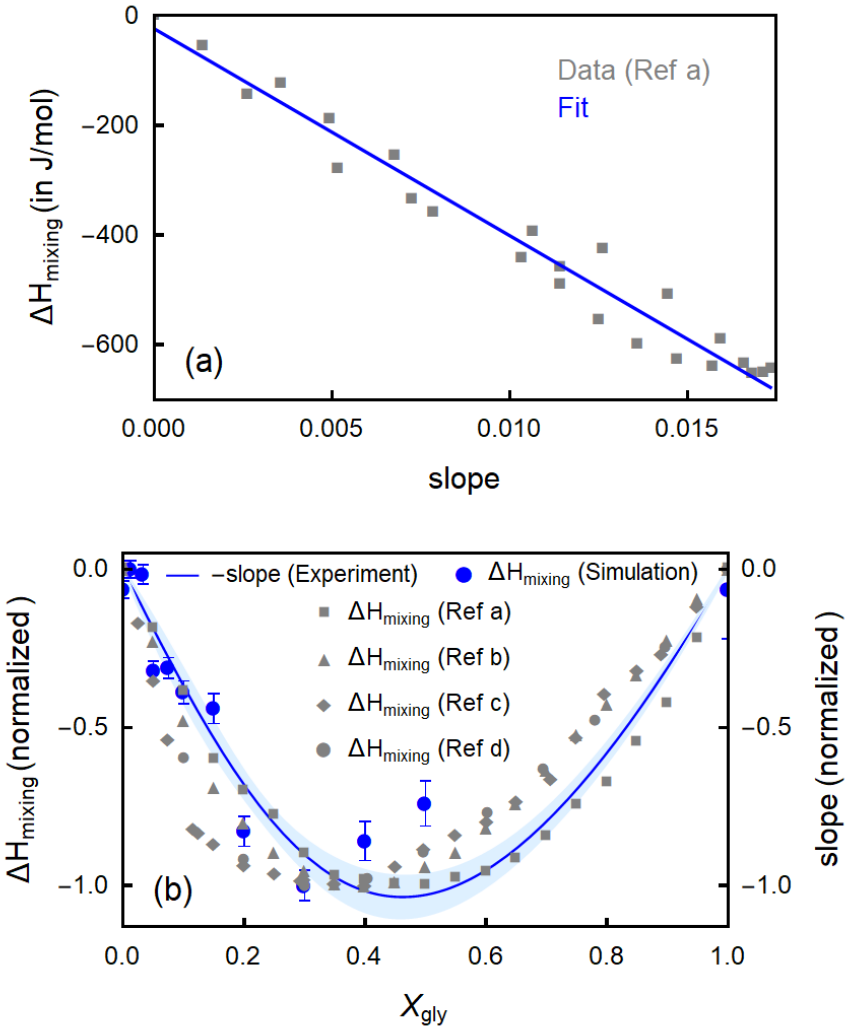

Figure 5. (a) Linear correlation between ΔH_{mixing} (from Ref a,[12] as an example) and the experimentally determined slope. (b) The trend of the experimentally determined slope (associated to the bound water population) as a function of glycerol content is compared with that of mixing enthalpy values (ΔH_{mixing}). ΔH_{mixing} values are taken from four literature data sets: Ref a; calculated,[12] Ref b; extrapolated,[12] Ref c,[51] Ref d,[51] as well as from the present MD simulations. All data sets are scaled (from 0 to -1) for better comparison.

References

1. Weng, L.; Stott, S. L.; Toner, M., Exploring Dynamics and Structure of Biomolecules, Cryoprotectants, and Water Using Molecular Dynamics Simulations: Implications for Biostabilization and Biopreservation. *Annu. Rev. Biomed. Eng.* **2019,** *21* (1), 1-31.
2. Towey, J. J.; Soper, A. K.; Dougan, L., Molecular Insight Into the Hydrogen Bonding and Micro-Segregation of a Cryoprotectant Molecule. *J. Phys. Chem. B* **2012,** *116* (47), 13898-13904.
3. Towey, J. J.; Soper, A. K.; Dougan, L., What happens to the structure of water in cryoprotectant solutions? *Faraday Discuss.* **2013,** *167* (0), 159-176.

4. Dashnau, J. L.; Nucci, N. V.; Sharp, K. A.; Vanderkooi, J. M., Hydrogen Bonding and the Cryoprotective Properties of Glycerol/Water Mixtures. *J. Phys. Chem. B* **2006,** *110* (27), 13670-13677.
5. Gekko, K.; Timasheff, S. N., Mechanism of protein stabilization by glycerol: preferential hydration in glycerol-water mixtures. *Biochemistry* **1981,** *20* (16), 4667-4676.
6. Chéron, N.; Naepels, M.; Pluhařová, E.; Laage, D., Protein Preferential Solvation in Water:Glycerol Mixtures. *J. Phys. Chem. B* **2020,** *124* (8), 1424-1437.
7. Hirai, M.; Ajito, S.; Sugiyama, M.; Iwase, H.; Takata, S.-i.; Shimizu, N.; Igarashi, N.; Martel, A.; Porcar, L., Direct Evidence for the Effect of Glycerol on Protein Hydration and Thermal Structural Transition. *Biophysical Journal* **2018,** *115* (2), 313-327.
8. Weng, L.; Chen, C.; Zuo, J.; Li, W., Molecular Dynamics Study of Effects of Temperature and Concentration on Hydrogen-Bond Abilities of Ethylene Glycol and Glycerol: Implications for Cryopreservation. *J. Phys. Chem. A* **2011,** *115* (18), 4729-4737.
9. Toner, M.; Cravalho, E. G.; Karel, M., Thermodynamics and kinetics of intracellular ice formation during freezing of biological cells. *J. Appl. Phys.* **1990,** *67* (3), 1582-1593.
10. Karlsson, J. O. M.; Cravalho, E. G.; Toner, M., A model of diffusion-limited ice growth inside biological cells during freezing. *J. Appl. Phys.* **1994,** *75* (9), 4442-4455.
11. Egorov, G. I.; Makarov, D. M.; Kolker, A. M., Volume properties of liquid mixture of water+glycerol over the temperature range from 278.15 to 348.15K at atmospheric pressure. *Thermochimica Acta* **2013,** *570*, 16-26.
12. Marcus, Y., Some thermodynamic and structural aspects of mixtures of glycerol with water. *Phys. Chem. Chem. Phys.* **2000,** *2* (21), 4891-4896.
13. Hayashi, Y.; Puzenko, A.; Feldman, Y., Slow and fast dynamics in glycerol–water mixtures. *J. Non-Cryst. Solids* **2006,** *352* (42), 4696-4703.
14. To, E. C. H.; Davies, J. V.; Tucker, M.; Westh, P.; Trandum, C.; Suh, K. S. H.; Koga, Y., Excess Chemical Potentials, Excess Partial Molar Enthalpies, Entropies, Volumes, and Isobaric Thermal Expansivities of Aqueous Glycerol at 25°C. *J. Solution Chem.* **1999,** *28* (10), 1137-1157.
15. Grasso, R.; Musumeci, F.; Gulino, M.; Scordino, A., Exploring the behaviour of water in glycerol solutions by using delayed luminescence. *PLOS ONE* **2018,** *13* (1), e0191861.
16. Mallamace, F.; Mallamace, D.; Chen, S.-H.; Lanzafame, P.; Papanikolaou, G., Hydrophilic and Hydrophobic Effects on the Structure and Themodynamic Properties of Confined Water: Water in Solutions. *Int. J. Mol. Sci.* **2021,** *22* (14), 7547.
17. Weeraratna, C.; Amarasinghe, C.; Lu, W.; Ahmed, M., A Direct Probe of the Hydrogen Bond Network in Aqueous Glycerol Aerosols. *J. Phys. Chem. Lett.* **2021,** *12* (23), 5503-5511.
18. Charkhesht, A.; Lou, D.; Sindle, B.; Wen, C.; Cheng, S.; Vinh, N. Q., Insights into Hydration Dynamics and Cooperative Interactions in Glycerol–Water Mixtures by Terahertz Dielectric Spectroscopy. *J. Phys. Chem. B* **2019,** *123* (41), 8791-8799.
19. Behrends, R.; Fuchs, K.; Kaatze, U.; Hayashi, Y.; Feldman, Y., Dielectric properties of glycerol/water mixtures at temperatures between 10 and 50°C. *J. Chem. Phys.* **2006,** *124* (14), 144512.
20. Hayashi, Y.; Puzenko, A.; Balin, I.; Ryabov, Y. E.; Feldman, Y., Relaxation Dynamics in Glycerol−Water Mixtures. 2. Mesoscopic Feature in Water Rich Mixtures. *J. Phys. Chem. B* **2005,** *109* (18), 9174-9177.
21. Puzenko, A.; Hayashi, Y.; Ryabov, Y. E.; Balin, I.; Feldman, Y.; Kaatze, U.; Behrends, R., Relaxation Dynamics in Glycerol−Water Mixtures: I. Glycerol-Rich Mixtures. *J. Phys. Chem. B* **2005,** *109* (12), 6031-6035.
22. Leavesley, A.; Wilson, C. B.; Sherwin, M.; Han, S., Effect of water/glycerol polymorphism on dynamic nuclear polarization. *Phys. Chem. Chem. Phys.* **2018,** *20* (15), 9897-9903.

