# Water Structure-Forming Capabilities Are Temperature Shifted for Different Models

Roman Shevchuk, Diego Prada-Gracia, and Francesco Rao*

Freiburg Institute for Advanced Studies, University of Freiburg, Freiburg, Germany

ABSTRACT: A large number of water models exist for molecular simulations. They differ in the ability to reproduce specific features of real water instead of others, like the correct temperature for the density maximum or the diffusion coefficient. Past analysis mostly concentrated on ensemble quantities, while few data were reported on the different microscopic behavior. Here, we compare seven widely used classical water models (SPC, SPC/E, TIP3P, TIP4P, TIP4P-Ew, TIP4P/2005, and TIP5P) in terms of their local structure-forming capabilities through hydrogen bonds for temperatures ranging from 210 to 350 K by the introduction of a set of order parameters taking into account the configuration of up to the second solvation shell. We found that all models share the same structural pattern up to a temperature shift. When this shift is applied, all models overlap onto a master curve. Interestingly, increased stabilization of fully coordinated structures extending to at least two solvation shells is found for models that are able to reproduce the correct position of the density maximum. Our results provide a self-consistent atomic-level structural comparison protocol, which can be of help in elucidating the influence of different water models on protein structure and dynamics.
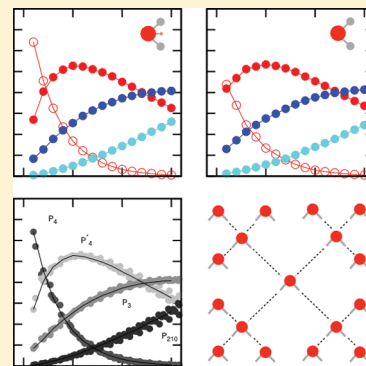
I. INTRODUCTION

At the fundamental level, water directly influences several biologically relevant processes including protein folding,[1] protein–protein association,[1–4] and amyloid aggregation.[5] Surprisingly, relatively simple models with fixed charges and geometry are able to reproduce the phase diagram as well as many of the anomalies of water with good accuracy.[6,7] For example, all popular classical water models present a density maximum.[8,9] However, only those that explicitly included this information in the fitting of the potential are able to correctly reproduce the experimental value located at around 277 K at ambient pressure.[10]

Due to their improved speed, biomolecular simulations in explicit water were traditionally run with TIP3P[11] or SPC.[12] Nowadays, more elaborated models can be easily used and their impact on the calculation assessed.[13] Optimized four-site models reproducing the experimental temperature of maximum density seem to improve the quality of biomolecular simulations. Best and collaborators showed that predicted helical propensities are in better agreement with experiments when a TIP4P/2005 water model is chosen in place of the traditional TIP3P.[14] Others reported that TIP4P-Ew provides better free-energy estimations compared to conventional water models.[15] In both studies, the improved behavior was not connected to a clear microscopic property of the water model. To this aim, one limitation is the lack of a common framework to compare the structural behavior of liquid water at the atomic level.

Here, seven popular classical water models, namely, SPC,[12] SPC/E,[16] TIP3P,[11] TIP4P,[17] TIP4P-Ew,[18] TIP4P/2005,[9] and TIP5P,[19] are investigated in terms of their local structure-forming capabilities, that is, their ability to form structured or partially structured environments of the size of up to two solvation shells through hydrogen bonds. This approach represents a simplified version of the recent complex network analysis introduced to study water structural inhomogeneities.[20,21]

II. METHODS

A. Simulations. Molecular dynamics simulations were run with the program GROMACS[22] with an integration time step of 2 fs. The water box consisted of 1024 water molecules in the NPT ensemble with pressure of 1 atm and temperatures ranging from 210 to 350 K with steps of 10 K. The Berendsen barostat,[23] velocity rescale thermostat,[24] and PME[25] were used for pressure coupling, temperature coupling, and long-range electrostatics, respectively.

