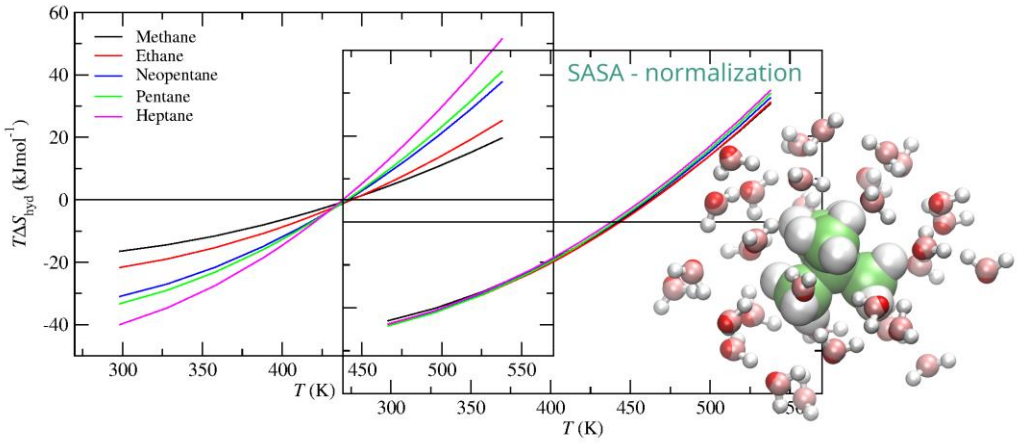
Protein Denaturation, Zero Entropy Temperature, and the Structure of Water around Hydrophobic and Amphiphilic Solutes

Kazimieras Tamoliūnas[a], Nuno Galamba[a,*]

[a] *Centre of Chemistry and Biochemistry and Biosystems and Integrative Sciences Institute, Faculty of Sciences of the University of Lisbon, C8, Campo Grande, 1749-016 Lisbon, Portugal.*

[*]Corresponding author. Electronic mail: njgalamba@fc.ul.pt



Significance

The hydrophobic effect, loosely defined as the disaffinity between oil and water, plays a pivotal role in many chemistry and biological phenomena, ranging from the low solubility of nonpolar molecules to protein folding. Here, we show, from molecular dynamics simulations, that the hydration entropy, normalized by the solvent accessible surface area, is nearly system size independent for hydrophobic, but not for amphiphilic molecules. A comprehensive molecular picture of the anomalous temperature dependence of hydrophobic hydration is then built upon information theory, through analysis of solute-water interactions and the reorganization of water's structure around hydrophobic groups. Kauzmann's "hydrocarbon model" on protein folding/denaturation is also discussed, with relatively large amphiphilic molecules exhibiting a more similar behavior to globular proteins, than hydrocarbons.

Abstract

The hydrophobic effect plays a key role in many chemical and biological processes, including protein folding. Nonetheless, a comprehensive picture of the effect of temperature on hydrophobic hydration and protein denaturation remains elusive. Here, we study the effect of temperature on the hydration of model hydrophobic and amphiphilic solutes, through molecular dynamics, aiming at getting insight on the singular behavior of water, concerning the zero-entropy temperature, T_S, and entropic convergence, also observed for some proteins, upon denaturation. We show that, similar to hydrocarbons, polar amphiphilic solutes exhibit a T_S, although strongly dependent upon solute-water interactions, opposite to hydrocarbons. Further, the temperature dependence of the hydration entropy, normalized by the solvent accessible surface area, is shown to be nearly solute size independent for hydrophobic, but not for amphiphilic solutes, for similar reasons. These results are further discussed in the light of information theory (IT) and the structure of water around hydrophobic groups. The latter shows that the tetrahedral enhancement of some water molecules around hydrophobic groups, associated with the reduction of water defects, leads to the strengthening of the weakest hydrogen bonds, relative to bulk water. However, a larger tetrahedrality is found in low density water populations, demonstrating that pure water has encoded structural information similar to that associated with hydrophobic hydration, consistent with IT assumptions. The source of the differences between Kauzmann's "hydrocarbon model" on protein denaturation and hydrophobic hydration is discussed, with relatively large amphiphilic hydrocarbons displaying a more similar behavior to globular proteins, than aliphatic hydrocarbons.

I. Introduction

The hydrophobic effect, loosely defined as the disaffinity between oil and water, plays a central role in many chemical and biological processes, including protein folding and association[1–4]. Hallmarks of hydrophobic hydration include a positive hydration free energy, ΔG_{hyd}, or excess chemical potential ($\Delta\mu^{ex}$), and an increase of the heat capacity, $\Delta C_p > 0$, also observed upon protein denaturation[3–8]. The source of this $\Delta G_{\mathrm{hyd}} > 0$ is a negative hydration entropy, ΔS_{hyd}, in spite of an also negative hydration enthalpy, ΔH_{hyd}. Hydrophobic hydration is also characterized by an unusual temperature dependence, exhibiting (extrapolated) temperatures, T_S and T_H, at which ΔS_{hyd} and ΔH_{hyd} are zero, respectively[4,5,7–10]. Thus, above these temperatures, hydration would be favored, instead, by entropy and disfavored by enthalpy. A more puzzling aspect is the fact that entropic convergence is observed at a temperature, T_S^*, close to T_S ~ 400 K, for different hydrocarbons[5,7,9,10]. Entropic convergence was also observed for some globular proteins at a temperature similar to T_S, thus, suggesting that the hydrophobic contribution to the entropy of protein unfolding is zero at this temperature[11,5,7–10,12]. The latter is apparently consistent with the analogy proposed by Kauzmann[13], between the exposure of the hydrophobic core of globular proteins to water, upon denaturation, and the transfer of a hydrocarbon from a nonpolar solvent into water. The transfer of a hydrophobic residue from water, to form the hydrophobic core of a protein, thus, interacting with other hydrophobic residues through van der Waals interactions, is then seen as the driving force (i.e., hydrophobic interactions) for protein folding[12–14]. The resemblances between this "hydrocarbon model" and protein (un)folding have, however, long been questioned[4,7,15–17]. For instance, while a small difference is observed between T_S^* and T_H^* for proteins, T_H is significantly lower than T_S, for hydrocarbons[4,10,18]. In addition, analysis of a large protein dataset by Robertson and Murphy[19] showed a non-generalized entropic convergence, suggesting that the importance of hydrophobic interactions may not be a universal facet of protein (un)folding.

Garde *et al.*[8] proposed an explanation for the convergence temperature, T_S^*, based on an information theory (IT), connecting it to the "weak temperature dependence of occupancy fluctuations for molecular scale volumes in water". They found a T_S^* ~ 410 K, slightly below T_S for model hard sphere solutes of sizes comparable to the noble gases and methane. A convergence temperature below T_S for hard spheres[20] and soft spherical solutes[21] was also previously observed. A recent molecular simulation study[22] by one of the authors, in turn, found a $T_S^* \sim T_S$ at ~475 K for various model aromatic hydrocarbons in liquid subcritical water at 100 atm.

The negative ΔS_{hyd} that characterizes hydrophobic hydration is believed to be associated with the work of cavity formation to insert the solute, whereas, the negative ΔH_{hyd} is related to solute-water interactions[23,24], although alternative pictures have been proposed[4,24–26]. The solvent excluded volume, related to the formation of the cavity, induces a reorganization of water's hydrogen bond (HB) network. This cavity work was argued to be especially large because of the small size of water molecules, as opposed to organic solvents, with larger voids, thus, involving the reorganization of less molecules to create a suitable cavity[24,27]. Nonetheless, although more sharply defined in water, the most probable size cavities in hexane and dodecane were found to be about the same size as those for water[28]. There should be an additional contribution to the entropy and enthalpy, related to the water reorganization upon solute insertion. Lee[24,27] proposed that the process of hydrophobic hydration could be broken into (a) the formation of a suitable cavity to lodge the solute and (b) the transfer of the solute into this cavity, involving the "turning on" of solute-water interactions. According to Lee[24], "changes in the thermodynamic quantities upon introduction of a cavity are entirely due to the solvent reorganization". Furthermore, the solvent reorganization process upon "turning on" solute-solvent interactions should be an exactly compensating process[24]. In addition, it has been argued that the solvent reorganization around the cavity should also be characterized by an exact entropy-enthalpy compensation[24,29–32], and the cavity formation free energy entirely attributed to the entropy decrease associated with volume exclusion[33].

While not observed through neutron diffraction experiments[34–36] a tetrahedral enhancement of some water molecules next to small hydrophobic and amphiphilic molecules has been recently observed through molecular dynamics[22,37–45], Raman scattering measurements with multivariate curve resolution[46,47], and infrared spectroscopy[48,49]. The significance of these structural changes on the hydration thermodynamics, including its temperature dependence, remains, however, poorly understood. Notice that any structural enhancement related to the reorganization of water upon cavity formation and solute insertion, while possibly not significantly contributing to the free energy (enthalpy-entropy compensation), still contributes to the hydration entropy and enthalpy. Nonetheless, disentangling the possible contributions from the cavity alone and from the solute-water interactions, to this tetrahedral enhancement, and to ΔS_{hyd} and ΔH_{hyd}, is not straightforward. Water molecules without interstitial water neighbors[50] ("water defects" associated with the existence of a fifth water neighbor in the interstitial region in-between the first and second coordination spheres) are more tetrahedral than water molecules with defects[51]. Thus, since larger cavities should be found near water molecules with a lower number of neighbors (i.e., lower density), similar to water next to nonpolar solutes, where the fifth water neighbor is already well beyond the girth of the first coordination sphere[22], this suggests that this tetrahedral enhancement

should already arise with the formation of the cavity. Explicit account of this contribution is generally neglected in solvation theories, and scale particle theory[3,24,52] (SPT), Pratt and Chandler's theory[53], and IT[8,54], quantitatively account for the ΔG_{hyd} of small hard spheres, without regarding any structural enhancement of water around the solute[55]. Thus, IT, for instance, can quantitatively describe ΔG_{hyd} for rare gases, based exclusively on the oxygen-oxygen radial distribution function (rdf) and the density of neat water[54].

We anticipate that a larger tetrahedral enhancement in pure water populations, without interstitial water molecules, was observed in this study, relative to water next to hydrophobic solutes, indicating that any structural enhancement, related with the solute, is already found in pure water due to structural fluctuations (or heterogeneities).

Following the above discussion, concerning the cavity work, ΔS_{hyd} should decrease (more negative) with the solute size. This result is consistent with molecular simulations and is reproduced by various theories, including SPT and IT. Entropic convergence implies, however, the reversal of this size dependence, with larger solutes exhibiting larger, positive, ΔS_{hyd}, at $T \geq T_S$. This, in turn, suggests that the temperature dependence of ΔS_{hyd} cannot be explained by the cavity work alone, even if suitable cavities for an arbitrarily large solute, formed in water around this temperature. Thus, although this crossover is predicted by IT for hard spheres, a molecular-level understanding of this positive entropy and its system size dependence remains elusive. The fact that the hydration entropy of polar and non-polar solutes of similar size (e.g., methane and methanol)[56] exhibit a different temperature dependence, suggests that solute-water and/or water-water interactions are important to explain $\Delta S_{\text{hyd}}(T)$, and, therefore, the molecular origin of entropic convergence in hydrophobic solutes and in some proteins. For instance, while ΔS_{hyd} of methane and methanol are similar near the melting point of water, the entropy of the former increases at a higher rate with the temperature[56].

Here, we study the hydration thermodynamics of model hydrocarbons and amphiphilic molecules, as simple prototypes of the core of globular proteins, aiming at understanding the molecular source of T_S and entropic convergence. Further, solvation analysis is carried out to probe the effects of temperature and solute size on the HB network of water, next to hydrophobic and amphiphilic solutes, to assess possible reasons associated with the inversion of the hydration entropy dependence on the solute size, above the entropy convergence temperature.

II. Methods

Molecular dynamics (MD) simulations of various OPLS-aa[57] model hydrocarbons and alcohols

in water were performed with the program GROMACS 5.1.4[58]. The following solutes were studied: methane, ethane, neopentane, pentane, heptane, benzene, naphthalene, anthracene, pyrene, methanol, ethanol, and neopentyl alcohol (neopentanol), chosen to include aliphatic and aromatic hydrocarbons as well as amphiphilic solutes of similar sizes. The hydration free energies of some solutes were first calculated at 298 K and 1 atm for distinct water models, namely, TIP3P[59], SPC/E[60], TIP4P/Ew[61], and TIP4P/2005[62], for comparison purposes. Although the TIP3P water model provides in general more accurate ΔG_{hyd} values, the TIP4P/2005 water gives the most accurate free energies among the other water models (see Table S1). Thus, because it provides a better description of liquid water, than TIP3P, the simulations were carried out with the TIP4P/2005 model.

To directly probe the hydration free energy maxima, avoiding extrapolation at high temperatures, especially for alcohols, for which $T_S >> T_b$, where T_b = 373 K is the normal boiling point of water, the simulations were carried out at 100 atm, allowing observing this maximum below T_b ~580 K, at this pressure, for most systems. While T_S is expected to decrease with the size of the solute, prohibitively large solutes would be required to observe a T_S close to T_b, at 1 atm, for amphiphilic solutes. Furthermore, whereas free energy simulations may be performed along the experimental liquid–vapor coexistence density, T_S^*, for spherical model solutes, has been shown to decrease by 60 K, at 1 atm, relative to water simulated at the coexistence density[21]. The effect of the pressure increase (100 atm), as discussed below for methane, is a shift of T_S by ~30 K to higher temperatures, relative to 1 atm, whereas a mild density increase is observed. Thus, in spite protein denaturation entropies and, therefore, T_S^* are not available at 100 atm, we believe a similar qualitative behavior should be found for protein denaturation at this pressure, as well as for large hydrophobic and amphiphilic solutes at 1 atm.