23. Nagasaki, Y.; Ishii, T.; Sunaga, Y.; Watanabe, Y.; Otsuka, H.; Kataoka, K., Novel Molecular Recognition via Fluorescent Resonance Energy Transfer Using a Biotin-PEG/Polyamine Stabilized CdS Quantum Dot. *Langmuir* **2004,** *20*, 6396-6400.
24. Zelent, B.; Nucci, N. V.; Vanderkooi, J. M., Liquid and Ice Water and Glycerol/Water Glasses Compared by Infrared Spectroscopy from 295 to 12 K. *J. Phys. Chem. A* **2004,** *108* (50), 11141-11150.
25. Chelli, R.; Procacci, P.; Cardini, G.; Califano, S., Glycerol condensed phases Part II.A molecular dynamics study of the conformational structure and hydrogen bonding. *Phys. Chem. Chem. Phys.* **1999,** *1* (5), 879-885.
26. Chelli, R.; Procacci, P.; Cardini, G.; G. Della Valle, R.; Califano, S., Glycerol condensed phases Part I. A molecular dynamics study. *Phys. Chem. Chem. Phys.* **1999,** *1* (5), 871-877.
27. Conti Nibali, V.; Pezzotti, S.; Sebastiani, F.; Galimberti, D.; Schwaab, G.; Heyden, M.; Gaigeot, M.-P.; Havenith, M., Wrapping up hydrophobic hydration: Locality matters. *J. Phys. Chem. Lett.* **2020,** *11* (12), 4809-4816.
28. Morawietz, T.; Marsalek, O.; Pattenaude, S. R.; Streacker, L. M.; Ben-Amotz, D.; Markland, T. E., The Interplay of Structure and Dynamics in the Raman Spectrum of Liquid Water over the Full Frequency and Temperature Range. *J. Phys. Chem. Lett.* **2018,** *9* (4), 851-857.
29. Pezzotti, S.; Sebastiani, F.; van Dam, E. P.; Ramos, S.; Conti Nibali, V.; Schwaab, G.; Havenith, M., Spectroscopic Fingerprints of Cavity Formation and Solute Insertion as a Measure of Hydration Entropic Loss and Enthalpic Gain. *Angewandte Chemie International Edition* **2022,** *61* (29), e202203893.
30. Das Mahanta, D.; Patra, A.; Samanta, N.; Luong, T. Q.; Mukherjee, B.; Mitra, R. K., Non-monotonic dynamics of water in its binary mixture with 1,2-dimethoxy ethane: A combined THz spectroscopic and MD simulation study. *J. Chem. Phys.* **2016,** *145* (16), 164501.
31. Funke, S.; Sebastiani, F.; Schwaab, G.; Havenith, M., Spectroscopic fingerprints in the low frequency spectrum of ice (Ih), clathrate hydrates, supercooled water, and hydrophobic hydration reveal similarities in the hydrogen bond network motifs. *J. Chem. Phys.* **2019,** *150* (22), 224505.
32. Vondracek, H.; Alfarano, S.; Hoberg, C.; Kolling, I.; Novelli, F.; Sebastiani, F.; Brubach, J.-B.; Roy, P.; Schwaab, G.; Havenith, M., Urea's match in the hydrogen-bond network? A high pressure THz study. *Biophysical Chemistry* **2019,** *254*, 106240.
33. Böhm, F.; Schwaab, G.; Havenith, M., Mapping Hydration Water around Alcohol Chains by THz Calorimetry. *Angew. Chem., Int. Ed.* **2017,** *56* (33), 9981-9985.
34. Adams, E. M.; Pezzotti, S.; Ahlers, J.; Rüttermann, M.; Levin, M.; Goldenzweig, A.; Peleg, Y.; Fleishman, S. J.; Sagi, I.; Havenith, M., Local Mutations Can Serve as a Game Changer for Global Protein Solvent Interaction. *JACS Au* **2021,** *1* (7), 1076-1085.
35. Ahlers, J.; Adams, E. M.; Bader, V.; Pezzotti, S.; Winklhofer, K. F.; Tatzelt, J.; Havenith, M., The key role of solvent in condensation: Mapping water in liquid-liquid phase-separated FUS. *Biophysical Journal* **2021,** *120* (7), 1266-1275.
36. Ma, C. Y.; Pezzotti, S.; Schwaab, G.; Gebala, M.; Herschlag, D.; Havenith, M., Cation enrichment in the ion atmosphere is promoted by local hydration of DNA. *Phys. Chem. Chem. Phys.* **2021,** *23* (40), 23203-23213.
37. Starciuc, T.; Guinet, Y.; Hedoux, A.; Shalaev, E., Water content thresholds in glycerol/water system: Low- and high-wavenumber Raman spectroscopy study. *J. Mol. Liq.* **2021,** *321*, 114678.
38. Ben-Amotz, D., Hydration-Shell Vibrational Spectroscopy. *Journal of the American Chemical Society* **2019,** *141* (27), 10569-10580.