The present analysis was done over 25 000 snapshots obtained from a 100 ps long run after a 10 ns equilibration in the same conditions. The short length of the analyzed trajectory is justified because structural properties are calculated for each water molecule, effectively improving the sampling by 3 orders of magnitudes (there are 1024 water molecules). Repeating the analysis with a longer equilibration time reproduced the same results. TIP5P data were collected up to 230 K, just before the approaching of the glass transition.[26]

B. Density Maximum. The position of the maximum density was obtained from 1 ns long simulations after 10 ns of equilibration in the NPT ensemble. The temperature of

maximum density was extracted by polynomial fitting around the maximum. Variations from the literature (see Table 1) may

Table 1. Temperature of maximum density calculated from our simulations (TMD), as found in the literature (TMD^{ref}), the structural temperature shift (ΔT_s), and the temperature at which P_4^* is maximum for the seven water models investigated in this work

water model	TMD	TMD^{ref}	ΔT_s	$T_{\max(P4*)}$
TIP3P	199	182[8]	65	229
SPC	226	228[8]	42	247
SPC/E	250	241[33]	18	275
TIP4P	256	248[11]	20	268
TIP4P/2005	280	278[9]	0	287
TIP4P-Ew	273	274[18]	6	281
TIP5P	282	285[34]	n.a.	269

be due to size effects and a different treatment of the electrostatics. The location of the TIP3P density maximum was obtained by running further simulations at lower temperatures.

C. Energetics and Tetrahedral Order Parameter. The free energy of a configuration i is given by

$$\Delta F_i = -k_B T \log(P_i) \tag{1}$$

where k_B is the Boltzmann factor, T the temperature, and P_i the population of the selected configuration. The enthalpy is estimated by summing up all pairwise contributions between the water molecules belonging to the same configuration.

The tetrahedral order parameter for a water molecule i is calculated as

$$q_i = 1 - \frac{3}{8} \sum_{j=1}^{3} \sum_{k=j+1}^{4} \left(\cos \psi_{jik} + \frac{1}{3} \right)^2 \tag{2}$$

where j and k are any of the four nearest water molecules of i, and ψ_{jik} is the angle formed by their oxygens.[27] The averaged value of this order parameter over an ensemble of water molecules is denoted as Q.

III. RESULTS

A. Structure-Forming Capabilities. Water structure-forming capabilities were investigated by analyzing the hydrogen-bond network of each water molecule in the simulation box together with its first and second solvation shells. A maximum of four hydrogen bonds per molecule was considered. A bond is formed when the distance between oxygens and the angle O–H–O is smaller than 3.5 Å and 30°, respectively.[28] Water structures were grouped into four archetypal configurations of population $P_i^{(*)}$: the fully coordinated first and second solvation shells for a total of 16 hydrogen bonds (P_4, see Figure 1 for a schematic representation); the fully coordinated first shell, in which one or more hydrogen bonds between the first and the second shells are missing or loops are formed (P_4^*); the three coordinated water molecule (P_3); and the rest (P_{210}). Within this representation, the sum over the four populations is equal to one for each temperature.

In Figure 2, the temperature dependence of the four microscopic water structures is shown. Among the different water models, the qualitative behavior is strikingly similar. Three main types of temperature scalings were observed: increasing population with decreasing temperature (enthalpi-
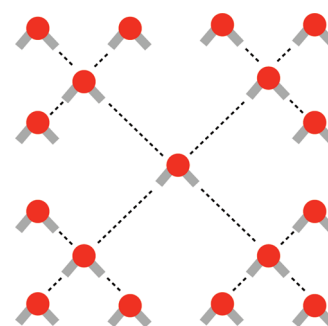

Figure 1. Schematic representation of the fully coordinated water configuration (P_4 population, see text). Dashed lines represent hydrogen bonds.