The aqueous systems were comprised of a single molecule of solute and 1000 water molecules in a cubic box with periodic boundary conditions. The systems were first equilibrated in the (N,P,T) ensemble for 10 ns after an 100 ps simulation in the (N,V,T) ensemble. The T and p were controlled with the thermostat of Bussi *et al.*[63] and the Parrinello-Rahman barostat[64], and the equations of motion were solved with the Verlet leap-frog algorithm with a 2 fs time-step. Electrostatic interactions were computed via the particle-mesh Ewald (PME) method[65]. A cut-off of 1 nm was used for non-bonded van der Waals and for the PME real space electrostatic interactions. Heavy atom-hydrogen covalent bonds were constrained with the LINCS algorithm[66].

The hydration free energies were obtained through “alchemical” free energy calculations[67] with the Bennett acceptance ratio (BAR)[68] method. The method involves the perturbation of the system based on the definition of a parameter, λ, taking values in the interval [0,1], allowing connecting the

end states of interest, A (λ=1) and B (λ=0), defined by the Hamiltonians $\mathbf{H}_A(r,p;\lambda)$ and $\mathbf{H}_B(r,p;\lambda)$. The transition from state *A*, the solution, to state B, the solvent, is performed by a number (N_λ) of different values of λ, corresponding to non-physical states. The ΔG_{hyd} calculated in this work, concern the transfer of a solute from a fixed position in the gas phase to a fixed position in water, following Ben-Naim and Marcus[69] standard, as opposed to the transfer from a non-polar environment to water.

A decoupling approach was used with $N_\lambda = 20$, connecting the states *A* and *B*. For the Coulombic interactions decoupling, a $\Delta\lambda = 0.25$ was adopted, whereas for the van der Waals interactions a $\Delta\lambda = 0.05$ was used. Langevin stochastic MD[70] were carried out and a soft-core potential was used for Lennard-Jones and electrostatic interactions to avoid numerical singularities at terminal λ values, with $\alpha = 0.5$, $\sigma = 0.3$, and a soft-core power of 1[67,71–74]. The simulations for each λ consisted of a steepest descent energy minimization step, followed by a 0.5 ns Langevin *NVT* simulation, and a 1 ns Langevin simulation in the *NPT* ensemble, using the Parrinello-Rahman barostat[64]. The hydration free energy was then computed from 2 independent Langevin *NPT* simulations, 10 ns long, for each λ. For some temperatures, where larger differences were observed, typically near the ΔG_{hyd} maxima, up to 5 independent simulations, 10 ns long, were carried out.

The entropy was assessed from, $\Delta S_{\text{hyd}} = -\left(\partial \Delta G_{\text{hyd}} / \partial T\right)_p$, where ΔG_{hyd} was fitted to a second order polynomial, and the hydration enthalpy was estimated from, $\Delta H_{\text{hyd}} = \Delta G_{\text{hyd}} + T\Delta S_{\text{hyd}}$. Every approach to assess either ΔH_{hyd} or ΔS_{hyd}, suffers from limitations[67,75] and, therefore, a larger uncertainty is associated with the hydration entropy and enthalpy.

Solvent accessible surface areas (SASA) were computed by rolling a solvent sphere[76,77] of radius 1.4 Å over the van der waals surface[78] of the solutes.

The tetrahedrality of water was assessed through the calculation of the orientational order parameter[79], *q*, in the rescaled form[80], $q = 1 - \frac{3}{8}\sum_{i=1}^{3}\sum_{j=i+1}^{4}\left(\cos\theta_{ij} + 1/3\right)^2$, where θ_{ij} is the angle formed by the lines joining the O atom of a given water molecule and those of its nearest neighbors, *i* and *j*. The average value of *q* varies between 0 (ideal gas) and 1 (perfect tetrahedral HB network).

Solvation water molecules, that is, those in the first hydration shell of the solutes, defined by the first minimum of the rdf, are separated in water molecules with 4 or more water neighbors (4MWN) and water molecules with less than 4 water neighbors (L4WN). The first population remains with a nearly tetrahedral coordination, similar to bulk water, whereas the second can form up to a maximum of 3 HBs, resembling water molecules near a water-vapor interface[81–84]. The latter

population (L4WN) is defined by water molecules that are closer to the solute (heavy atoms or hydrogen atoms) than to a fourth water neighbor, whereas the former (4MWN) is comprised by water molecules that are closer to any four water neighbors than to the solute[40]. The tetrahedrality was calculated by sampling each population every 50 fs from 20 ns simulations in the *NPT* ensemble. The tetrahedrality of bulk water was further disentangled in water molecules with interstitial water molecules (IWM) and no interstitial water molecules (NIWM), where an interstitial water molecule was defined as a fifth water neighbor found up to 3.7 Å[85], around a central water molecule.

III. Results

1. Thermodynamic Parameters

The temperature dependence of ΔG_{hyd}, ΔS_{hyd}, and ΔH_{hyd} for the aliphatic hydrocarbons is shown in Fig. 1. The temperature of the ΔG_{hyd} maximum, ~440 K, is nearly solute size independent, consistent with experimental data extrapolations at 1 atm, although for the transfer of a hydrocarbon from the pure liquid, rather than from the gas phase, into water[5,7]. A minor effect of the pressure on ΔG_{hyd} of methane is observed, indicating that our results should be comparable to those extrapolated at 1 atm; a similar effect was observed for the other hydrocarbons at 298 K. T_S is the temperature at which ΔG_{hyd} displays a maximum and, thus, $\Delta S_{\text{hyd}} = 0$. For ΔG_{hyd} of the form of a second order polynomial (see Fig. 1(a_1)),

$$\Delta G_{\text{hyd}} = a + bT + cT^2 \qquad (1)$$

this temperature is given by $T_S = -b/2c$. The T_S obtained through this equation for methane, ethane, neopentane, pentane, and heptane are, respectively, 442K, 444K, 441K, 441K, and 439 K. Thus, entropic convergence is observed at $T_S^* \sim 440 \pm 10$ K, marginally lower, although within the range of T_S, because of the T_S near solute size independence.

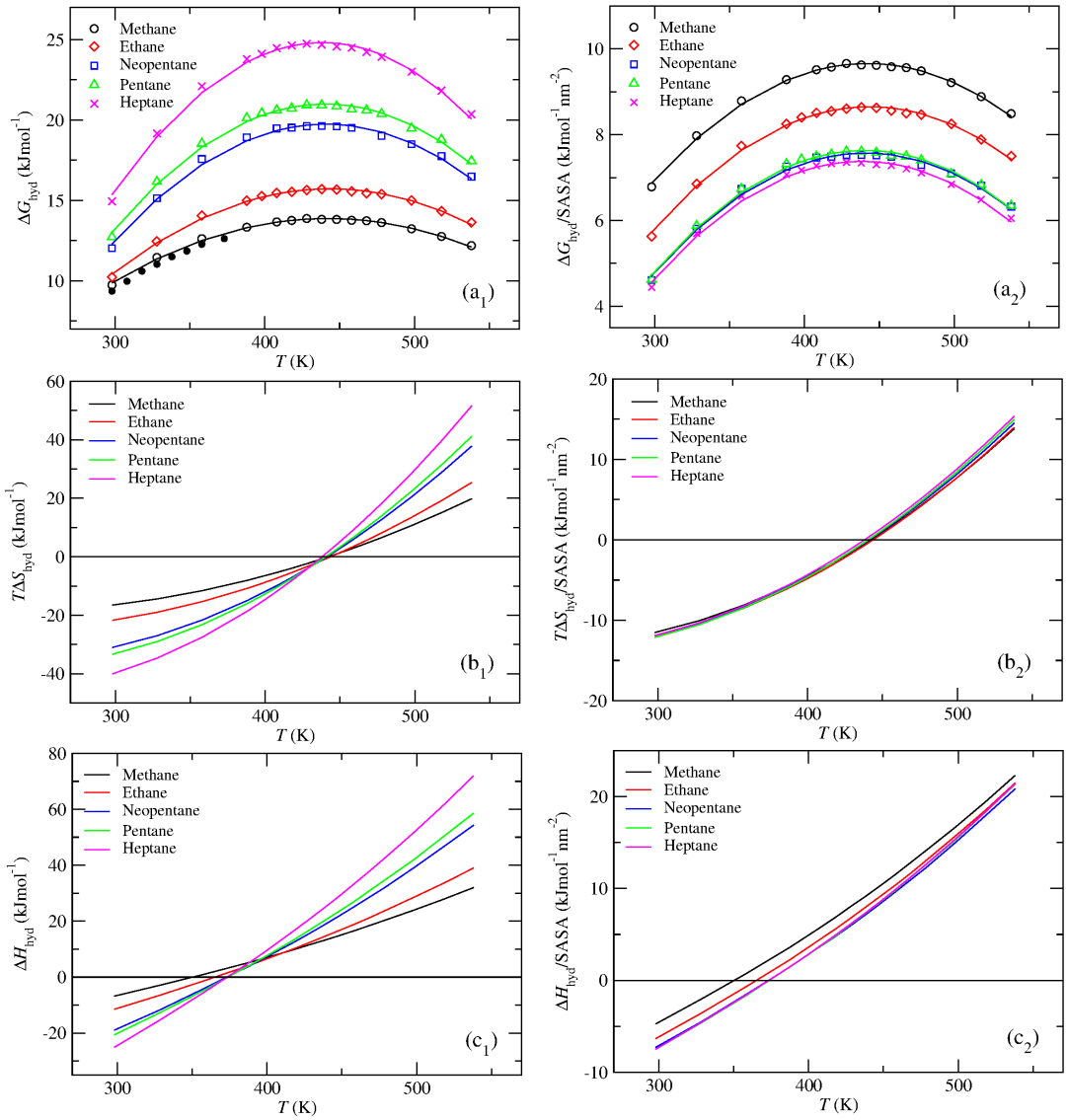


Figure 1 – Temperature dependence of (a$_1$) ΔG_{hyd}, (b$_1$) ΔS_{hyd}, and (c$_1$) ΔH_{hyd}, for the distinct hydrocarbons, at 100 atm. The ΔS_{hyd} convergence region extends around ~440±10 K. The filled circles in (a$_1$) are for methane at 1 atm. The entropy was obtained from the temperature derivative of the free energy, fitted to a second order polynomial, displayed as solid curves in (a$_1$). The standard deviations of the free energy obtained from the independent simulations are comparable to the symbols' size and are omitted. The same plots, normalized by the SASA (see Table S2) are shown, respectively, in (a$_2$), (b$_2$), and (c$_2$).

The temperature dependence of ΔS_{hyd} and ΔH_{hyd} for methane at 1 atm was also assessed, allowing comparing with available "experimental" (see Fig. S1) data[56,86,87]. A good agreement is observed. Furthermore, a T_S of 412 K (extrapolated), close to the experimental, ~400 K, was found, at 1 atm.

Fig. 1(b$_1$) shows that the hydration entropy (i.e., $T\Delta S_{hyd}$) at 298 K decreases (more negative) with the solute size, consistent with the fact that a larger cavity must form to host larger solutes. However, the rate of increase of the entropy with the temperature increases with the solute size, and above T_S^* an inversion occurs, with ΔS_{hyd} increasing with the solute size. Thus, the fact that ΔS_{hyd}

increases with the solute size, above T_S^*, and that $\Delta S_{\text{hyd}} > 0$ above T_S, suggests that ΔS_{hyd} should no longer be governed by the cavity work.

ΔH_{hyd} exhibits a similar behavior, although T_H increases with the solute size, opposite to T_S. Furthermore, T_H is lower than T_S for every solute and convergence occurs at T_H^* = 390±20 K << T_S^* and $\Delta H_{\text{hyd}}(T_H^*) > 0$. The solute size dependence of the rate of increase of the hydration enthalpy with the temperature (i.e., $\Delta C_p = \left(d\Delta H_{\text{hyd}} / dT\right)_p$) is consistent with the temperature dependence of the solute-water van de Waals interactions (see Fig. S2). However, unlike ΔH_{hyd}, solute-water van der Waals interactions do not converge at T_H^*, suggesting that the reversal of the solute size dependence of ΔH_{hyd} should be further connected with water-water interactions. Electrostatic interactions are significantly weaker and appear to exhibit a convergence behavior with the temperature; noteworthy, for methane and ethane, the solute-water Coulombic repulsion first decreases with the temperature, opposite to the larger solutes (see Fig. S2).