39. Galamba, N., Water's Structure around Hydrophobic Solutes and the Iceberg Model. *J. Phys. Chem. B* **2013,** *117* (7), 2153-2159.
40. Chaimovich, A.; Shell, M. S., Tetrahedrality and structural order for hydrophobic interactions in a coarse-grained water model. *Phys. Rev. E* **2014,** *89* (2), 022140.
41. Dahanayake, J. N.; Mitchell-Koch, K. R., Entropy connects water structure and dynamics in protein hydration layer. *Phys. Chem. Chem. Phys.* **2018,** *20* (21), 14765-14777.
42. Wu, X.; Lu, W.; Streacker, L. M.; Ashbaugh, H. S.; Ben-Amotz, D., Temperature-Dependent Hydrophobic Crossover Length Scale and Water Tetrahedral Order. *J. Phys. Chem. Lett.* **2018,** *9* (5), 1012-1017.
43. Stock, P.; Monroe, J. I.; Utzig, T.; Smith, D. J.; Shell, M. S.; Valtiner, M., Unraveling Hydrophobic Interactions at the Molecular Scale Using Force Spectroscopy and Molecular Dynamics Simulations. *ACS Nano* **2017,** *11* (3), 2586-2597.
44. Monroe, J. I.; Shell, M. S., Decoding signatures of structure, bulk thermodynamics, and solvation in three-body angle distributions of rigid water models. *J. Chem. Phys.* **2019,** *151* (9), 094501.
45. Errington, J. R.; Debenedetti, P. G., Relationship between structural order and the anomalies of liquid water. *Nature* **2001,** *409* (6818), 318-321.
46. Lum, K.; Chandler, D.; Weeks, J. D., Hydrophobicity at small and large length scales. *J. Phys. Chem. B* **1999,** *103* (22), 4570-4577.
47. Huang, D. M.; Chandler, D., The Hydrophobic Effect and the Influence of Solute−Solvent Attractions. *J. Phys. Chem. B* **2002,** *106* (8), 2047-2053.
48. Huang, D. M.; Geissler, P. L.; Chandler, D., Scaling of Hydrophobic Solvation Free Energies. *J. Phys. Chem. B* **2001,** *105* (28), 6704-6709.
49. Di, W.; Gao, X.; Huang, W.; Sun, Y.; Lei, H.; Liu, Y.; Li, W.; Li, Y.; Wang, X.; Qin, M.; Zhu, Z.; Cao, Y.; Wang, W., Direct Measurement of Length Scale Dependence of the Hydrophobic Free Energy of a Single Collapsed Polymer Nanosphere. *Phys. Rev. Lett.* **2019,** *122* (4), 047801.
50. Zehioua, R.; Coquelet, C.; Soo, C.-B.; Richon, D.; Meniai, A.-H., Experimental and predicted excess molar enthalpies of some working pairs for absorption cycles. *Thermochimica Acta* **2009,** *495* (1), 72-80.
51. Batov, D. V.; Zaichikov, A. M.; Slyusar, V. P.; Korolev, V. P., Enthalpies of Mixing and State of Components in Aqueous-Organic Mixtures with Nets of Hydrogen Bonds. *Russian Journal of General Chemistry* **2001,** *71* (8), 1208-1214.
52. Hasan, M.; Fayter, A. E. R.; Gibson, M. I., Ice Recrystallization Inhibiting Polymers Enable Glycerol-Free Cryopreservation of Microorganisms. *Biomacromolecules* **2018,** *19* (8), 3371-3376.
53. Garman, E. F.; Owen, R. L., Cryocooling and radiation damage in macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **2006,** *62* (Pt 1), 32-47.
54. Georgieva, E. R.; Roy, A. S.; Grigoryants, V. M.; Borbat, P. P.; Earle, K. A.; Scholes, C. P.; Freed, J. H., Effect of freezing conditions on distances and their distributions derived from Double Electron Electron Resonance (DEER): A study of doubly-spin-labeled T4 lysozyme. *J. Mag. Res.* **2012,** *216*, 69-77.
55. Schwaab, G.; Sebastiani, F.; Havenith, M., Ion Hydration and Ion Pairing as Probed by THz Spectroscopy. *Angewandte Chemie International Edition* **2019,** *58* (10), 3000-3013.
56. Blieck, J.; Affouard, F.; Bordat, P.; Lerbret, A.; Descamps, M., Molecular dynamics simulations of glycerol glass-forming liquid. *Chem. Phys.* **2005,** *317* (2), 253-257.
57. Izadi, S.; Anandakrishnan, R.; Onufriev, A. V., Building Water Models: A Different Approach. *J. Phys. Chem. Lett.* **2014,** *5* (21), 3863-3871.
58. Jahn, D. A.; Akinkunmi, F. O.; Giovambattista, N., Effects of Temperature on the Properties of Glycerol: A Computer Simulation Study of Five Different Force Fields. *J. Phys. Chem. B* **2014,** *118* (38), 11284-11294.

59. Darden, T.; York, D.; Pedersen, L., Particle mesh Ewald: An N· log (N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993,** *98* (12), 10089-10092.
60. Eastman, P.; Pande, V., OpenMM: A Hardware-Independent Framework for Molecular Simulations. *Computing in Science & Engineering* **2010,** *12* (4), 34-39.
61. Martínez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M., PACKMOL: A package for building initial configurations for molecular dynamics simulations. *Journal of Computational Chemistry* **2009,** *30* (13), 2157-2164.
62. Roe, D. R.; Cheatham, T. E., PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and Computation* **2013,** *9* (7), 3084-3095.
63. Luzar, A.; Chandler, D., Structure and hydrogen bond dynamics of water–dimethyl sulfoxide mixtures by computer simulations. *J. Chem. Phys.* **1993,** *98* (10), 8160-8173.

Supporting information for

Local Water Structures Govern the Mixing Thermodynamics of Glycerol-Water Solutions

Debasish Das Mahanta[1,2], Dennis Robinson Brown[3], Simone Pezzotti[1], Songi Han[4,5], Gerhard Schwaab[1], M. Scott Shell[3,*], Martina Havenith[1,2,*]

[1]Lehrstuhl für Physikalische Chemie II, Ruhr-Universität Bochum, 44780 Bochum, Germany

[2]Department of Physics, Technische Universität Dortmund, 44227 Dortmund, Germany

[3]Department of Chemical Engineering, University of California, Santa Barbara, California 93106-5080, United States

[4]Department of Chemical Engineering, University of California, Santa Barbara, California 93106-5080, United States

[5]Department of Chemistry and Biochemistry, University of California, Santa Barbara, California 93106-9510, United States

*Corresponding author

Email: shell@ucsb.edu, Martina.Havenith@ruhr-uni-bochum.de

Section 1: *Noise estimation in* the extinction coefficient spectra

Our signal contrast is determined by the signal size relative to the signal difference with respect to water. We have estimated the noise level from the principal component analysis (PCA). PCA is a method of multivariate data analysis, provides mutually orthogonal principal components with an order according to their decreasing importance. We have broadly applied PCA to analyse THz spectra in our previous studies (see e.g. ref.[1] for more details). In a nutshell, this is a statistical approach to decompose a large multivariate data set of partly correlated variables to a set of uncorrelated variables i.e., principal components through orthogonal transformations. For glycerol-water mixtures we found there exist two most important components which can retrieve almost 99% of the original spectra. The higher order principal components correspond to the statistical noise and small systematic error. We have reconstructed the THz extinction spectra from the two most important principal components (see figure S1 below). We also have calculated the residual part of the spectra by subtracting the reconstructed spectra from the original measured spectra. This residual part estimates the noise level. As shown in the right panel of the figure below, the noise level is 100 times lower than the original measured spectra.