cally stabilized); increasing population with increasing temperature (entropically stabilized); with a maximum, where a turnover between enthalpic and entropic stabilization takes place at a model-dependent temperature. All four water configurations fall into one of these three main classes. The population of the fully ordered structure, P_4, increases with decreasing temperature (Figure 2, red empty circles). Consequently, this configuration is enthalpically stabilized. This is not the case when defects in the hydrogen bond structure are introduced (P_4^*, filled red circles). For this configuration, the population increases with decreasing temperature until it reaches a maximum in correspondence to a rapid increase of the population of the fully coordinated configuration. The maximum is located in a temperature range close to the temperature of maximum density of the model under consideration (dashed vertical line). Finally, both P_3 and P_{210} are mainly entropically stabilized, showing larger populations at higher temperatures. Taken together, these results indicate that specific water configurations dominate at each temperature range: full coordination extending to at least two solvation shells at low temperatures, four-coordinated configurations with no spatial extension at intermediate temperatures, and mainly disordered ones at higher temperatures.

Despite these similarities, an important difference among the models is the temperature range at which the relative configurations become dominant. For example, the maximum population of P_4^* for the SPC model was observed around ~245 K (Table 1). This is not the case for TIP4P/2005, where the maximum is located at a 40 K larger temperature. The same behavior was observed comparing the temperatures at which P_4 and P_3 are equal (e.g., around 270 K for TIP4P/2005). These observations suggested that a temperature shift factor (ΔT_s) exists among the models. Using TIP4P/2005 as a reference, we found a temperature shift factor for each model ranging from 65 to 6 K (see Figure 3 and Table 1). TIP4P/2005 was chosen as a reference for its ability to reproduce the density curve.[6] Applying this shift to the data allowed the superposition of all models onto four master curves, one for each structural configuration, as shown in the monochrome plot at the bottom right of Figure 2. Our observation is consistent with previously found phase diagram shifts among different water models[29,30] as well as in the presence of ions,[31] but in this case we could superimpose all models onto a master curve. Unfortunately, TIP5P had to be excluded from the superposition because all points show an increased curvature with respect to the other models, consistent with the increased curvature of the isobaric density at 1 atm.[8] The structural temperature shift is larger for
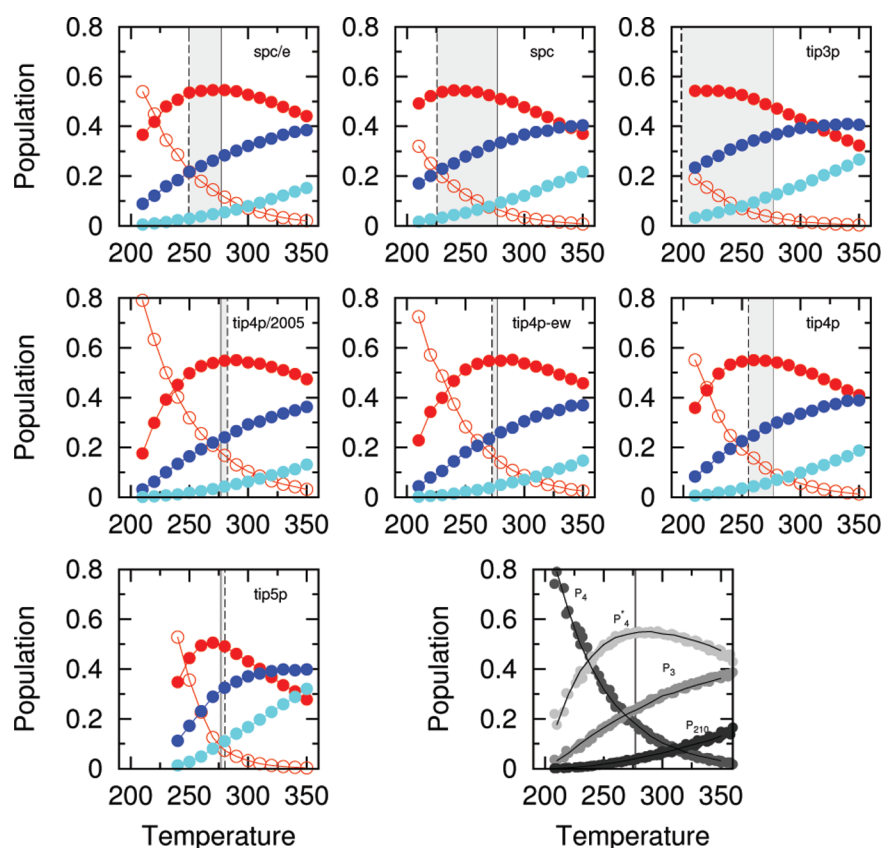
Figure 2. Temperature dependence of water structure populations for seven classical water models. P_4, P_4^*, P_3, and P_{210} are shown in red empty, filled red, blue, and light blue circles, respectively (see text for details). The gray stretch highlights the temperature difference between the calculated position of the temperature of maximum density (vertical dashed line, see also Table 1) and the experimental value at 277 K (solid line). The bottom right monochrome plot shows the superposition of all models after temperature shifting each data set (TIP4P/2005 data was used as reference). For each temperature, the sum over the four groups is equal to one.
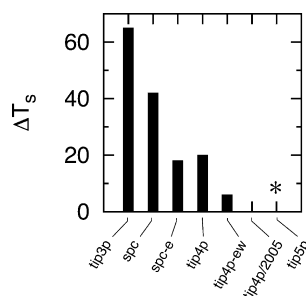