Figure 1 also shows the same thermodynamic parameters normalized by the solvent accessible surface areas (SASA). Remarkably, $T\Delta S_{\text{hyd}} / SASA$ exhibits a nearly solute size independent behavior, while the rate of increase of ΔH_{hyd} and solute-water van der Waals interactions (see Fig. S2(b)), normalized by the *SASA*, are now more similar for the distinct solutes. Notice that, in addition to a curvature homogenization of $\Delta G_{\text{hyd}} / SASA$, a reversal of the system size dependence occurs (see Figs 1(a1) and 1(a2)). Figure 1(b2) shows that the SASA allows normalizing the solute size dependence of the entropy in both regimes, $\Delta S_{\text{hyd}}(T < T_S) < 0$ and $\Delta S_{\text{hyd}}(T > T_S) > 0$. However, because the SASA encloses information on the volume of the cavity and on the extent of solute-water interactions and HB perturbations, it does not allow distinguishing between the temperature dependence contributions associated with the cavity work and solute-water interactions.

Notice that ΔG_{hyd} grows linearly with the solute volume, for small solutes, whereas for large hydrophobic surfaces (> ~1 nm; not studied in this work) it grows linearly with the surface area[2]. This behavior is believed to be associated with the formation of a liquid-vapor like interface next to large solutes, with water molecules moving away from the solute (i.e., microscopic dewetting)[2], opposite to small solutes. The structural transformations of water and their putative connection with the distinct rates of increase of ΔS_{hyd} and ΔH_{hyd}, with the temperature, are discussed in sub-section 2. Before, however, we discuss the results for the model aromatic hydrocarbon and alcohol aqueous systems.

Figure 2 shows similar plots for the aromatic hydrocarbons, exhibiting a T_S^* ~ 485 K (except for

pyrene), thus, ~ 45 K larger than for the aliphatic hydrocarbons. This temperature is similar to that recently observed by one of the authors[22] for the generalized amber force field[88] model of benzene, naphthalene, and anthracene, in TIP4P-Ew[61] water. The T_S obtained from the second order polynomial fit for benzene, naphthalene, anthracene, and pyrene, are respectively, 469 K, 472 K, 474 K, and 476 K. Thus, opposite to aliphatic hydrocarbons, $\Delta S_{\text{hyd}}(T_S^*) > 0$, consistent with the experimental T_S^* for some aromatic and cyclic hydrocarbons[5]. In addition, $T_H^* > T_S^*$, opposite to aliphatic hydrocarbons.

A similar qualitative behavior to that displayed in Fig. 1 is observed for ΔS_{hyd} and ΔH_{hyd}, normalized by the SASA. Generalized curves for $\Delta S_{\text{hyd}} / SASA$ of aliphatic and aromatic hydrocarbons are given in Fig. S3 and Table S3. The higher rate of increase of ΔH_{hyd} with the solute size is again consistent with the respective rate of increase of the solute-water van der Waals interactions (see Fig. S4). However, electrostatic interactions are now more important, significantly contributing to ΔH_{hyd} and, therefore, to ΔG_{hyd}. Thus, opposite to aliphatic hydrocarbons a reversal of the system size dependence is not observed for $\Delta G_{\text{hyd}} / SASA$, relative to ΔG_{hyd}. This is because ΔG_{hyd} is now governed by ΔH_{hyd}, via solute-water interactions, and a significantly larger difference between ΔH_{hyd} and ΔG_{hyd}, but not ΔS_{hyd}, is observed among the solutes. Notice that the solute size dependence of $\Delta G_{\text{hyd}} / SASA$, unlike ΔG_{hyd}, is governed by $\Delta H_{\text{hyd}} / SASA$, for both aliphatic and aromatic hydrocarbons, in the sense that $\Delta S_{\text{hyd}} / SASA$ is nearly solute size independent. For pyrene ($C_{16}H_{10}$), electrostatic interactions are less repulsive than for anthracene ($C_{14}H_{10}$), because of the higher rate of carbon/hydrogen atoms, explaining the similar rate of increase of ΔH_{hyd} and, thus, the non-convergence of the enthalpy at high temperatures (see Fig. 2(c1)).

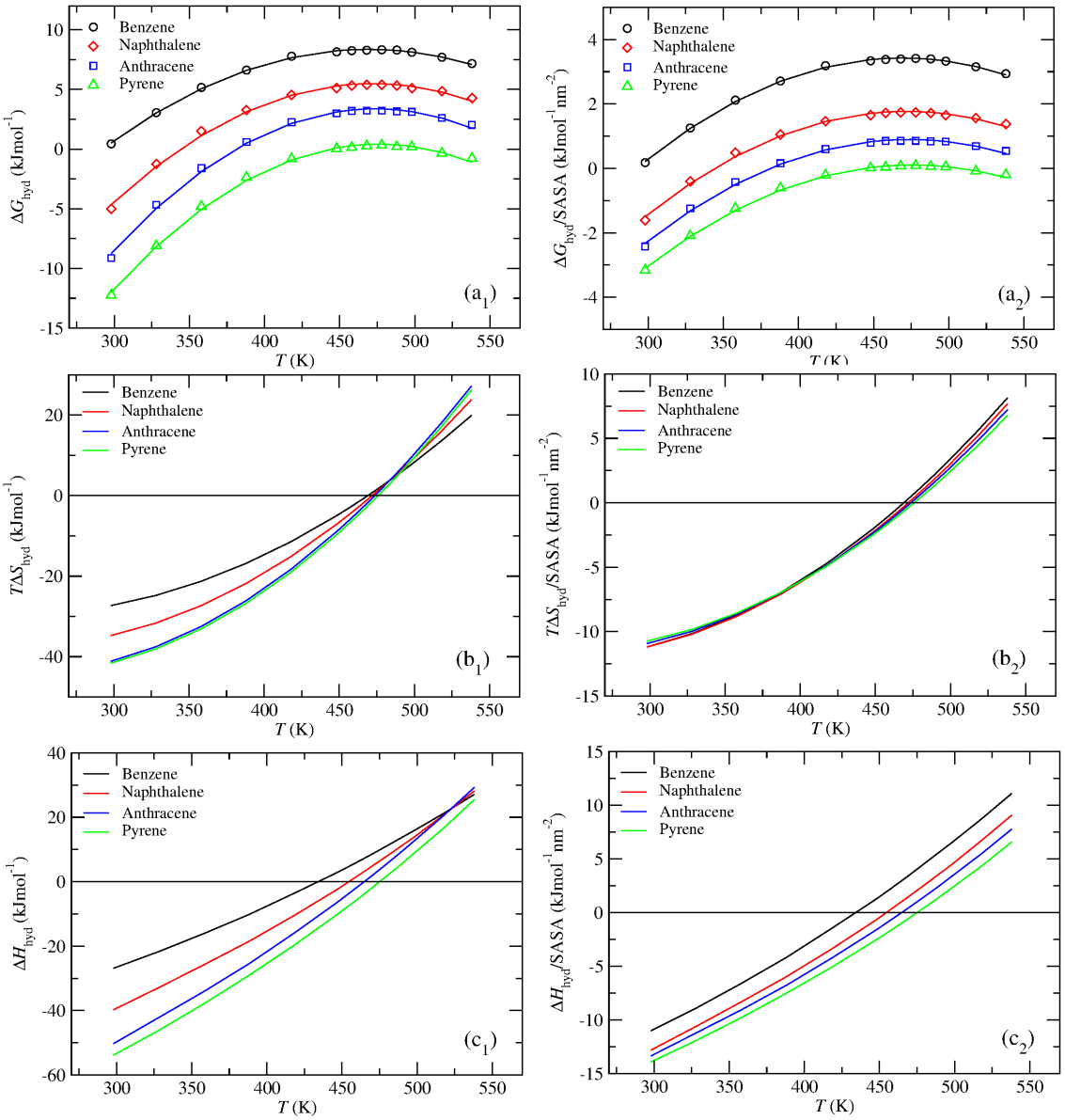


Figure 2 - Temperature dependence of (a$_1$) ΔG_{hyd}, (b$_1$) ΔS_{hyd}, and (c$_1$) ΔH_{hyd}, of benzene, naphthalene, anthracene, and pyrene at 100 atm. The entropy was obtained from the temperature derivative of the free energy fitted to a second order polynomial displayed as solid curves in (a$_1$). The same plots, normalized by the SASA (see Table S2) are shown, respectively, in (a$_2$), (b$_2$), and (c$_2$).

The ΔS_{hyd} increase with the temperature for both aliphatic and aromatic solutes indicates a facilitated insertion of the solute, which could be explained by the larger number of broken HBs in water, increasing the size of the cavities, and thus, the probability of finding a suitable cavity to lodge the solute. For hard spheres this probability is related to the hydration free energy, by $\Delta G_{hyd} = -k_B T \ln p_0(R)$, where R is the solvent accessible radius, given by $R = R_S + R_W$, and R_S and R_W are the radius of the solute and water, modeled as hard spheres, whereas $p_0(R)$ is the probability that a sphere of radius R randomly inserted in water is devoid of water molecules[3,8,89].

Garde *et al.*[8] analyzed the entropy convergence based on an information theory where the hydration free energy of small hard spheres is approximated by a function of $T\rho^2(T)$,

$$\Delta G_{\text{hyd}}^{\text{HS}} = T\rho^2\left\{k_B v^2 / 2\sigma^2\right\} + T\left\{k_B \ln(2\pi\sigma^2)/2\right\}$$
$$\Delta G_{\text{hyd}}^{\text{HS}} = T\rho^2(T)x(v) + Ty(v) \quad (2)$$

Equation (2) was obtained from the probability of finding exactly *n* solvent molecules in the cavity volume $v = 4\pi R^3/3$, modeled by a Gaussian distribution, $p_n \approx \exp(-\delta n^2/2\sigma^2)/\sigma\sqrt{2\pi}$, with variance, $\sigma^2 = \langle \delta n^2 \rangle$ and $\delta n = n - \langle n \rangle$. From the second equation, *x*(*v*) and *y*(*v*) only depend on the excluded volume.[8] This equation follows from the observation by Garde *et al.*[8] of a weak temperature dependence of occupancy fluctuations for solute excluded volumes in water, $(\sigma^2(T,v) \approx \sigma^2(v))$ not expected in non-polar organic solvents. The latter offers an explanation to the singular ability of water to exhibit a zero-entropy convergence temperature, contradicting the view that significantly larger voids should form in water associated with a more broken HB network. Notice the above equation does not depend on any specific structural transformation of water around the solute, neither associated with volume exclusion nor with solute-water interactions, but rather only on the density and density fluctuations of water with the temperature. The hydration or excess entropy obtained from eq. (2) gives,

$$\Delta S_{\text{hyd}} = \rho^2 x(v)(2T\alpha - 1) - y(v) \quad (3)$$

where α is the thermal expansion coefficient. Now, the temperature of zero entropy is given by,

$$T_S^{HS} = \frac{y(v) + \rho^2 x(v)}{2\rho^2 x(v)\alpha} = \frac{b^{HS}}{2c^{HS}} \quad (4)$$

which has a similar form to the empirical T_S obtained from eq. (1), i.e., $T_S = -b/2c$.

The empirical parameters *b* and *c* in eq. (1) contain information on the density fluctuations of water as well as on the solute-water interactions, absent in b^{HS} and c^{HS}. The inclusion of van der Waals interactions is expected to shift T_S^{HS} to higher temperatures[3]. However, the normalization of ΔS_{hyd}, defined through eq. (3), by the SASA, should still exhibit a solute size dependence, nearly absent in hydrocarbons. This indicates that $\Delta S_{\text{hyd}}(T)$ should depend not only on the size of the cavity and density fluctuations, accounted by IT, but also on solute-water interactions, both related with the SASA. In addition, while T_S is nearly solute size independent (for small solutes), it should be determined by water's density fluctuations and solute-water interactions; this point is further discussed below. For larger solutes (not studied here) a decrease of T_S is expected, since T_S decreases monotonically with the solute radius, for spherical solutes.[3,15,21]

Turning attention to the model amphiphilic solutes, Fig. 3 shows the thermodynamic functions for methanol, ethanol, phenol, and neopentanol. The temperature dependence of ΔS_{hyd} and ΔH_{hyd} for methanol, at 1 atm, was also assessed, allowing comparing with available "experimental" (see

Fig. S1) data[56,90,91]; a reasonable agreement is found, in spite of significant differences in the curvature of ΔS_{hyd}.