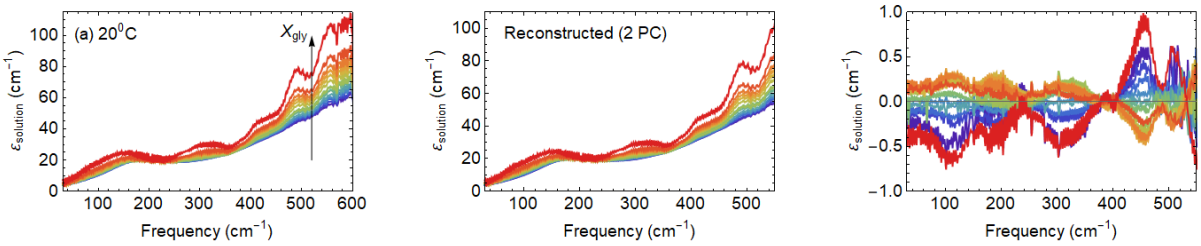

Figure S1. The left most panel shows the original measured THz extinction spectra at 20^{0}C. We have reconstructed the spectra from the two most important principal components (in the middle panel). The residual of the spectra is calculated by subtracting the reconstructed spectra from the original measured spectra (in the right most panel).

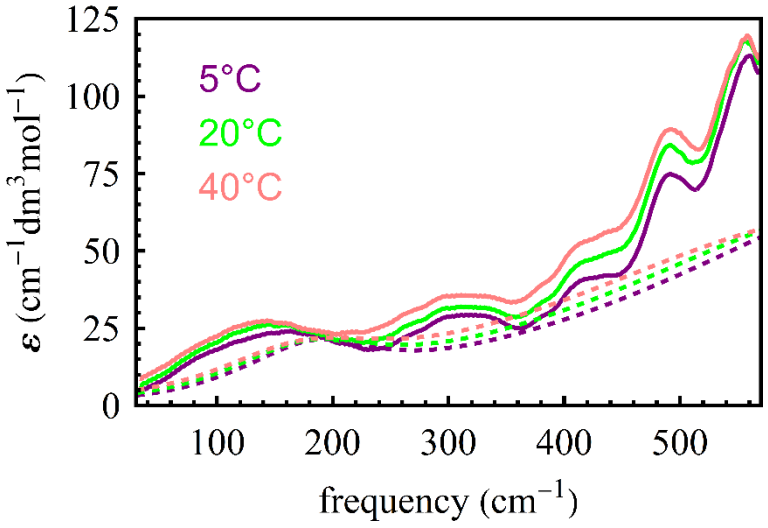

Figure S2. Average molar extinction coefficient of pure glycerol (solid line) and pure water (dashed line) at three different temperatures.

Section 2: *Theoretical THz spectra and slope*

The THz intensity (I_{THz}) can be directly calculated from MD simulations by[2]:

$$I_{THz} = \frac{2\pi\beta}{3cV}\int dt\, e^{i\omega t}\,\langle \dot{M}(t)\dot{M}(0)\rangle \qquad (S1)$$

where $\beta = \frac{1}{k_B T}$ (k_B is the Boltzmann constant and T is the temperature), c is speed of light and V is the volume of the system. $\langle \dot{M}(t)\dot{M}(0)\rangle$ is a correlation function, with $\dot{M}$ being the time derivative of the total dipole moment of the system. This can be further decomposed into molecular contributions by:

$$I_{THz} = \frac{2\pi\beta}{3cV}\int dt\, e^{i\omega t}\sum_{i=1}^{N_{mol}}\sum_{j=1}^{N_{mol}}\langle \dot{\mu}_i(t)\dot{\mu}_j(0)\rangle \qquad (S2)$$

where μ is the dipole of each molecule in the system and the sum over i and j run over all molecules in the system. The contribution to the THz spectrum of bound water molecules can be hence separately calculated from theory by restricting the summation over i to water belonging to the bound population only (N_{bound}):

$$I_{THz} = \frac{2\pi\beta}{3cV}\int dt\, e^{i\omega t}\sum_{i=1}^{N_{bound}}\sum_{j=1}^{N_{mol}}\langle \dot{\mu}_i(t)\dot{\mu}_j(0)\rangle = \sum_{i=1}^{N_{bound}} I_{bound} \qquad (S3)$$

where $I_{bound} = \frac{2\pi\beta}{3cV}\int dt\, e^{i\omega t}\sum_{j=1}^{N_{mol}}\langle \dot{\mu}_i(t)\dot{\mu}_j(0)\rangle$ is the THz spectrum of one bound water molecule. By introducing $\bar{I}_{bound}$ as the average spectrum of one bound water molecule, we hence obtain:

$$I_{THz} = N_{bound}\bar{I}_{bound} \qquad (S4)$$

The contribution of bound water molecules to the slope discussed in the main text can be hence theoretically determined by means of equation S4. To this end, we have calculated the average $\bar{I}_{bound}$spectrum for a bound water molecule H-bonded to a –OH group by means of equation S3 from a previously performed DFT-MD calculation on a model system where one alcohol (tBuOH) molecule is solvated in water, see references[2-3] for more details. The $\bar{I}_{bound}$ spectrum is shown in figure 3b of the main text (N_{bound} =1 spectrum, for which $I_{THz} = \bar{I}_{bound}$). In the same figure, we show the evolution of I_{THz} as a function of the number of bound waters (from N_{bound} = 1 to N_{bound} = 5), as calculated from equation S4. The results show a linear dependence of the slope on N_{bound} (inset of figure 3b, as expected from the linear relationship expressed by equation S4).