Figure 3. Structural temperature shift ΔT_s with respect to the TIP4P/2005 model. TIP5P was excluded from the superposition analysis (see text for details).

Figure 4. Overlap of the P_i data when a Skinner definition[32] for the hydrogen bond is used.

three-site models (Figure 3) with a spread of up to 65 K for TIP3P. On the other hand, four-site models deviate less. In general, models providing a better estimation of the position of the density maximum deviate less.

To check the robustness of the P_i overlap with the hydrogen bond definition, the recent definition of Skinner and collaborators[32] was applied. Figure 4 shows that the overlap between the curves is independent from the hydrogen bond definition. Moreover, the temperature shifts calculated in this case are very similar to the ones reported in Table 1. It is interesting to note that this bond definition is very different from the one used in Figure 2, being based on an empirical fit of the electronic occupancy.[32]
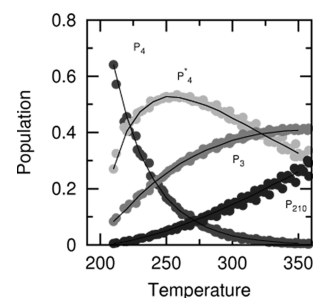
B. Stabilization of the Fully Coordinated Configuration. At all temperatures, water models with smaller shifts provide an improved stabilization of the fully coordinated configuration. (Alternatively, it can be said that these models destabilize poorly hydrogen-bonded configurations.) To make this point clearer, the free energy of the fully coordinated configuration at 230 K was calculated (Figure 5a and Methods). At this temperature P_4 is appreciably large for all water models. Comparison with the temperature shifts of Figure 3 indicates a remarkable correlation where even the small differences between SPC-E and TIP4P are respected. The correlation decreases when considering only the enthalpy, as shown in Figure 5b (see Methods). Interestingly, enthalpy and free energy correlate very well within the same model family. This is
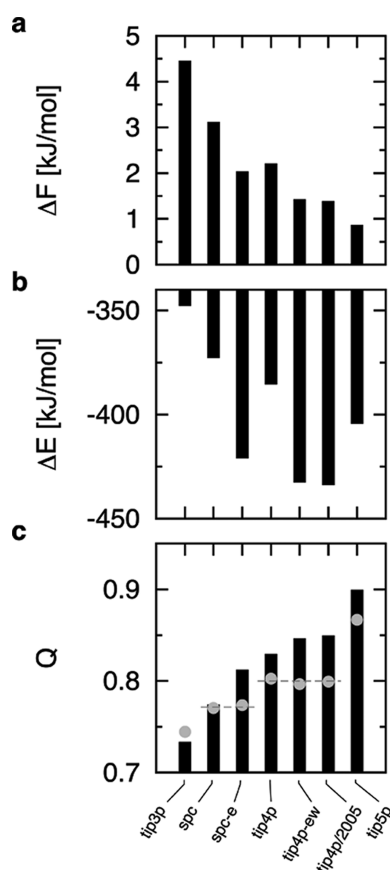
Figure 5. Comparison of water models with respect to the fully coordinated configuration at 230 K. (a) The value of the free energy. (b) Average enthalpy. (c) Average value of the tetrahedral order parameter Q. The value of Q for the case $P_4 = 0.2$ is shown as gray-filled circles.