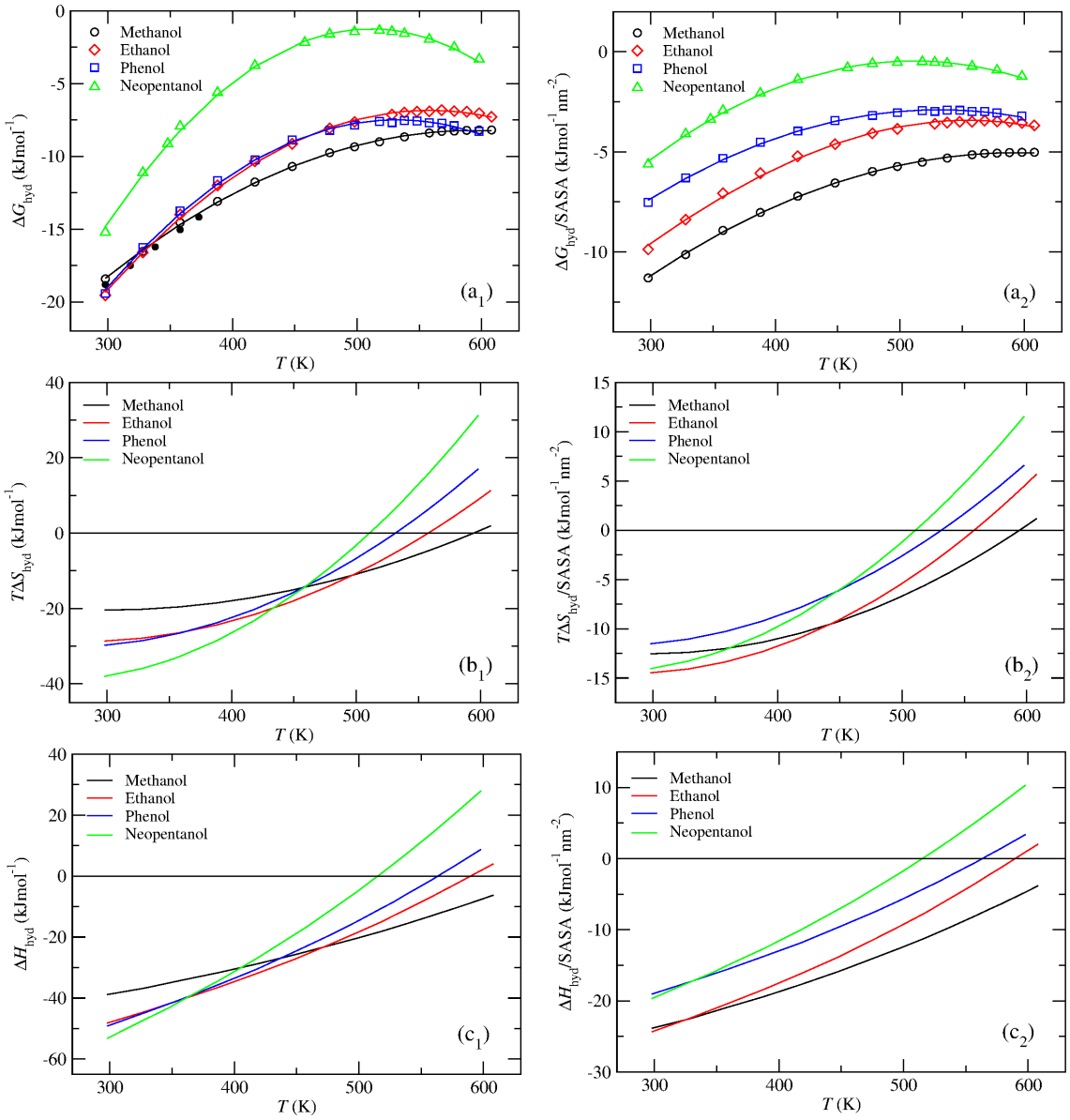


Figure 3 - Temperature dependence of (a$_1$) ΔG_{hyd}, (b$_1$) ΔS_{hyd}, and (c$_1$) ΔH_{hyd} of methanol, ethanol, phenol, and neopentanol at 100 atm. The filled circles in (a) are for methanol at 1 atm. The entropy was obtained from the temperature derivative of the free energy fitted to a second order polynomial displayed as solid curves in (a$_1$). The same plots, normalized by the SASA (see Table S2) are shown, respectively, in (a$_2$), (b$_2$), and (c$_2$).

The solutes exhibit significantly different T_S, opposite to hydrocarbons. Furthermore, T_S are larger than for the hydrocarbons. The T_S obtained from the second order polynomial fit for methanol, ethanol, phenol, and neopentanol, are respectively, 594 K, 557 K, 531 K, and 511 K.

Entropic convergence can be observed at T_S^*~458 K, with the exception of ethanol. Entropic convergence was previously observed for some linear alcohols from analysis of experimental data

at 1 atm[32]; interestingly, convergence was also observed at negative values of ΔS_{hyd}, and no convergence was observed between methanol and ethanol up to 400 K[92]. Electrostatic interactions are now dominant (see Fig. S5) and these are stronger for methanol and ethanol than for phenol and neopentanol, opposite to van der Waals interactions.

In spite of the apparent solute size dependence (see Fig. 3(b1)), this is not the reason behind the significantly different T_S, since a near system size independence was already demonstrated for hydrocarbons. To further demonstrate this point we calculated ΔG_{hyd}, ΔS_{hyd}, and ΔH_{hyd} for neopentane, benzene, and neopentanol, neglecting electrostatic interactions. Thus, non-bonded solute-water interactions were restricted to van der Waals interactions. Figure 4 shows that a similar T_S is now obtained for the distinct solutes. T_S values of 442 K, 449 K and 442 K were found, respectively, for neopentane, benzene, and neopentanol modeled exclusively by van der Waals interactions. This indicates that the different T_S are connected with the solute-water interactions, especially, electrostatic interactions, almost absent in aliphatic hydrocarbons.

Furthermore, the solute size dependence of $\Delta S_{hyd} / SASA$, not observed for hydrocarbons, is connected with the fact that electrostatic interactions are less sensible to the SASA, since these are long ranged.

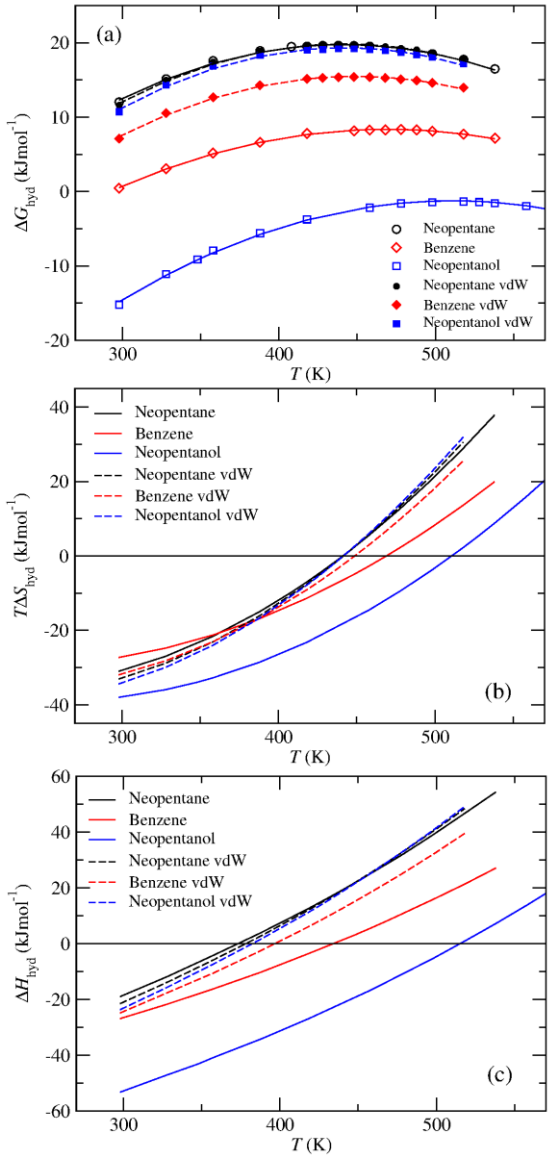


Figure 4 - Temperature dependence of the (a) ΔG_{hyd}, (b) ΔS_{hyd}, and (c) ΔH_{hyd} of model neopentane, benzene, and neopentanol at 100 atm, with and without (vdW) including solute-water electrostatic interactions.

Now, opposite to the aliphatic hydrocarbons it can be seen that T_H is larger than T_S for the amphiphilic solutes. A distinction between protein denaturation and the "hydrocarbon model" is the fact that a much smaller difference is observed between T_S^* and T_H^* for proteins depicting a T_S^* similar to T_S for hydrocarbons, relative to the difference between T_S and T_H for the latter[4,10,18]. This reflects the differences between a hydrophobic group in the core of a globular protein and a hydrocarbon dissolved in a nonpolar solvent. As can be seen in Figs 1-3, $T_S > T_H$ for aliphatic hydrocarbons, $T_S \sim T_H$ for the larger polycyclic aromatic hydrocarbons and neopentanol, and $T_S < T_H$ for the smaller alcohols. Thus, although $T_S \sim T_H$ for neopentanol and the aromatic solutes studied, should be larger than $T_S^* \sim T_H^*$, for proteins, at 100 atm, the former decreases with the "size" of the hydrophobic group, via a weakening of solute-water interactions, suggesting that the

thermodynamics of protein (un)folding, at 1 atm, should resemble more closely the thermodynamics of solvation of an amphiphilic molecule, bearing a large hydrophobic group, than the solvation of a hydrocarbon. The latter is consistent with the fact that the core of globular proteins is not exclusively populated by aliphatic residues.

2. Molecular Solvation Analysis

In the previous section we analyzed the differences between the temperature dependence of hydrophobic and amphiphilic hydration, with respect to solute-water interactions and solute size dependence. These solute-water interactions, and water-water interactions (neglected until now), depend on the structural transformations of water near the solutes. Thus, we now discuss such structural transformations.

Figure 5(a) shows a tetrahedrality enhancement of water molecules with 4 or more water neighbors (4MWN), next to neopentane, relative to bulk water; similar results were observed for the other solutes. An even larger tetrahedrality is observed, however, for bulk water molecules with no interstitial water molecules (NIWM) (see Fig. 5(b)). A comparison between these populations in bulk water at 1 atm and 100 atm is given in Fig. S7, displaying similar results. The presence of interstitial water molecules is sensitive to pressure, which disrupts tetrahedrality[93]. Figure S7 shows an almost negligible decrease/increase of the NIWM/IWM at 100 atm, relative to bulk water at 1 atm, indicating that density fluctuations are not significantly different at 100 atm, relative to bulk water at normal pressure. Furthermore, the density of water at 1 atm (0.997 $g \cdot cm^{-3}$) and 100 atm (1.001 $g \cdot cm^{-3}$) are only marginally different.

This demonstrates that the structural transformations of water, associated with the hydration of small hydrophobic solutes, already occur spontaneously in pure water, due to density fluctuations. Thus, any (compensating) entropic and enthalpic contributions associated with this water reorganization are already encoded in pure water, consistent with the fact that IT only requires information on the pair distribution function and density fluctuations of pure water.

The hydrophobic related tetrahedral enhancement (Fig. 5(a)), consistent with previous simulation[22,37,38,40–45] and experimental[46–48] studies for different hydrophobic molecules and groups, but at odds with neutron diffraction experiments[34–36], is completely lost at high temperatures for methane and neopentane, but not for methanol and neopentanol (see Fig. 5c).

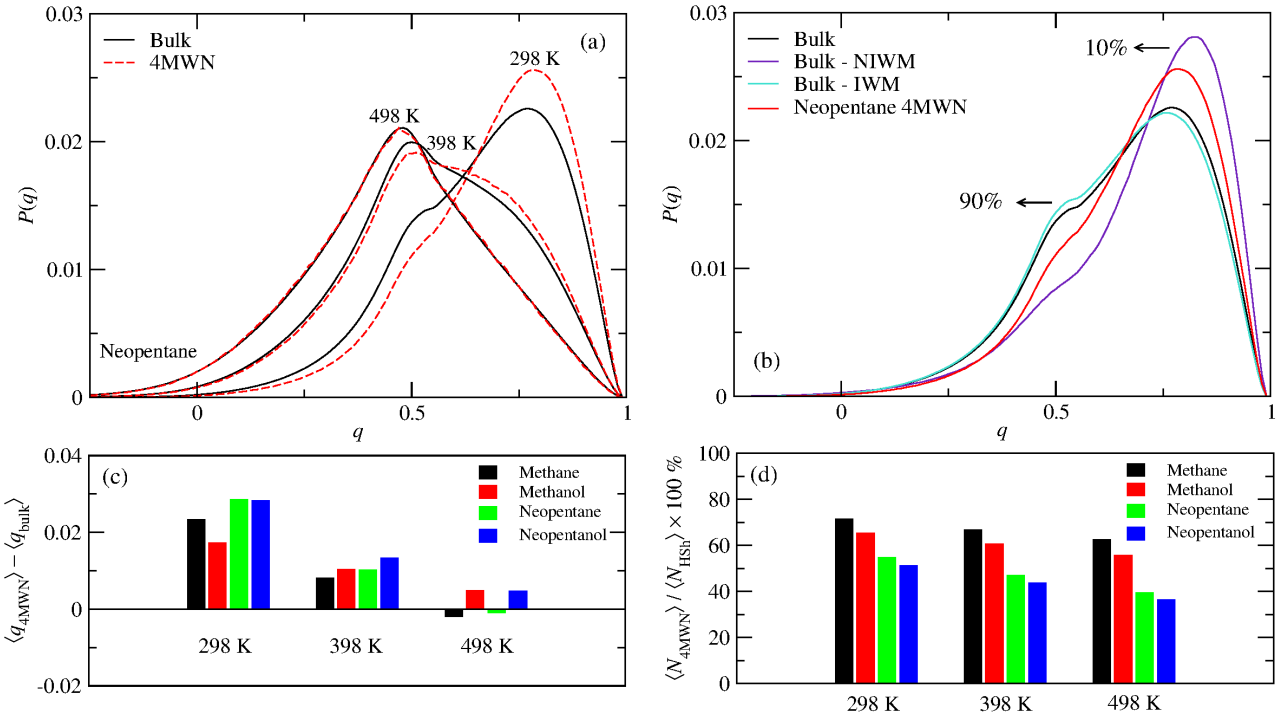


Figure 5 – Temperature dependence of the (a) tetrahedrality (q) distributions of water molecules with 4 or more water neighbors (4MWN) in the first coordination sphere of the methyl groups of neopentane and in bulk water, (b) comparison between q distributions in the neopentane coordination sphere and water molecules with no interstitial water molecules (NIWM) and with interstitial water molecules (IWM) in pure water at 298 K; the respective populations are shown (c) tetrahedrality difference between the water population with 4MWN and bulk water, and (d) fraction of the water population in the coordination sphere with 4MWN.