Section 3: *Raman measurements in the OH-stretching frequency range*

In order to corroborate our interpretation (discussed in the main text) of the molecular arrangement of glycerol-water mixtures into three water populations, i.e., bulk, bound and wrap populations, we performed Raman measurements in the OH-stretching, 2500-3700 cm^{-1} frequency range (figure S3). Pure glycerol shows three features in this frequency range: a broad OH-stretching mode around 3350 cm^{-1} and the two CH_2 stretching modes at 2886 cm^{-1} (symmetric) and 2946 cm^{-1} (asymmetric), respectively. We have analysed the spectra of the glycerol-water mixtures with the multivariate curve resolution (MCR-ALS) technique pioneered by Ben Amotz and co-workers[4] and described by Frenich et al.[5] The method requires the choice of a reference point, where both, glycerol and water have no or negligible intensity. Since both glycerol and water contribute to the OH stretch band, we choose the frequency range 2000-3700 cm^{-1} to make sure that the range includes the reference points with zero intensity lying outside the OH-stretch bands for both water and glycerol. The dissimilarity method described by Frenich et al.[5] leads to three relevant components for the dissection. Using three main components as deduced by the MCR-ALS technique yields the results shown in figure S4. The shape of the spectra and the mol fraction dependency of the weights of each of these three components (see figure S5) allows us to assign them to bulk-water, bulk-glycerol, and glycerol-water interactions, respectively.

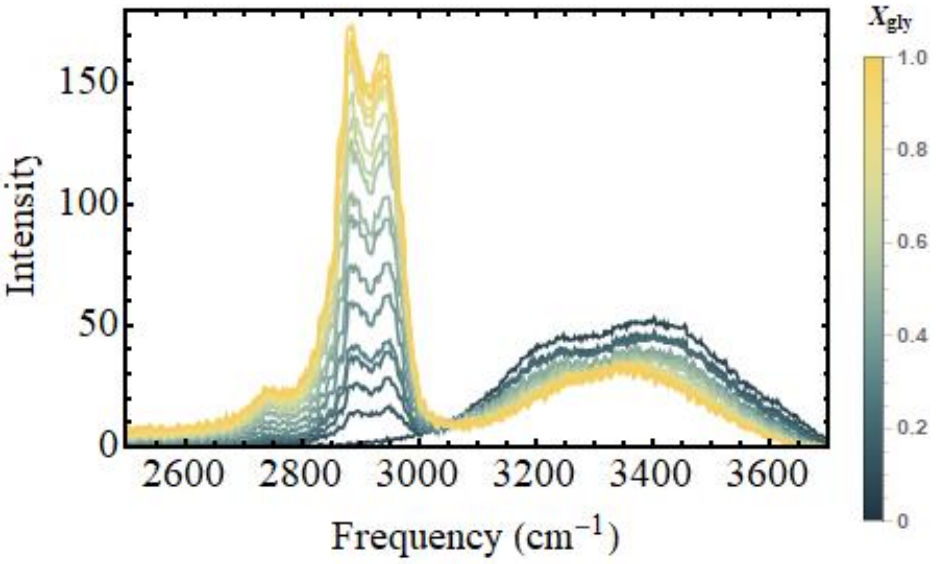

Figure S3. Raman intensity of glycerol-water mixtures.

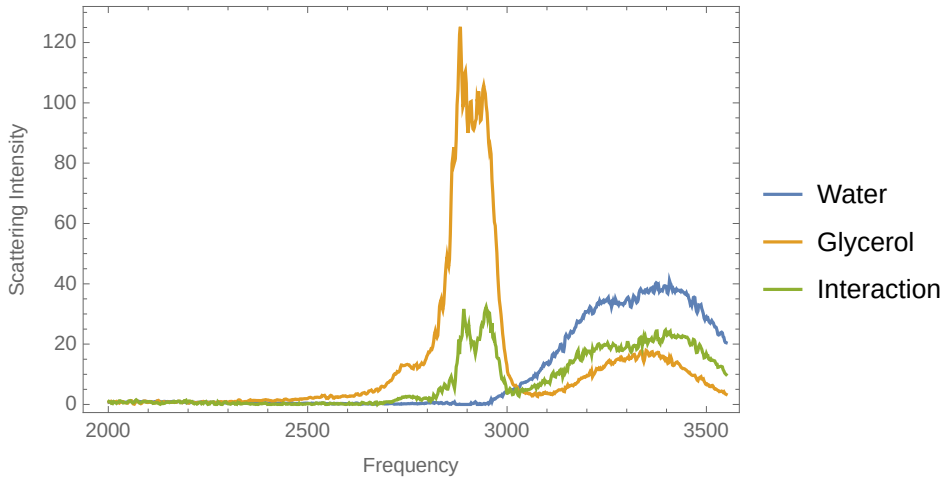

Figure S4. Three components as deduced by MCR-ALS of the Raman spectrum of glycerol-water mixtures.

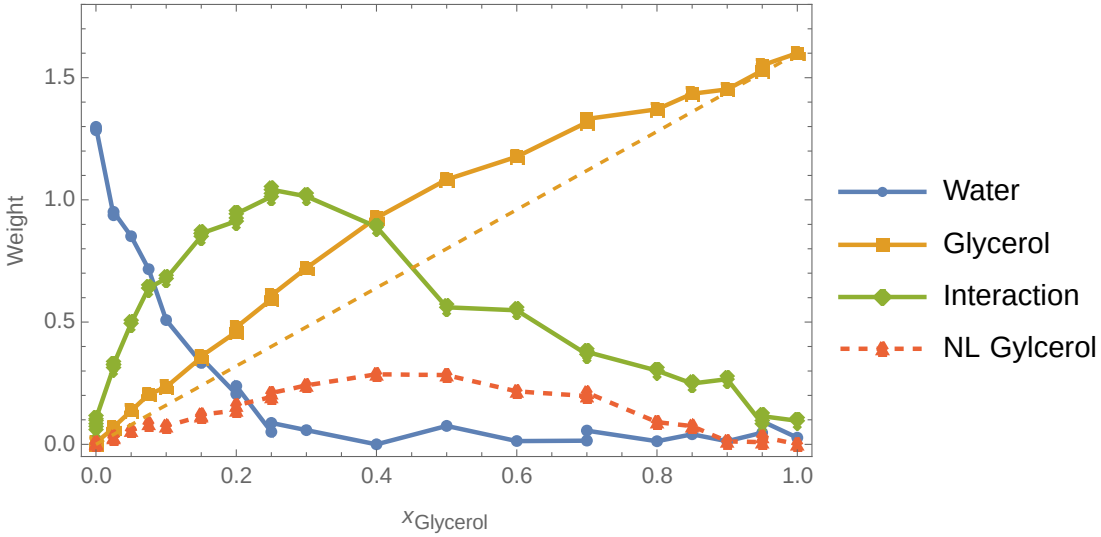

Figure S5. Relative populations of the three components as deduced by MCR-ALS as a function of X_{gly}.

The component attributed to bulk water (blue in figure S5) vanishes around X_{gly}=0.2. This is in nice agreement with the simulation data (figure 3d of the main text). By construction, this part is not present in the THz spectrum.