Figure 6. Comparison between the structural temperature shifts (ΔT_s) and the position of the density maximum ($\Delta T_{density}$). Four-site models were compared to TIP4P/2005 (filled circles). Three-site models were compared to SPC/E (empty circles). Crosses refer to the case when TIP4P/2005 was used as a reference for the three-site models.

energy and enthalpy for the different water models (Figure 5a,b).

To better elucidate the connection with the temperature shifts, the radial distribution function (RDF) was calculated. At 270 K the RDF for the various models shows a different structural signature (Figure 7a). Within the same family, only

Figure 7. Radial distribution function (RDF) (a) at 270 K and (b) after the application of the temperature shift ΔT_s (see Table 1), taking the TIP4P/2005 model at 270 K as reference. TIP3P, SPC, and TIP4P families are shown in red, light blue, and blue, respectively.

particularly clear when looking at the three site models (i.e., the trend for TIP3P, SPC, and SPC/E), suggesting a different entropic contribution between three- and four-site models which is systematic.

Finally, the average value of the tetrahedral order parameter Q of the fully coordinated configuration calculated at the same temperature is shown in Figure 5c. In a first approximation, the parameter correlates well with the structural shift although not as good as the free energy.

C. Comparison with the Position of the Density Maximum. It is worth commenting on the relation between the structural temperature shifts found in this work and the model-dependent temperature of maximum density. As shown in Figure 6 and Table 1, the relationship between the structural ΔT_s and the density $\Delta T_{density}$ temperature shifts is linear within the three- or the four-site models (filled and empty circles in Figure 6). However, when comparing all models together using TIP4P/2005 as a reference, a small systematic deviation is observed (filled circles and crosses in the figure). This is due to the relation that exists between the populations P_i and the density. It is noted that the relative position of the P_4^* maximum with respect to the temperature of maximum density (dashed line in Figure 2, see also Table 1) depends on the model family. For four-site models, the two temperatures are identical, while for three- and five-site models the maximum of P_4^* is found at a higher and a lower temperature, respectively. This behavior might be connected with the systematic deviations between free
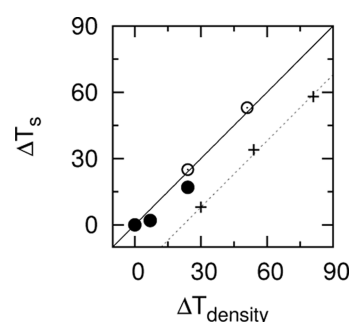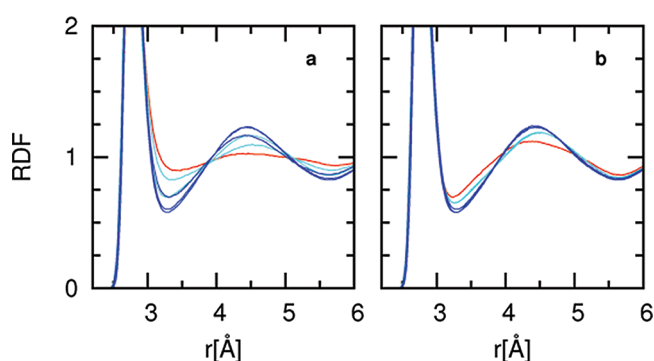
the curves for TIP4P/2005 and TIP4P-Ew overlap, the two models being very similar. In Figure 7b, the RDF was recalculated at temperatures shifted according to ΔT_s using as a reference TIP4P/2005 at 270 K. The figure shows that all models with the same geometry perfectly overlap (e.g., TIP4P, TIP4P-Ew, and TIP4P/2005, blue lines), suggesting that model reparametrizations act as an effective shift in temperature space. On the other hand, changes in the geometry or the number of sites affect the general shape of the radial distribution function and consequently the density.