A similar structural loss is observed in bulk water for the NIWM population, at high temperatures, exhibiting a lower tetrahedrality than the IWM population (see Table S4). This is connected with the fact that at high temperatures water loses much of its tetrahedral geometry and an interstitial water molecule increases the number of ways in which a tetrahedron can form[51]. Figure 5(d) shows the temperature dependence of the fraction of water molecules with 4MWN. This is larger for the smaller solutes, consistent with the fact that the HB network of water is less perturbed upon dissolution of a small solute.

The observed tetrahedral enhancement leads to the formation of stronger HBs, especially with the third and fourth nearest water neighbors, as shown in Fig. 6; similar results were recently found by one of the authors for aromatic solutes[22]. Notice that the L4WN population exhibits the stronger HB, with the nearest neighbor. However, pair interactions with the third nearest neighbor are already weaker than in the in the bulk, even at room T, among this water population. This is expected, since the third water neighbor already appears at an average distance, longer than that found in bulk water[22].

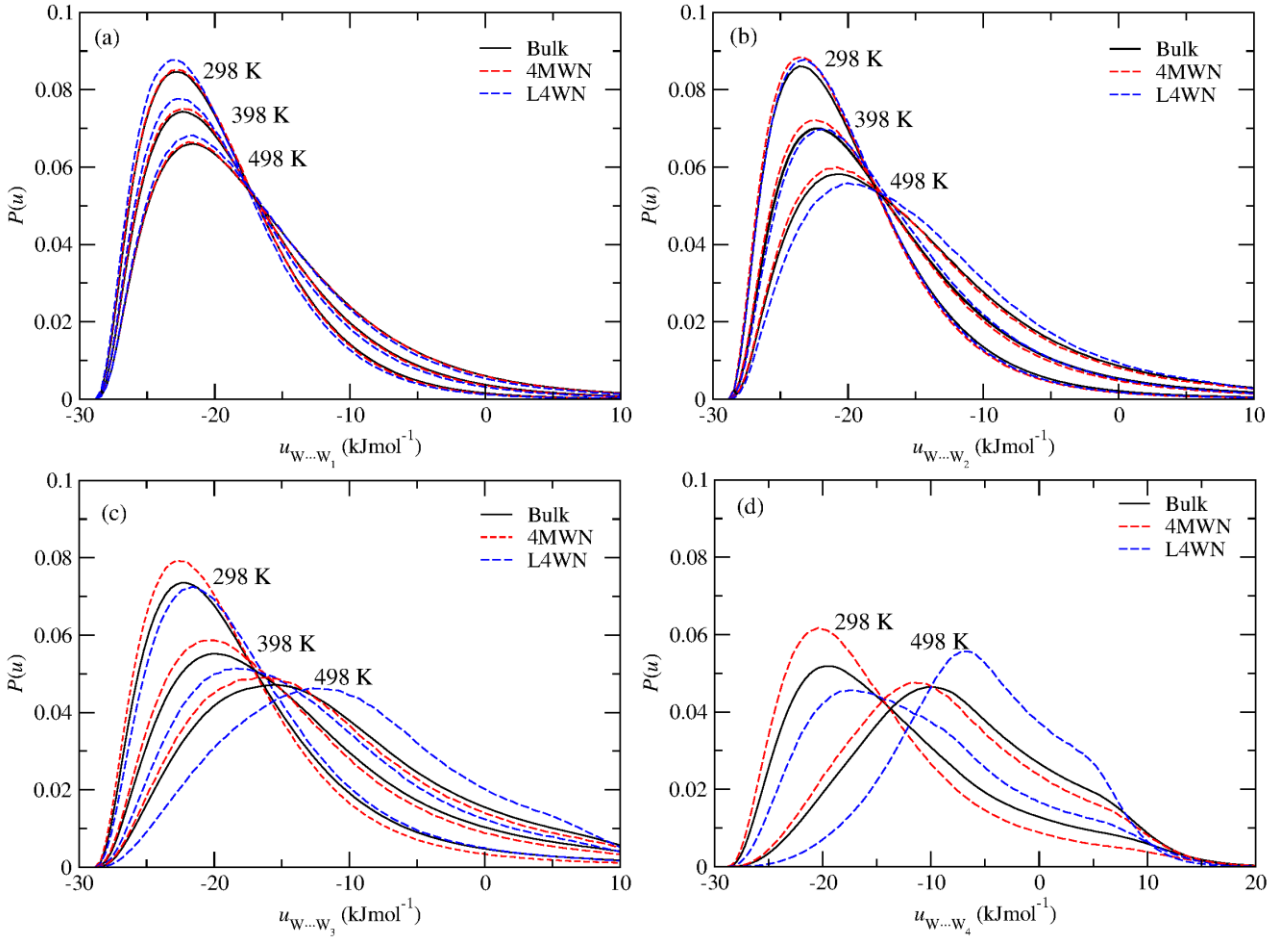


Figure 6 – Temperature dependence of the pair interaction energy distributions for the first, second, third, and fourth nearest water neighbors in the water populations with 4MWN and less than 4 water neighbors (L4WN) in the first hydration shell (HSh) of the methyl groups of neopentane and in bulk water. The distributions at 398 K for the fourth nearest water neighbor are omitted for clarity.

As can be seen, the temperature affects more dramatically water molecules with L4WN, that is, those nearest to the solute. Furthermore, this population increases with the temperature, along with a mild dehydration of the solutes (see Table S5). The fact that the 4MWN water population decreases at a slightly higher rate with the temperature, for the larger solutes, along with a slightly higher rate of dehydration (see Table S5), is consistent with the larger (positive) ΔS_{hyd} of neopentane, relative to methane, above the entropic convergence temperature, ~440 K (see Fig. 1(b_1)). For the amphiphilic solutes a similar behavior is observed. Entropic convergence occurs, however, at $T < T_S$ and a less negative ΔS_{hyd} is observed for neopentanol, relative to methanol, above ~ 460 K (see Fig. 3(b_1)). This negative entropy, as opposed to methane and neopentane, results from electrostatic interactions, as previously discussed, which are also responsible for the solute size dependence of ΔS_{hyd} / *SASA*, opposite to hydrocarbons.

The above results indicate that larger hydrophobic groups induce a more liquid-vapor like interface at every temperature, characterized by more water molecules with dangling OH groups

(L4WN population), and this difference is intensified with the temperature. This solute size dependence, disappears, however, for $\Delta S_{hyd}/SASA$, since the latter accounts for the number of water molecules in the hydration shell (see Table S6), thus, normalizing the solute size dependence of ΔS_{hyd} with respect to solute-water (van der Waals) and water-water interactions, in the first coordination sphere. Thus, while solute-water interactions greatly influence the temperature dependence of the rate of increase of ΔS_{hyd}, and, therefore, T_S, which is nearly size independent because of the different rates of increase of ΔS_{hyd}, water-water interactions, associated with structural perturbations next to the solutes, also play an important role, contributing to the reversal of the solute-size dependence of ΔS_{hyd}, above the entropic convergence temperature. This picture is fully supported by the fact that solute-water interactions do not converge at any temperature (see Fig. S6), opposite to ΔH_{hyd}, thus indicating that water-water interactions should be important to explain the reversal of the solute size dependence of both ΔH_{hyd} and ΔS_{hyd}.

IV. Conclusions

The hydration of hydrophobic and amphiphilic solutes was studied through molecular dynamics, aiming at gaining insight on the singular behavior of water, concerning hydrophobic zero entropy, T_S, and entropic convergence, also observed upon protein denaturation. Our results show that the difference between the T_S of hydrophobic (aliphatic and aromatic hydrocarbons) and amphiphilic solutes is associated with solute-water interactions, namely, electrostatic interactions. Further, the reversal of the solute-size dependence of ΔS_{hyd}, above the entropic convergence temperature, seems to be closely associated with a higher rate of dehydration of larger solutes, influencing both, solute-water and water-water interactions. Remarkably, ΔS_{hyd}, normalized by the SASA, shows a nearly solute size independence, for hydrocarbons, below and above T_S. For amphiphilic solutes, ΔS_{hyd}/SASA, does not exhibit a similar behavior because of long- range electrostatic interactions, much less sensible to the SASA. We also show that the tetrahedral enhancement of water next to hydrophobic groups is already encoded in neat water populations lacking interstitial waters, and that this is lost at high temperatures, both in pure water and next to hydrophobic solutes.

The source of the differences between Kauzmann's "hydrocarbon model" on protein denaturation and hydrophobic hydration, regarding the zero entropy and enthalpy temperatures, is also discussed, with relatively large amphiphilic (e.g. neopentanol) and polycyclic aromatic hydrocarbons, displaying a more similar behavior to globular proteins, than aliphatic hydrocarbons

or small amphiphilic molecules. This result is consistent with the fact that the core of globular proteins is not exclusively populated by aliphatic residues.

Acknowledgements

KT would like to acknowledge an undergraduate ERASMUS scholarship. NG would like to acknowledge financial support from Fundação para a Ciência e a Tecnologia of Portugal (CEEC/2018). NG would also like to thank Prof. Hank Ashbaugh for discussions on the zero entropy and information theory, and for reading the manuscript. Work supported by UIDB/04046/2020 and UIDP/04046/2020 centre grants from FCT, Portugal (to BioISI) and by the Portuguese National Distributed Computing Infrastructure (http://www.incd.pt).

Supporting Information

Various results, discussed in the manuscript, including, the ΔG_{hyd} for different water models at 298 K and 1 atm (Table S1), SASA for distinct molecules (Table S2), parameters for the generalized equation of $\Delta S_{hyd} / SASA$ (Table S3), average tetrahedrality of different water populations at 298 K and 498 K (Table S4), T dependence of the coordination numbers of different water populations (Tables S5 and S6), the T dependence of the thermodynamic functions for methane and methanol at 1 atm (Fig. S1), the T dependence of the solute-water interaction energy for aliphatic hydrocarbons (Fig. S2), T dependence of $\Delta S_{hyd} / SASA$ for hydrocarbons (Fig. S3), the T dependence of the solute-water interaction energy for aromatic hydrocarbons (Fig. S4) and amphiphilic solutes (Fig. S5), and a plot of the T dependence of the solute-water energy of hydrocarbons and amphiphilic solutes (Fig. S6), are available, free of charge in the Supporting Information.

References

1 W. Blokzijl and J. B. F. N. Engberts, Hydrophobic Effects. Opinions and Facts, *Angew. Chem. Int. Ed. Engl.*, 1993, **32**, 1545–1579.
2 D. Chandler, Interfaces and the driving force of hydrophobic assembly, *Nature*, 2005, **437**, 640–647.
3 H. S. Ashbaugh and L. R. Pratt, Colloquium: Scaled particle theory and the length scales of hydrophobicity, *Rev. Mod. Phys.*, 2006, **78**, 159–178.
4 N. T. Southall, K. A. Dill and A. D. J. Haymet, A View of the Hydrophobic Effect, *J. Phys. Chem. B*, 2002, **106**, 521–533.
5 P. L. Privalov and S. J. Gill, The hydrophobic effect: a reappraisal, *Pure Appl. Chem.*, 1989, **61**, 1097–1104.
6 G. I. Makhatadze and P. L. Privalov, Hydration effects in protein unfolding, *Biophys. Chem.*, 1994, **51**, 291–309.
7 R. L. Baldwin, Temperature dependence of the hydrophobic interaction in protein folding., *Proc. Natl. Acad. Sci.*, 1986, **83**, 8069–8072.