The water-glycerol interaction component (displayed in green) peaks at X_{gly}=0.25 which differs from the maximum of X_{gly}=0.4 that we find for bound water from our THz measurements. Interestingly, the “bulk” glycerol component (displayed in yellow) shows a nonlinear behaviour which indicates that it is not merely glycerol, but also includes water-glycerol contributions. In the following we tried to separate the linear glycerol contribution from the nonlinear water-glycerol interaction part. The dashed yellow line shows the linear interpolation between X_{gly}=0 and X_{gly}=1. If we subtract this from the yellow curve, we obtain the dashed red curve (non-linear) which peaks at X_{gly}=0.4, in agreement with the bound water

component as deduced by THz spectroscopy. We speculate that the non-linear contribution in the Raman spectrum in the frequency range of the CH stretch region can be attributed to bound water, which affects the intensity of the C-H stretch by strong interaction with glycerol.

Now the question is how do we explain that water-glycerol interaction component deduced by Raman (shown in green) differs from the bound water contribution as deduced by THz. We propose that the green "interaction" component is a spectroscopic signature of both, the "bound" water as well as the "wrap water". We found that the wrap and bound contributions as a function of X_{gly} peak at around 0.15 (from theory, figure 3d of the main text) and 0.4 (from both theory and THz measurements, figure 3), respectively. As a consequence, the maximum of the Raman scattering cross section is intermediate between the two mol fractions, i.e., at X_{gly}=0.25. In THz, both water populations, the "bound" and the "cavity wrap" population are spectroscopically well separated. In the case of glycerol, the "cavity wrap" contribution is almost bulk like in the THz range and therefore not visible in this case.

In summary, both the methods yield consistent and complementary information on the water glycerol interaction. The Raman spectra provide a signature of the wrap water population, which is predicted to be present in the simulation but missing in the THz spectra (as discussed in the main text), while the slope of the librational band in the THz/FIR region provides an unambiguous spectroscopic signature of bound water. By combination of both techniques, both contributions can be dissected.

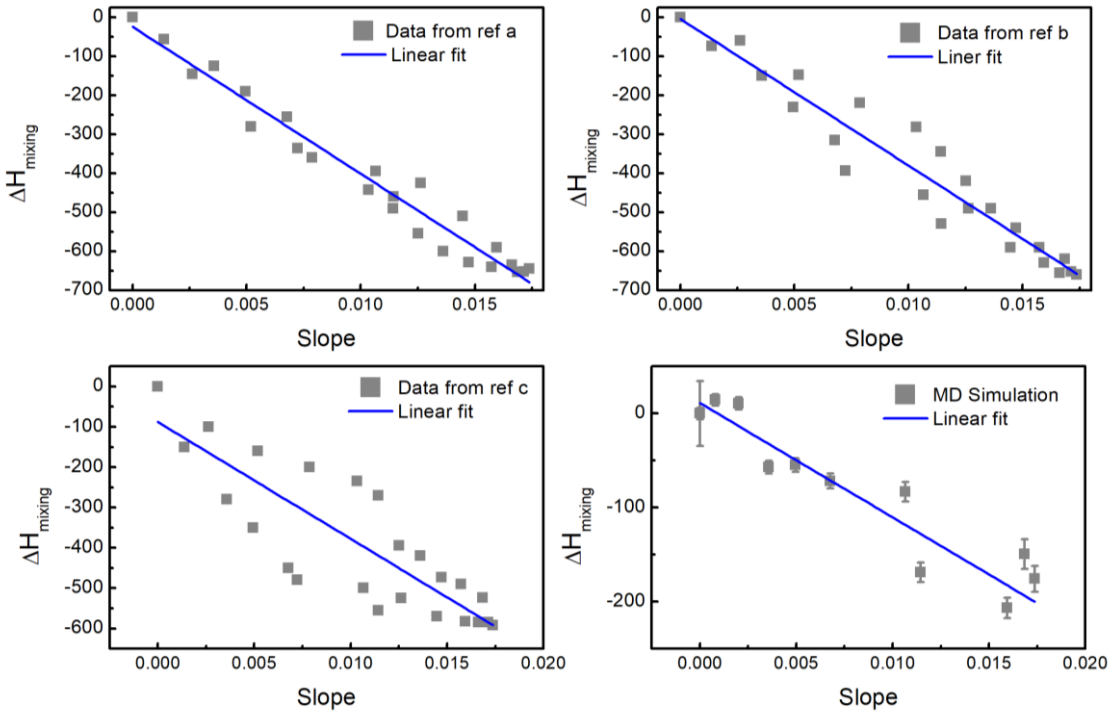

Figure S6. Linear correlation between ΔH_{mixing} data and the experimentally obtained slope. Ref a, b and c are mentioned in the main text.

Section 4: *Defining the hydration shell*

To motivate our choice of hydration shell, we compute the average local density between glycerol heavy atoms and all water oxygens

$$\rho_{gly-Ow}(r) = \frac{1}{N_{heavy}} \sum_{i=1}^{N_{heavy}} \sum_{j=1}^{N_{wat}} \delta(\vec{r} - \vec{r}_j) = \frac{1}{N_{heavy}} \sum_{i=1}^{N_{heavy}} \rho^i_{gly-Ow}(r) \qquad (S5)$$

where N_{heavy}, N_{wat}, $\vec{r}_j$, $\delta(\vec{r} - \vec{r}_j)$, and $\rho^i_{gly-Ow}(r)$ are the number of glycerol heavy atoms, the number of water oxygens, the position of the j[th] glycerol heavy atom, the Dirac delta function centred at $\vec{r}_j$, and the local water density relative to the i[th] glycerol heavy atom. We obtain the radial distribution function by dividing $\rho_{gly-Ow}(r)$ by the bulk value $\rho^{bulk}_{gly-Ow} = \rho_{gly-Ow}(r \geq 14\text{Å})$. In figure S7, we observe a systematic increase in the amplitude of $g_{gly-Ow}(r)$ peaks as X_{gly} increases, indicating a monotonic increase in the local density of water near glycerol. However, the positions of short-range ($r \leq 5$ Å) maxima and minima remain essentially constant for the entire range of concentrations. Therefore, the underlying hydration shell structure remains qualitatively similar for all considered mixtures. To further motivate our chosen cut-off radius, we investigate the impact of the cut-off radius on the characterization of bound and wrap water populations.

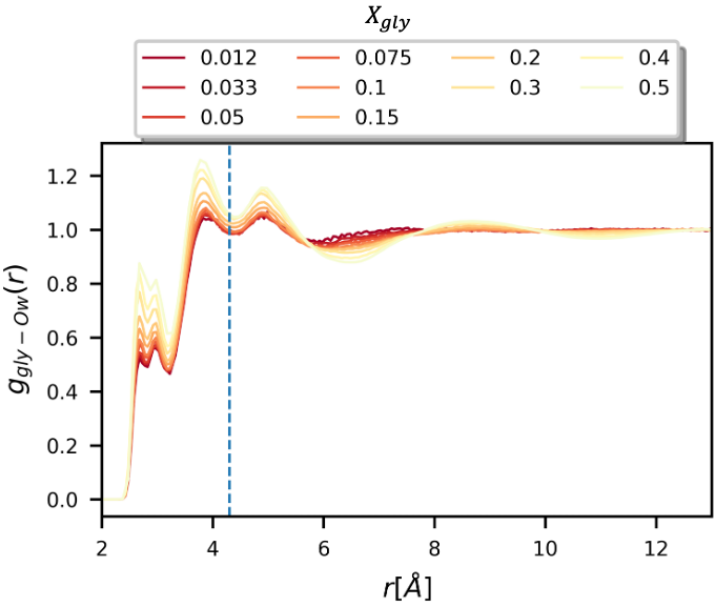

Figure S7. As the glycerol concentration increases, the amplitude of the peaks increases systematically. For r <5 Å, the location of the minima and maxima remain essentially constant

for the entire concentration range from X_{gly} =0.012 to 0.5. The 4.2 Å cut-off radius mentioned in the text is indicated here with a vertical dashed line.