Similar conclusions can be deduced when calculating the tetrahedral order parameter Q for temperatures at which the fully coordinated structure has the same probability for all models (P_4 around 0.2, gray-filled circles in Figure 5c). Q takes the same value within a given family, but it is influenced by the change of the molecular geometry, indicating that the structure corresponding to fully coordinated waters depends on the model family.

IV. CONCLUSIONS

In conclusion, we found that seven among the most used classical water models are characterized by very similar hydrogen-bond structure-forming capabilities up to a temperature shift. All models but TIP5P perfectly overlap onto a master curve when this shift is applied. This behavior does not depend on the hydrogen-bond definition. Our findings suggest that model reparametrization acts as an effective shift in temperature space. On the other hand, changes in the geometry or the number of sites cannot be fully reconducted to temperature shifts alone as shown by the analysis of the density as well as the radial distribution function. As such, although the hydrogen bond topology is universal when applying a certain temperature shift, this is not the case for the structure, each model family being characterized by its own signature.

We found that the three water models optimized to reproduce the position of the density maximum (i.e., TIP4P-Ew, TIP4P/2005, and TIP5P) systematically improve the stabilization of fully coordinated water configurations with an extension of at least two solvation shells. On the basis of this observation, we speculate that the improvements of these models for biomolecular simulations[14,15] are connected to the higher stabilization of *ordered* water. This property has important implications for the solvation of biomolecules, changing the balance between solute–solute and solute–solvent interactions.

Development of improved force fields strongly depends on this balance. Our analysis provides a microscopic and reductionist approach to face this challenge.

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: francesco.rao@frias.uni-freiburg.de.

Notes
The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Levy, Y.; Onuchic, J. N. Water mediation in protein folding and molecular recognition. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35* (1), 389−415.

(2) Papoian, G. A.; Ulander, J.; Wolynes, P. G. Role of water mediated interactions in protein-protein recognition landscapes. *J. Am. Chem. Soc.* **2003**, *125* (30), 9170−9178.

(3) Hummer, G. Molecular binding: Under water's influence. *Nat. Chem.* **2010**, *2* (11), 906−907.

(4) Ahmad, M.; Gu, W.; Geyer, T.; Helms, V. Adhesive water networks facilitate binding of protein interfaces. *Nat. Commun.* **2011**, *2*, 261.

(5) Thirumalai, D.; Reddy, G.; Straub, J. E. Role of water in protein aggregation and amyloid polymorphism. *Acc. Chem. Res.* **2011**, *45*, 83−92.

(6) Abascal, J. L. F.; Sanz, E.; Vega, C. Triple points and coexistence properties of the dense phases of water calculated using computer simulation. *Phys. Chem. Chem. Phys.* **2009**, *11*, 556−562.

(7) Aragones, J. L.; MacDowell, L. G.; Siepmann, J. I; Vega, C. Phase diagram of water under an applied electric field. *Phys. Rev. Lett.* **2011**, *107*, 155702.

(8) Vega, C.; Abascal, J. L. F. Relation between the melting temperature and the temperature of maximum density for the most common models of water. *J. Chem. Phys.* **2005**, *123* (14), 144504.

(9) Abascal, J. L. F.; Vega, C. A general purpose model for the condensed phases of water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123* (23), 234505.