8 S. Garde, G. Hummer, A. E. García, M. E. Paulaitis and L. R. Pratt, Origin of Entropy Convergence in Hydrophobic Hydration and Protein Folding, *Phys. Rev. Lett.*, 1996, **77**, 4966–4968.
9 P. L. Privalov and S. J. Gill, in *Advances in Protein Chemistry*, Elsevier, 1988, vol. 39, pp. 191–234.
10 P. L. Privalov, in *Advances in Protein Chemistry*, Elsevier, 1979, vol. 33, pp. 167–241.
11 P. L. Privalov and N. N. Khechinashvili, A thermodynamic approach to the problem of stabilization of globular protein structure: A calorimetric study, *J. Mol. Biol.*, 1974, **86**, 665–684.
12 K. Murphy, P. Privalov and S. Gill, Common features of protein unfolding and dissolution of hydrophobic compounds, *Science*, 1990, **247**, 559–561.
13 W. Kauzmann, Some factors in the interpretation of protein denaturation, *Adv. Protein Chem.*, 1959, **14**, 1–63.
14 R. L. Baldwin and G. D. Rose, How the hydrophobic factor drives protein folding, *Proc. Natl. Acad. Sci.*, 2016, **113**, 12462–12466.
15 D. M. Huang and D. Chandler, Temperature and length scale dependence of hydrophobic effects and their possible implications for protein folding, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 8324–8327.
16 A. Ben-Naim, *Molecular Theory of Water and Aqueous Solutions: Part II: The Role of Water in Protein Folding, Self-Assembly and Molecular Recognition*, World Scientific Publishing Company, 2011.
17 R. L. Baldwin, Gas-liquid transfer data used to analyze hydrophobic hydration and find the nature of the Kauzmann-Tanford hydrophobic factor, *Proc. Natl. Acad. Sci.*, 2012, **109**, 7310–7313.
18 J. A. Schellman, Temperature, stability, and the hydrophobic interaction, *Biophys. J.*, 1997, **73**, 2960–2964.
19 A. D. Robertson and K. P. Murphy, Protein Structure and the Energetics of Protein Stability, *Chem. Rev.*, 1997, **97**, 1251–1268.
20 S. Garde and H. S. Ashbaugh, Temperature dependence of hydrophobic hydration and entropy convergence in an isotropic model of water, *J. Chem. Phys.*, 2001, **115**, 977–982.
21 F. Sedlmeier, D. Horinek and R. R. Netz, Entropy and enthalpy convergence of hydrophobic solvation beyond the hard-sphere limit, *J. Chem. Phys.*, 2011, **134**, 055105.
22 N. Galamba, A. Paiva, S. Barreiros and P. Simões, Solubility of Polar and Nonpolar Aromatic Molecules in Subcritical Water: The Role of the Dielectric Constant, *J. Chem. Theory Comput.*, 2019, **15**, 6277–6293.
23 R. L. Baldwin, The new view of hydrophobic free energy, *FEBS Lett.*, 2013, **587**, 1062–1066.
24 B. Lee, Solvent reorganization contribution to the transfer thermodynamics of small nonpolar molecules, *Biopolymers*, 1991, **31**, 993–1008.
25 T. Lazaridis, Solvent Size vs Cohesive Energy as the Origin of Hydrophobicity, *Acc. Chem. Res.*, 2001, **34**, 931–937.
26 R. L. Baldwin, Dynamic hydration shell restores Kauzmann's 1959 explanation of how the hydrophobic factor drives protein folding, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 13052–13056.
27 B. Lee, The physical origin of the low solubility of nonpolar solutes in water, *Biopolymers*, 1985, **24**, 813–823.
28 L. R. Pratt and A. Pohorille, Theory of hydrophobicity: Transient cavities in molecular liquids, *Proc. Natl. Acad. Sci.*, 1992, **89**, 2995–2999.
29 A. Ben-Naim, Hydrophobic interaction and structural changes in the solvent, *Biopolymers*, 1975, **14**, 1337–1355.
30 E. Grunwald and C. Steel, Solvent Reorganization and Thermodynamic Enthalpy-Entropy Compensation, *J. Am. Chem. Soc.*, 1995, **117**, 5687–5692.
31 A. Ben-Naim, A simple model for demonstrating the relation between solubility, hydrophobic interaction, and structural changes in the solvent, *J. Phys. Chem.*, 1978, **82**, 874–885.

32 G. Graziano, Entropy Convergence in the Hydration Thermodynamics of *n* -Alcohols, *J. Phys. Chem. B*, 2005, **109**, 12160–12166.
33 B. Lee, A procedure for calculating thermodynamic functions of cavity formation from the pure solvent simulation data, *J. Chem. Phys.*, 1985, **83**, 2421–2425.
34 J. Turner and A. K. Soper, The effect of apolar solutes on water structure: Alcohols and tetraalkylammonium ions, *J. Chem. Phys.*, 1994, **101**, 6116.
35 A. K. Soper and J. L. Finney, Hydration of methanol in aqueous solution, *Phys. Rev. Lett.*, 1993, **71**, 4346–4349.
36 P. Buchanan, N. Aldiwan, A. K. Soper, J. L. Creek and C. A. Koh, Decreased structure on dissolving methane in water, *Chem. Phys. Lett.*, 2005, **415**, 89–93.
37 M. P. S. Mateus, N. Galamba and B. J. C. Cabral, Structure and electronic properties of a benzene-water solution, *J. Chem. Phys.*, 2012, **136**, 014507.
38 N. Galamba, Water's Structure around Hydrophobic Solutes and the Iceberg Model, *J. Phys. Chem. B*, 2013, **117**, 2153–2159.
39 N. Galamba, Reply to "Comment on 'Water's Structure around Hydrophobic Solutes and the Iceberg Model,'" *J. Phys. Chem. B*, 2014, **118**, 2600–2603.
40 N. Galamba, Water Tetrahedrons, Hydrogen-Bond Dynamics, and the Orientational Mobility of Water around Hydrophobic Solutes, *J. Phys. Chem. B*, 2014, **118**, 4169–4176.
41 H. F. M. C. Martiniano and N. Galamba, Fast and slow dynamics and the local structure of liquid and supercooled water next to a hydrophobic amino acid, *Phys Chem Chem Phys*, 2016, **18**, 27639–27647.
42 B. Song and V. Molinero, Thermodynamic and structural signatures of water-driven methane-methane attraction in coarse-grained mW water, *J. Chem. Phys.*, 2013, **139**, 054511.
43 H. S. Ashbaugh, J. W. Barnett, A. Saltzman, M. E. Langrehr and H. Houser, Communication: Stiffening of dilute alcohol and alkane mixtures with water, *J. Chem. Phys.*, 2016, **145**, 201102.
44 S. Daschakraborty, How do glycerol and dimethyl sulphoxide affect local tetrahedral structure of water around a nonpolar solute at low temperature? Importance of preferential interaction, *J. Chem. Phys.*, 2018, **148**, 134501.
45 F. Merzel and F. Avbelj, Why do water molecules around small hydrophobic solutes form stronger hydrogen bonds than in the bulk?, *Biochim. Biophys. Acta BBA - Gen. Subj.*, 2020, **1864**, 129537.
46 J. G. Davis, K. P. Gierszal, P. Wang and D. Ben-Amotz, Water structural transformation at molecular hydrophobic interfaces, *Nature*, 2012, **491**, 582–585.
47 X. Wu, W. Lu, L. M. Streacker, H. S. Ashbaugh and D. Ben-Amotz, Methane Hydration-Shell Structure and Fragility, *Angew. Chem. Int. Ed.*, 2018, **57**, 15133–15137.
48 J. Grdadolnik, F. Merzel and F. Avbelj, Origin of hydrophobicity and enhanced water hydrogen bond strength near purely hydrophobic solutes, *Proc. Natl. Acad. Sci.*, 2017, **114**, 322–327.
49 G.-H. Deng, Y. Shen, H. Chen, Y. Chen, B. Jiang, G. Wu, X. Yang, K. Yuan and J. Zheng, Ordered-to-Disordered Transformation of Enhanced Water Structure on Hydrophobic Surfaces in Concentrated Alcohol–Water Solutions, *J. Phys. Chem. Lett.*, 2019, **10**, 7922–7928.
50 F. Sciortino, A. Geiger and H. E. Stanley, Effect of defects on molecular mobility in liquid water, *Nature*, 1991, **354**, 218–221.
51 N. Galamba, On the hydrogen-bond network and the non-Arrhenius transport properties of water, *J. Phys. Condens. Matter*, 2017, **29**, 015101.
52 F. H. Stillinger, Structure in aqueous solutions of nonpolar solutes from the standpoint of scaled-particle theory, *J. Solut. Chem.*, 1973, **2**, 141–158.
53 L. R. Pratt and D. Chandler, Theory of the hydrophobic effect, *J. Chem. Phys.*, 1977, **67**, 3683.
54 G. Hummer, S. Garde, A. E. Garcia, A. Pohorille and L. R. Pratt, An information theory model of hydrophobic interactions, *Proc. Natl. Acad. Sci.*, 1996, **93**, 8951.
55 B. J. Berne, Inferring the hydrophobic interaction from the properties of neat water., *Proc. Natl. Acad. Sci.*, 1996, **93**, 8800–8803.
56 G. Graziano, Contrasting the hydration thermodynamics of methane and methanol, *Phys. Chem.*

Chem. Phys., 2019, **21**, 21418–21430.
57 W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids, *J. Am. Chem. Soc.*, 1996, **118**, 11225–11236.
58 D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen, GROMACS: fast, flexible, and free, *J. Comput. Chem.*, 2005, **26**, 1701–1718.
59 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys.*, 1983, **79**, 926–935.
60 H. J. C. Berendsen, J. R. Grigera and T. P. Straatsma, The missing term in effective pair potentials, *J. Phys. Chem.*, 1987, **91**, 6269–6271.
61 H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura and T. Head-Gordon, Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew, *J. Chem. Phys.*, 2004, **120**, 9665–9678.
62 J. L. F. Abascal and C. Vega, A general purpose model for the condensed phases of water: TIP4P/2005, *J. Chem. Phys.*, 2005, **123**, 234505.
63 G. Bussi, D. Donadio and M. Parrinello, Canonical sampling through velocity rescaling, *J. Chem. Phys.*, 2007, **126**, 014101.
64 M. Parrinello, Polymorphic transitions in single crystals: A new molecular dynamics method, *J. Appl. Phys.*, 1981, **52**, 7182.
65 U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, A smooth particle mesh Ewald method, *J. Chem. Phys.*, 1995, **103**, 8577.
66 B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, LINCS: A linear constraint solver for molecular simulations, *J. Comput. Chem.*, 1997, **18**, 1463–1472.
67 G. Duarte Ramos Matos, D. Y. Kyu, H. H. Loeffler, J. D. Chodera, M. R. Shirts and D. L. Mobley, Approaches for Calculating Solvation Free Energies and Enthalpies Demonstrated with an Update of the FreeSolv Database, *J. Chem. Eng. Data*, 2017, **62**, 1559–1569.
68 C. H. Bennett, Efficient estimation of free energy differences from Monte Carlo data, *J. Comput. Phys.*, 1976, **22**, 245–268.
69 A. Ben-Naim and Y. Marcus, Solvation thermodynamics of nonionic solutes, *J. Chem. Phys.*, 1984, **81**, 2016–2027.
70 W. F. Van Gunsteren and H. J. C. Berendsen, A Leap-frog Algorithm for Stochastic Dynamics, *Mol. Simul.*, 1988, **1**, 173–185.
71 T. Steinbrecher, I. Joung and D. A. Case, Soft-core potentials in thermodynamic integration: Comparing one- and two-step transformations, *J. Comput. Chem.*, 2011, **32**, 3253–3263.
72 T. C. Beutler, A. E. Mark, R. C. van Schaik, P. R. Gerber and W. F. van Gunsteren, Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations, *Chem. Phys. Lett.*, 1994, **222**, 529–539.
73 V. Gapsys, D. Seeliger and B. L. de Groot, New Soft-Core Potential Function for Molecular Dynamics Based Alchemical Free Energy Calculations, *J. Chem. Theory Comput.*, 2012, **8**, 2373–2382.
74 T. T. Pham and M. R. Shirts, Identifying low variance pathways for free energy calculations of molecular transformations in solution phase, *J. Chem. Phys.*, 2011, **135**, 034114.
75 E. Gallicchio, M. M. Kubo and R. M. Levy, Enthalpy−Entropy and Cavity Decomposition of Alkane Hydration Free Energies: Numerical Results and Implications for Theories of Hydrophobic Solvation, *J. Phys. Chem. B*, 2000, **104**, 6271–6285.
76 A. Shrake and J. A. Rupley, Environment and exposure to solvent of protein atoms. Lysozyme and insulin, *J. Mol. Biol.*, 1973, **79**, 351–371.
77 F. Eisenhaber, P. Lijnzaad, P. Argos, C. Sander and M. Scharf, The double cubic lattice method: Efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies, *J. Comput. Chem.*, 1995, **16**, 273–284.
78 A. Bondi, van der Waals Volumes and Radii, *J. Phys. Chem.*, 1964, **68**, 441–451.
79 P.-L. Chau and A. J. Hardwick, A new order parameter for tetrahedral configurations, *Mol. Phys.*,

1998, **93**, 511–518.
80 J. R. Errington and P. G. Debenedetti, Relationship between structural order and the anomalies of liquid water, *Nature*, 2001, **409**, 318–321.
81 N. Galamba and B. J. C. Cabral, The Changing Hydrogen-Bond Network of Water from the Bulk to the Surface of a Cluster: A Born−Oppenheimer Molecular Dynamics Study, *J. Am. Chem. Soc.*, 2008, **130**, 17955–17960.
82 Q. Du, R. Superfine, E. Freysz and Y. R. Shen, Vibrational spectroscopy of water at the vapor/water interface, *Phys. Rev. Lett.*, 1993, **70**, 2313–2316.
83 E. A. Raymond, T. L. Tarbuck, M. G. Brown and G. L. Richmond, Hydrogen-Bonding Interactions at the Vapor/Water Interface Investigated by Vibrational Sum-Frequency Spectroscopy of HOD/H_2O/D_2O Mixtures and Molecular Dynamics Simulations, *J. Phys. Chem. B*, 2003, **107**, 546–556.
84 I. Benjamin, Vibrational Spectrum of Water at the Liquid/Vapor Interface, *Phys. Rev. Lett.*, 1994, **73**, 2083–2086.
85 E. Shiratani and M. Sasai, Growth and collapse of structural patterns in the hydrogen bond network in liquid water, *J. Chem. Phys.*, 1996, **104**, 7671–7680.
86 A. Ben-Naim and M. Yaacobi, Effects of Solutes on the Strength of Hydrophobic Interaction and Its Temperature Dependence, *J. Phys. Chem.*, 1974, **78**, 170–175.
87 H. Naghibi, S. F. Dec and S. J. Gill, Heat of solution of methane in water from 0 to 50.degree.C, *J. Phys. Chem.*, 1986, **90**, 4621–4623.
88 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, Development and testing of a general amber force field, *J. Comput. Chem.*, 2004, **25**, 1157–1174.
89 B. Widom, Potential-distribution theory and the statistical mechanics of fluids, *J. Phys. Chem.*, 1982, **86**, 869–872.
90 S. Cabani, P. Gianni, V. Mollica and L. Lepori, Group contributions to the thermodynamic properties of non-ionic organic solutes in dilute aqueous solution, *J. Solut. Chem.*, 1981, **10**, 563–595.
91 G. I. Makhatadze and P. L. Privalov, Heat capacity of alcohols in aqueous solutions in the temperature range from 5 to 125^{o}C, *J. Solut. Chem.*, 1989, **18**, 927–936.
92 G. Graziano and B. Lee, On the intactness of hydrogen bonds around nonpolar solutes dissolved in water, *J. Phys. Chem. B*, 2005, **109**, 8103–8107.
93 N. Galamba, On the Effects of Temperature, Pressure, and Dissolved Salts on the Hydrogen-Bond Network of Water, *J. Phys. Chem. B*, 2013, **117**, 589–601.