Given that all bound waters must lie within 3.5 Å of a glycerol hydroxyl group atom, increasing the cut-off radius beyond 3.5 Å does not alter the bound water trend shown in figure 3d. In figure S8, we demonstrate that increasing the cut-off radius systematically from 3.5 to 5.0 Å causes an enhancement of the wrap water peak and shifting the peak position to lower X_{gly}. The wrap water behavior is otherwise consistent for all considered cut-off radii.

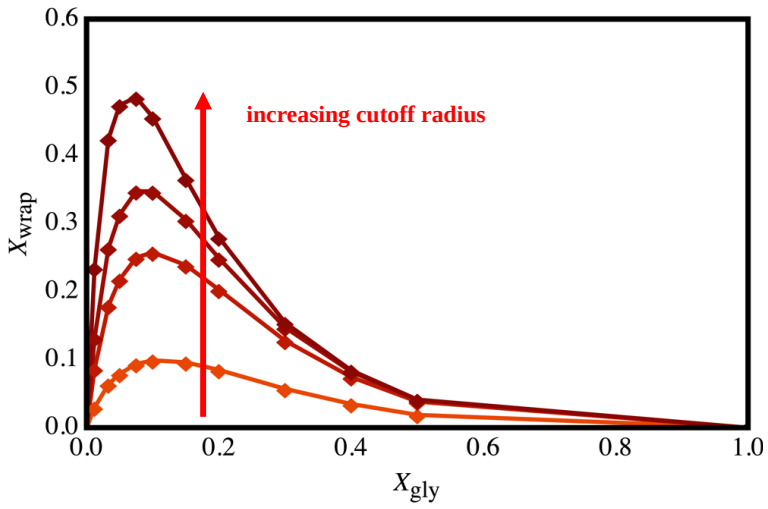

Figure S8. Here, we observe that a systematic increase in the cut-off radius (3.5, 3.9, 4.2, and 5.0 Å) yields an enhancement in the peak mole fraction of wrap waters while shifting the location of the peak towards lower X_{gly}.

Similarly, the choice of cut-off radius does not qualitatively impact the population of tetrahedrally-coordinated waters p_{tet}. For smaller cutoff radii, the wrap waters reside nearer to the glycerol molecules and hence form fewer three body angles (see *Results and Discussions* for further details). Thus, the wrap waters exhibit lower p_{tet} for a small cutoff radius (e.g., 3.5 Å) compared to a larger cut-off radius (e.g., 5.0 Å) [see figure S9].

To understand shifts in the hydration shell of glycerol with changing composition, we choose a cut-off radius that is within approximately the 2nd hydration shell of glycerol ($r \leq 4.2$ Å) while including the hydrogen bond distance criteria ($r \leq 3.5$ Å) mentioned in the *Materials and methods* section of the text. However, choosing a smaller (e.g., 3.5 Å) or larger (e.g., 5.0 Å) does not impact the essential trend in the p_{tet} of the wrap waters.

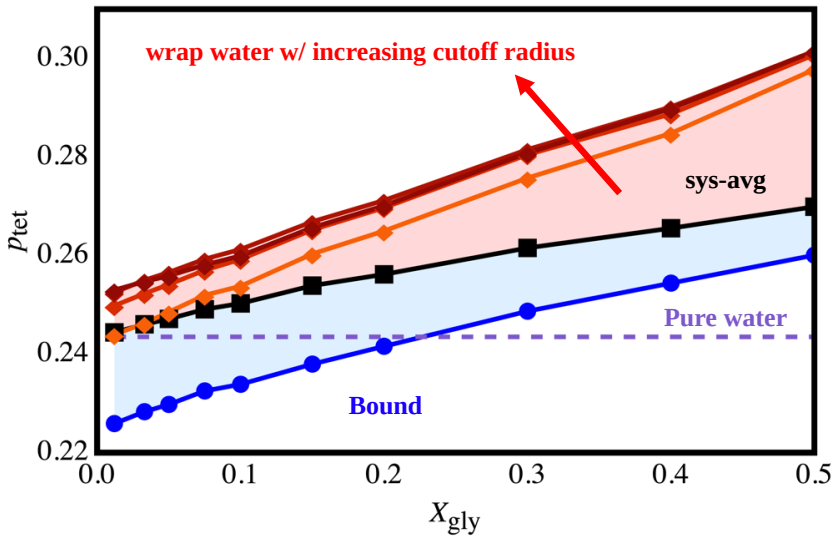

Figure S9. We observe the same trends in p_{tet} for bound, wrap, and the system-average waters as in the text (figure 4c). Here, we demonstrate that increasing the cutoff radius from 3.5 to 5.0 Å does not qualitatively impact the p_{tet} of the wrap waters.

Section 5: *Estimating standard error with block averaging*

We implement a block averaging scheme to estimate the standard error of the mean of quantities like the enthalpy of mixing (ΔH_{mixing}), the population of tetrahedrally coordinated waters (p_{tet}), and the mole fraction of bound/wrap waters as a function of glycerol concentration. As an example, we consider the block averaging procedure for the total energy of a box of OPC water molecules at 18°C and 1 bar: (1) split the time series of length N for an observed quantity into blocks of length $\tau < N/2$, (2) compute the mean of each block E_i (where $i = 1, 2, .., N/\tau$), (3) calculate the standard error of the block means s_E^τ, (4) repeat. For smaller block lengths, the individual samples (blocks) are more correlated, yielding lower standard error. We obtain the true s_E^τ (s_E) once the limiting behavior is attained (indicated with a horizontal dashed line in figure S10). After this point, s_E^τ becomes increasingly noisy due to a decreasing number of samples. To automate the calculation of s_E, we fit the data to a function

$$s_{E,fit}(\tau) = \sigma\sqrt{\frac{2}{N}}\left[aT_1\left(\frac{T_1\left(e^{-\frac{\tau}{T_1}}-1\right)}{\tau}+1\right)+(1-a)T_2\left(\frac{T_2\left(e^{-\frac{\tau}{T_2}}-1\right)}{\tau}+1\right)\right]^{1/2} \quad (S6)$$

with fit parameters a, T_1, T_2, and σ. Taking the limit as $\tau \to \infty$, we compute s_E

$$s_E = \lim_{\tau \to \infty} s_{E,fit}(\tau) = \sigma \sqrt{\frac{2}{N}} [aT_1 + (1-a)T_2]^{1/2} \qquad (S7)$$