(10) Kell, G. S. Density, thermal expansivity, and compressibility of liquid water from 0°c to 150°c. correlations and tables for atmospheric pressure and saturation reviewed and expressed on 1968 temperature scale. *J. Chem. Eng. Data* **1975**, *20* (1), 97−105.

(11) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79* (2), 926−935.

(12) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. Interaction models for water in relation to protein hydration. *Intermolecular Forces* **1981**, *11* (Suppl. 1), 331−338.

(13) Florová, P.; Sklenovský, P.; Banáš, P.; Otyepka, M. Explicit Water Models Affect the Specific Solvation and Dynamics of Unfolded Peptides While the Conformational Behavior and Flexibility of Folded Peptides Remain Intact. *J. Chem. Theory Comput.* **2010**, *6* (11), 3569−3579.

(14) Best, R. B.; Mittal, J. Protein simulations with an optimized water model: cooperative helix formation and temperature-induced unfolded state collapse. *J. Phys. Chem. B* **2010**, *114* (46), 14916−14923.

(15) Nerenberg, P. S.; Head-Gordon, T. Optimizing protein-solvent force fields to reproduce intrinsic conformational preferences of model peptides. *J. Chem. Theory Comput.* **2011**, *7*, 1220−1230.

(16) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **1987**, *91* (24), 6269−6271.

(17) Jorgensen, W. L.; Madura, J. D. Temperature and size dependence for monte carlo simulations of TIP4P water. *Mol. Phys.* **1985**, *56* (6), 1381−1392.

(18) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120* (20), 9665−9678.

(19) Mahoney, M. W.; Jorgensen, W. L. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys.* **2000**, *112* (20), 8910−8922.

(20) Rao, F.; Garrett-Roe, S.; Hamm, P. Structural inhomogeneity of water by complex network analysis. *J. Phys. Chem. B* **2010**, *114* (47), 15598−15604.

(21) Garrett-Roe, S.; Perakis, F.; Rao, F.; Hamm, P. Three-dimensional infrared spectroscopy of isotope-substituted liquid water reveals heterogeneous dynamics. *J. Phys. Chem. B* **2011**, *115* (21), 6976−6984.

(22) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005**, *26* (16), 1701−1718.

(23) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81* (8), 3684−3690.

(24) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126* (1), 014101.

(25) 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.

(26) Brovchenko, I.; Geiger, A.; Oleinikova, A. Liquid-liquid phase transitions in supercooled water studied by computer simulations of various water models. *J. Chem. Phys.* **2005**, *123* (4), 044515.

(27) Errington, J. R.; Debenedetti, P. G. Relationship between structural order and the anomalies of liquid water. *Nature* **2001**, *409* (6818), 318−321.

(28) Luzar, A.; Chandler, D. Effect of Environment on Hydrogen Bond Dynamics in Liquid Water. *Phys. Rev. Lett.* **1996**, *76*, 928−931.

(29) Sanz, E.; Vega, C.; Abascal, J. L. F.; MacDowell, L. G. Phase diagram of water from computer simulation. *Phys. Rev. Lett.* **2004**, *92*, 255701.

(30) Vega, C.; Sanz, E.; Abascal, J. L. F. The melting temperature of the most common models of water. *J. Chem. Phys.* **2005**, *122* (11), 114507.

(31) Corradini, D.; Gallo, P. Liquid-liquid coexistence in nacl aqueous solutions: a simulation study of concentration effects. *J. Phys. Chem. B* **2011**, *115* (48), 14161−14166.

(32) Kumar, R.; Schmidt, J. R.; Skinner, J. L. Hydrogen bonding definitions and dynamics in liquid water. *J. Chem. Phys.* **2007**, *126* (20), 204107.

(33) Rick, S. W. A reoptimization of the five-site water potential (tip5p) for use with ewald sums. *J. Chem. Phys.* **2004**, *120* (13), 6085−6093.

(34) Lisal, M.; Kolafa, J.; Nezbeda, I. An examination of the five-site potential (tip5p) for water. *J. Chem. Phys.* **2002**, *117* (19), 8892−8897.