Protein Denaturation, Zero Entropy Temperature, and the Structure of Water around Hydrophobic and Amphiphilic Solutes

Supporting Information

Kazimieras Tamoliūnas[a], Nuno Galamba[a,*]

[a] *Centre of Chemistry and Biochemistry and Biosystems and Integrative Sciences Institute, Faculty of Sciences of the University of Lisbon, C8, Campo Grande, 1749-016 Lisbon, Portugal.*

*Corresponding author. Electronic mail: njgalamba@fc.ul.pt

Table S1 – Hydration free energy at 298 K and 1 atm for various nonpolar and polar OPLS-aa model solutes in different water models.

	TIP3P ΔG_{hyd} (kJmol^{-1})	SPC/E ΔG_{hyd} (kJmol^{-1})	TIP4P-Ew ΔG_{hyd} (kJmol^{-1})	TIP4P-2005 ΔG_{hyd} (kJmol^{-1})	Experimental ΔG_{hyd} (kJmol^{-1})
Methane	+9.1 ± 0.02	+9.5 ± 0.1	+9.5 ± 0.02	+9.4 ± 0.04	+8.4
Methanol	-17.8 ± 0.1	-18.0 ± 0.1	-18.4 ± 0.1	-18.9 ± 0.15	-21.3
Benzene	-1.5 ± 0.1	- 0.2 ± 0.1	+ 0.1 ± 0.1	- 0.6 ± 0.1	-3.6
Phenol	-21.7 ± 0.02	-20.1 ± 0.2	-19.7 ± 0.1	-20.2 ± 0.1	-27.7
Naphthalene	- 7.8 ± 0.1	- 5.5 ± 0.1	- 5.0 ± 0.1	- 5.9 ± 0.2	-9.6
Anthracene	-13.2 ± 0.2	-10.1 ± 0.1	-9.2 ± 0.1	-10.7 ± 0.1	-17.7

Table S2 – Solvent accessible surface area (SASA) values for the different solutes.

Solute	SASA[1] (nm^2)	Solute	SASA (nm^2)	Solute	SASA (nm^2)
Methane	1.43(6)	Methanol	1.63(0)	Benzene	2.44(1)
Ethane	1.81(7)	Ethanol	1.97(9)	Naphthalene	3.10(3)
Neopentane	2.60(7)	Neopentanol	2.58(3)	Anthracene	3.76(4)
Pentane	2.75(0)	Phenol	2.70(5)	Pyrene	3.87(3)
Heptane	3.36(3)				

[1] Solvent sphere radius 1.4 Å.

Table S3 – Generalized parameters for the temperature dependence of $T\Delta S_{hyd}(\text{kJmol}^{-1})$, normalized by the SASA (nm^2), for aliphatic and aromatic hydrocarbons. The equation is of the form $T\Delta S / SASA = A + BT + CT^2$, where T is the temperature in K.

	A	B	C
Aliphatic	0.0185382	-0.123008	0.000278762
Aromatic	0.00437478	-0.100095	0.000211814

Table S4 – Mean tetrahedrality, <q>, of different water populations at 298 K and 498 K. The no interstitial water molecules (NIWN) population includes water molecules surrounded by no more than 4 water neighbors up to 3.7 Å. The interstitial water molecules (IWN) population includes water molecules surrounded by 5 water neighbors up to 3.7 Å. The 4 or more water neighbors (4MWN) population includes water molecules in the first coordination sphere of neopentane that retain four or more water neighbors, closer than any atom of the solute.

	<q>	
Water Population	298 K	498 K
Bulk	0.668	0.477
Bulk - NIWM	0.717	0.443
Bulk - IWM	0.662	0.489
Neopentane – 4MWN	0.697	0.476

Table S5 – Temperature dependence of the number of water molecules in the first hydration shell, N_{HSh}, and the respective number of water molecules with four or more water neighbors (4MWN), N_{4MWN}, for different solutes.

	298 K N_{HSh} [N_{4MWN}]	398 K N_{HSh} [N_{4MWN}]	498 K N_{HSh} [N_{4MWN}]	Δ(498K-298K) ΔN_{HSh} [ΔN_{4MWN}]
Methane	20.3[14.5]	18.7[12.5]	15.8[9.9]	4.5[4.6]
Methanol	20.3[13.3]	18.7[11.3]	16.2[9.0]	4.1[4.3]
Neopentane	25.9[14.2]	22.2[10.4]	17.7[7.0]	8.2[7.2]
Neopentanol	25.4[13.0]	22.3[9.7]	18.2[6.6]	7.2[6.4]

Table S6 – Temperature dependence of the number of water molecules in the first hydration shell, N_{HSh}, and the respective number of water molecules with four or more water neighbors (4MWN), N_{4MWN}, normalized by the SASA, for different solutes.

	298 K N_{HSh}/SASA [N_{4MWN}/SASA]	398 K N_{HSh}/SASA [N_{4MWN}/SASA]	498 K N_{HSh}/SASA [N_{4MWN}/SASA]	Δ(498K-298K) N_{HSh}/SASA [N_{4MWN}/SASA]
Methane	14.1[10.1]	13.0[8.7]	11.0[6.9]	3.1[3.2]
Methanol	12.5[8.2]	11.5[6.9]	9.9[5.5]	2.5[2.6]
Neopentane	9.9[5.4]	8.5[4.0]	6.8[2.7]	3.1[2.8]
Neopentanol	9.4[4.8]	8.2[3.6]	6.7[2.4]	2.7[2.4]

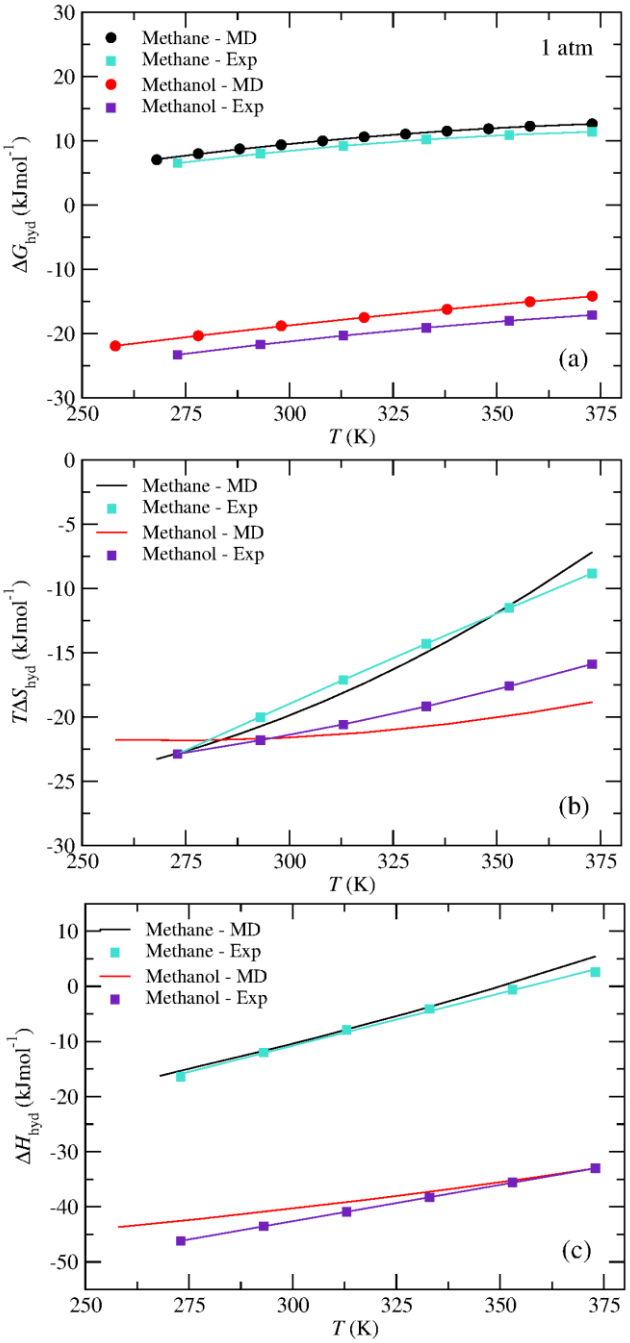


Figure S1 – Temperature dependence of the hydration (a) free energy, (b) entropy, and (c) enthalpy of methane and methanol at 1 atm. The values from molecular dynamics (MD) simulations are compared with available "experimental" data (Table 1 of ref. 51). The methane experimental thermodynamic parameters were obtained from the experimental values at 298 K and 1 atm and the experimental ΔC_p up to 323 K and linear extrapolation of ΔC_p up to 373 K. The heat capacities for methanol used to estimate the thermodynamic parameters were obtained using group additivity contributions.

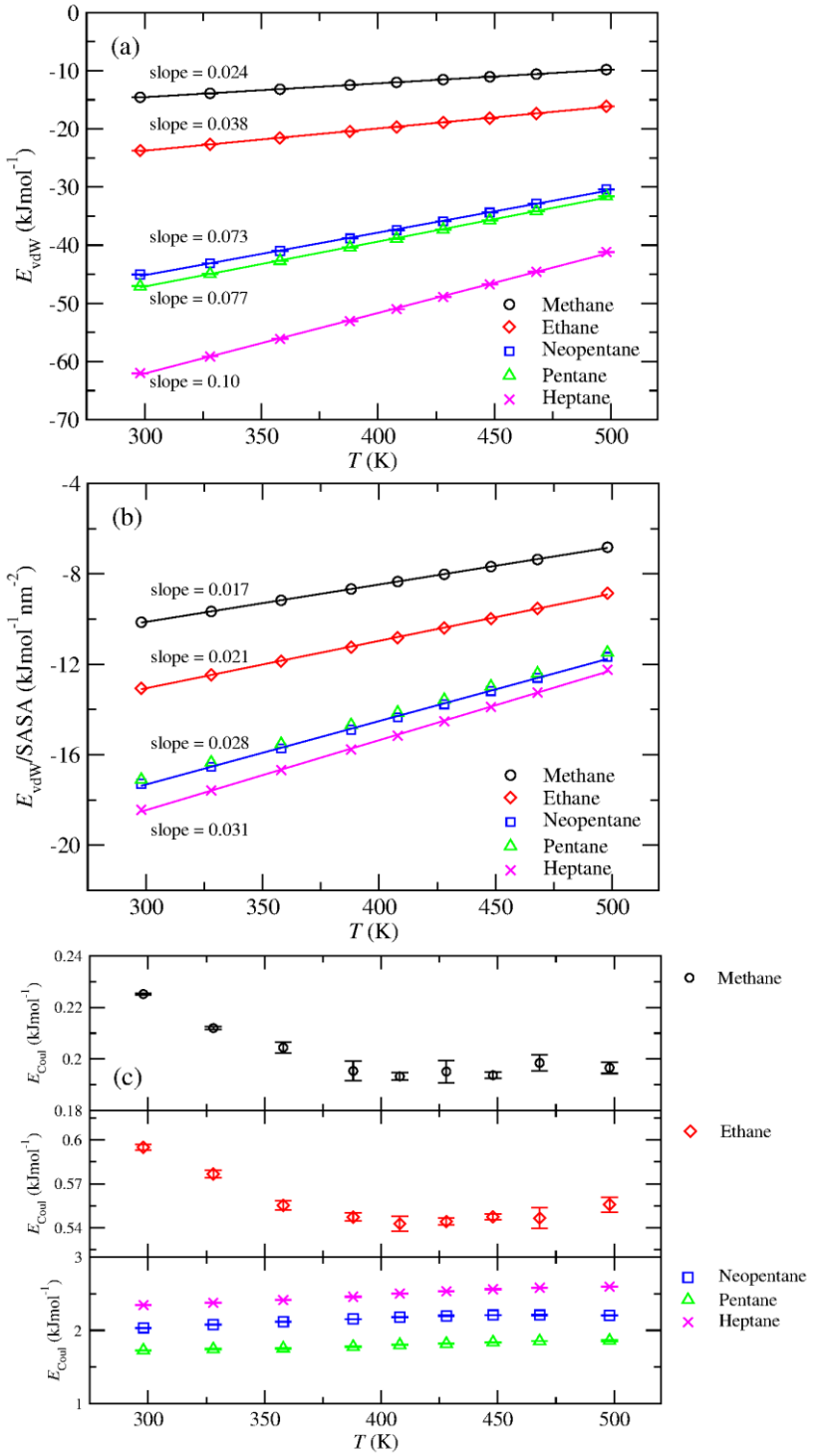


Figure S2 – Temperature dependence of the solute-water (a) van der Waals (vdW) interaction energy, (b) vdW interaction energy normalized by the SASA, and (c) short range (r < 10 Å) component of the Coulomb interaction energy for the distinct aliphatic hydrocarbons. The rate of increase of the vdW energy with the temperature increases with the solute size. A more similar rate of increase can be observed for the van der Waals energy normalized by the SASA. The long-range contribution of the electrostatic energy is not included because the PME reciprocal space component of the electrostatic potential energy cannot be separated into solute-water and water-water components. The slopes from linear fittings (solid lines) are shown. Errors bars are standard deviations calculated from four, 20 ns long, independent simulations.