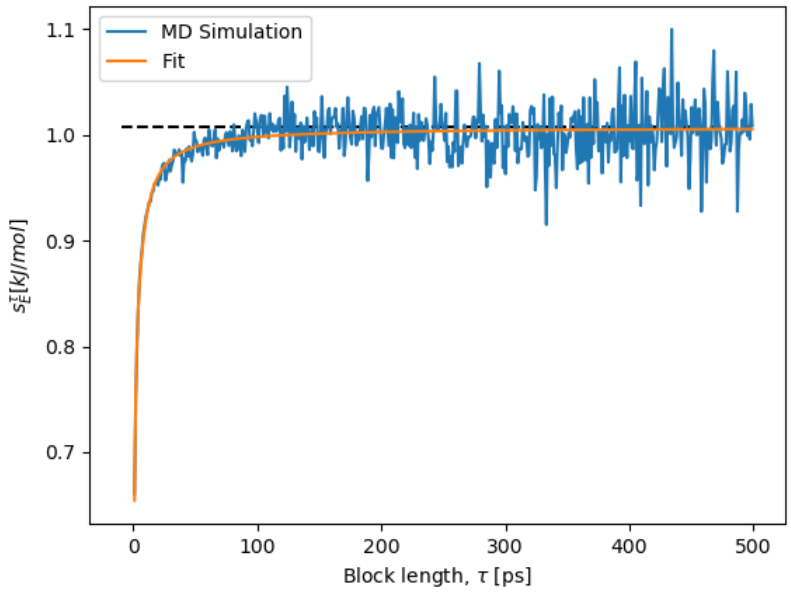

Figure S10. Here, we provide an illustrative example of the block averaging protocol for estimating the standard error, s_E, of the total energy of the system. The raw data, fit to equation S6, and limiting behavior are represented in blue, orange, and black, respectively.

Section 6: *Hydrogen bonding in glycerol-water*

The peaks in slope at X_{gly}= 0.4 [figure 3c] and the mole fraction of bound waters, X_{bound}, at X_{gly}=0.3 [figure 3d] originate partly from an increased availability of water-glycerol hydrogen bonds with increasing glycerol content. Normalizing the number of glycerol-water hydrogen bonds by the total number of hydrogen bonds $\frac{N_{HB,gly-wat}}{N_{HB,Tot}}$, we observe an apparent maximum at equimolar concentration [figure S11]. However, the location of the peak in $\frac{N_{HB,gly-wat}}{N_{HB,Tot}}$ differs from the observed maximum in slope (figure 3c) or X_{bound} (figure 3d).

Instead, we observe a monotonic increase in $\frac{N_{HB,gly-wat}}{N_{HB,Tot}}$ and $\frac{N_{HB,gly-gly}}{N_{HB,Tot}}$ and simultaneous decrease in $\frac{N_{HB,wat-wat}}{N_{HB,Tot}}$ with higher glycerol content (figure S11). These trends in hydrogen bond populations yield increasing in X_{bound} for $X_{gly} \leq 0.30$. At X_{gly}= 0.3, most water molecules are characterized as bound or wrap waters [$X_{boumd} + X_{wrap} = 0.69$]; hence, further increasing the concentration of glycerol leads to a simultaneous decrease in X_{bound} and X_{wrap}.

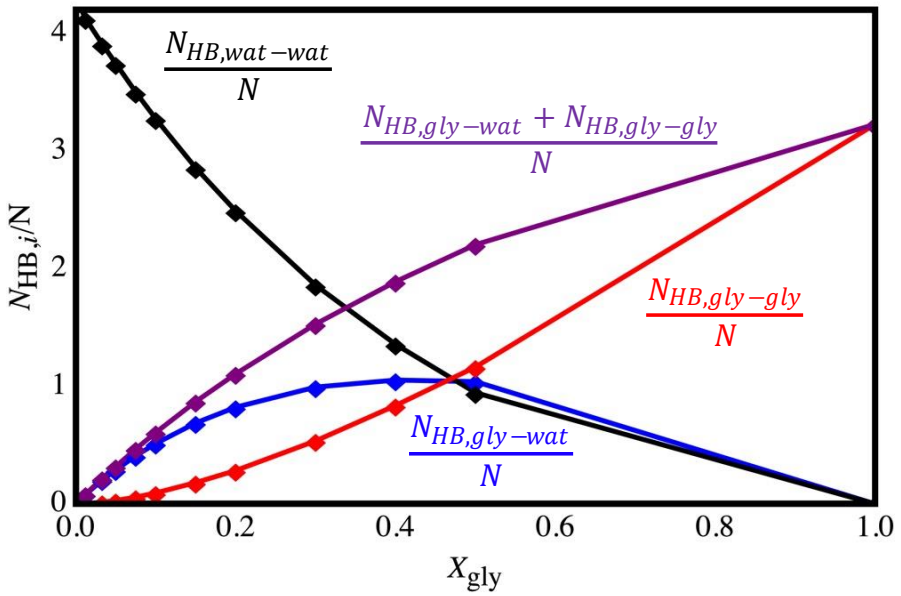

Figure S11. We depict the number of all three types of hydrogen bond pairs (i) water-water [black], (ii) glycerol-water [blue], and (iii) glycerol-glycerol [red] relative to the total number of molecules $N = N_{wat} + N_{gly}$. In purple, we show the fraction of glycerol-water and glycerol-glycerol hydrogen bonds $(N_{HB,gly-wat} + N_{HB,gly-gly})/N$.

References

1. Schwaab, G.; Sebastiani, F.; Havenith, M., Ion Hydration and Ion Pairing as Probed by THz Spectroscopy. *Angewandte Chemie International Edition* **2019,** *58* (10), 3000-3013.
2. Conti Nibali, V.; Pezzotti, S.; Sebastiani, F.; Galimberti, D.; Schwaab, G.; Heyden, M.; Gaigeot, M.-P.; Havenith, M., Wrapping up hydrophobic hydration: Locality matters. *J. Phys. Chem. Lett.* **2020,** *11* (12), 4809-4816.
3. Pezzotti, S.; Sebastiani, F.; van Dam, E. P.; Ramos, S.; Conti Nibali, V.; Schwaab, G.; Havenith, M., Spectroscopic Fingerprints of Cavity Formation and Solute Insertion as a Measure of Hydration Entropic Loss and Enthalpic Gain. *Angewandte Chemie International Edition* **2022,** *61* (29), e202203893.
4. Ben-Amotz, D., Hydration-Shell Vibrational Spectroscopy. *Journal of the American Chemical Society* **2019,** *141* (27), 10569-10580.
5. Garrido Frenich, A.; Martínez Galera, M.; Martínez Vidal, J. L.; Massart, D. L.; Torres-Lapasió, J. R.; De Braekeleer, K.; Wang, J.-H.; Hopke, P. K., Resolution of multicomponent peaks by orthogonal projection approach, positive matrix factorization and alternating least squares. *Analytica Chimica Acta* **2000,** *411* (1), 145-155.