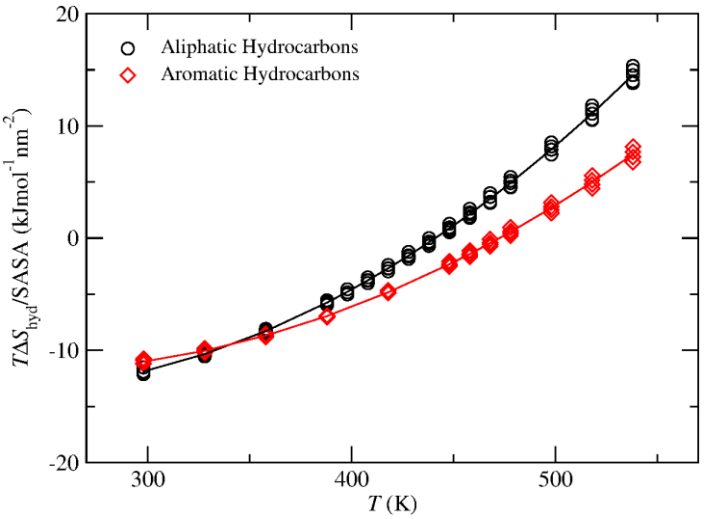


Figure S3 – Generalized (every solute) equations (solid lines) for the hydration entropy of aliphatic and aromatic hydrocarbons, normalized by the SASA. The data was fitted to a second order polynomial, $T\Delta S / SASA = A + BT + CT^2$; parameters are given in Table S3.

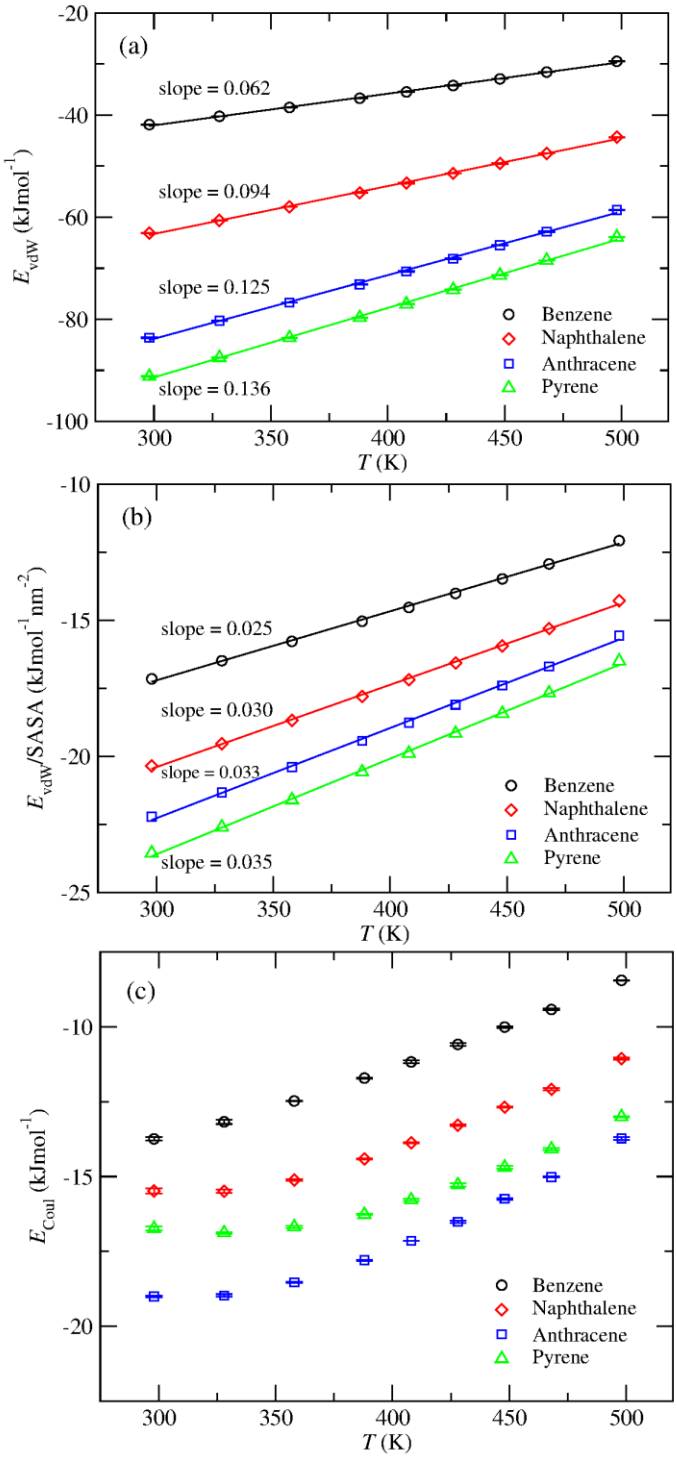


Figure S4 – Temperature dependence of the solute-water (a) van der Waals (vdW) interaction energy, (b) vdW interaction energy normalized by the SASA, and (c) short range (r < 10 Å) component of the Coulomb interaction energy for the distinct aromatic hydrocarbons. The rate of increase of the vdW energy with the temperature increases with the solute size. A more similar rate of increase can be observed for the van der Waals energy normalized by the SASA. The slopes from linear fittings (solid lines) are shown. Errors bars are standard deviations calculated from four, 20 ns long, independent simulations.

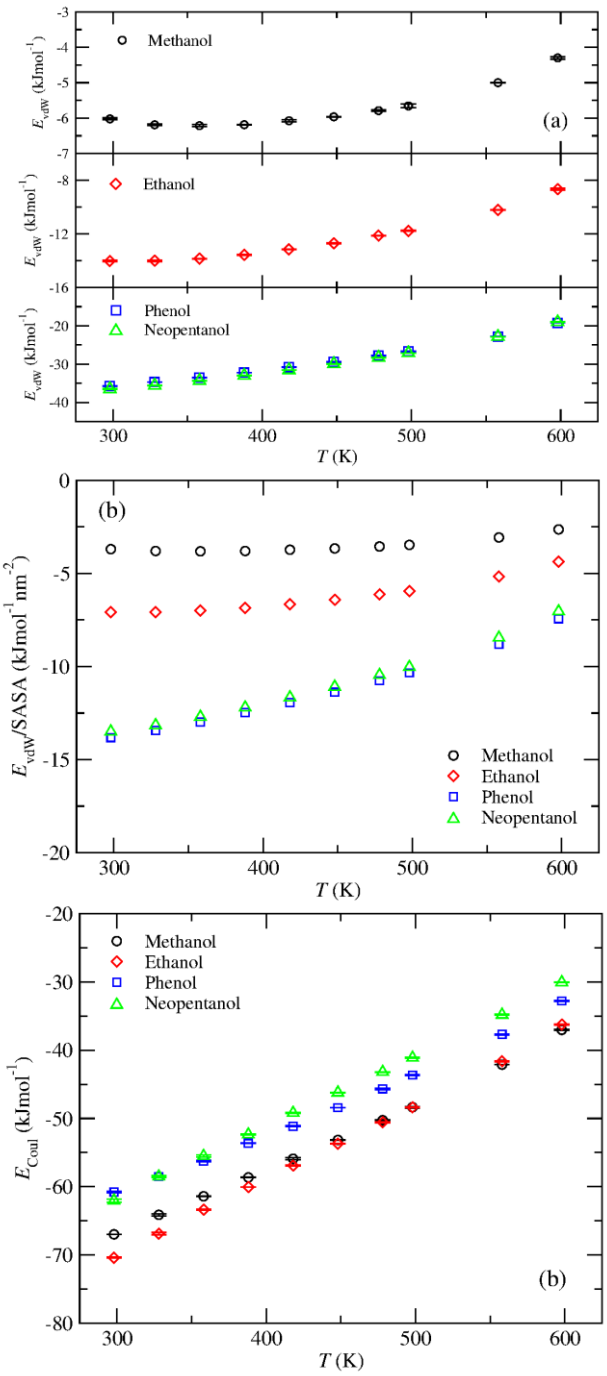


Figure S5 – Temperature dependence of the solute-water (a) van der Waals (vdW) interaction energy, (b) vdW interaction energy normalized by the SASA, and (c) short range (r < 10 Å) component of the Coulomb interaction energy for the distinct alcohols. Errors bars are standard deviations calculated from four, 20 ns long, independent simulations.

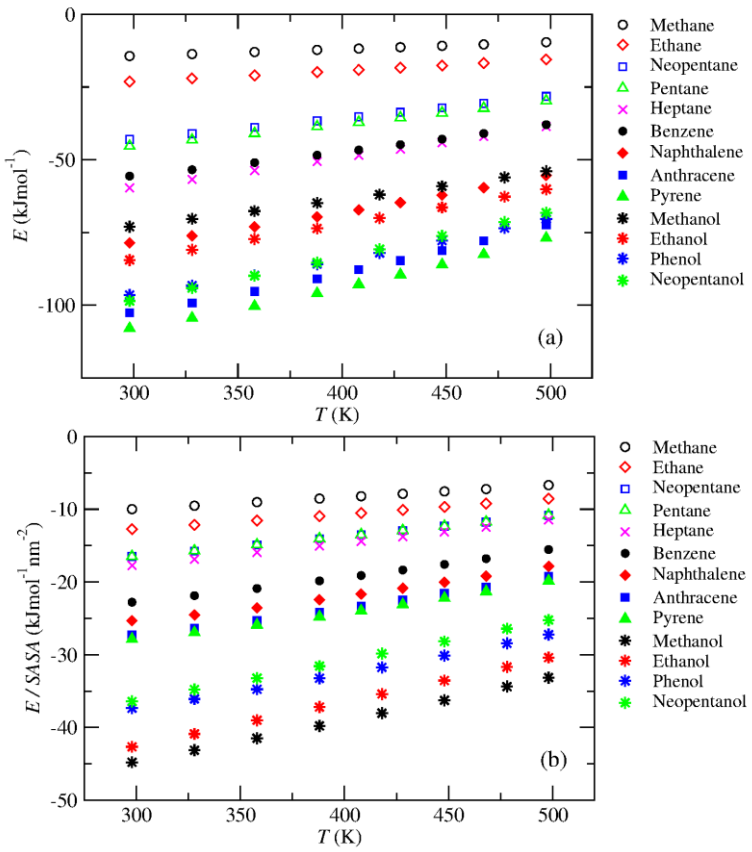


Figure S6 – Temperature dependence of the solute-water, (a) van der Waals and short-range component of the Coulomb interaction energy, *E*, and (b) *E* normalized by the SASA.

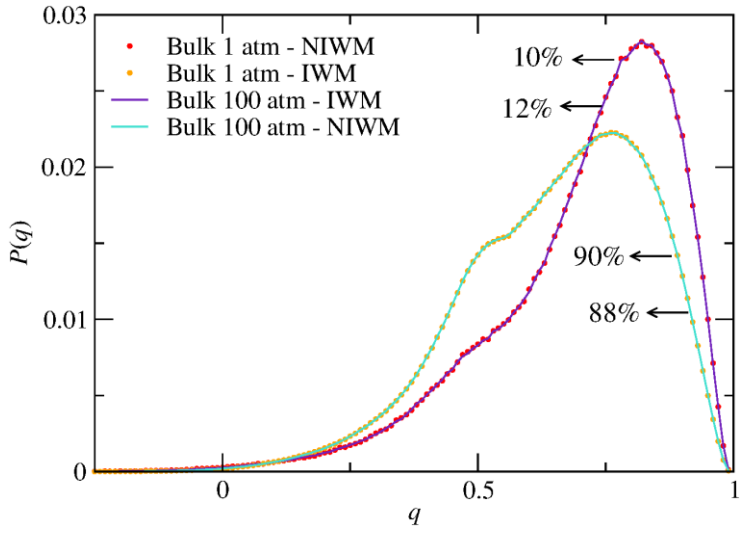


Figure S7 – Tetrahedrality of bulk water populations with no interstitial water neighbors (NIWN) and with interstitial water neighbors (IWN) at 298 K and 1 atm and 100 atm. The distributions are nearly indistinguishable, whereas the NIWM populations are 10% and 12%, at 1 atm and 100 atm, respectively. The density of TIP4P/2005 water at 298 K and 1 atm and 100 atm, is, respectively, 0.997 g·cm^{-3} and 1.001 g·cm^{-3}.