Structural Inhomogeneity of Water by Complex Network Analysis

Francesco Rao,*,†,‡ Sean Garrett-Roe,§ and Peter Hamm*,§

Freiburg Institute for Advanced Studies (FRIAS), University of Freiburg, Freiburg, Germany, Laboratoire de Chimie Biophysique/ISIS, Université de Strasbourg, Strasbourg, France, and Physikalisch-Chemisches Institut, Universität Zürich, Zürich, Switzerland

Received: July 1, 2010; Revised Manuscript Received: August 24, 2010

There is still an open debate regarding the structure forming capabilities of water at ambient conditions. To probe the presence of such inhomogeneities, we apply complex network analysis methods to a molecular dynamics simulation at room temperature. This study provides both a structural and quantitative characterization of kinetically homogeneous substates present in bulk water. We find that the conformation-space network is highly modular, and that structural properties of water molecules are spatially correlated over at least two solvation shells. From a kinetic point of view, the free energy surface is characterized by multiple heterogeneous metastable regions with different populations and marginal barriers separating them. The typical time scale of hopping between them is 200−400 fs. A scanning in temperature reveals that those substates can be stabilized either entropically or enthalpically. The latter resembles an icelike domain that extends for at least two solvation shells.

I. Introduction

The many abnormalities of water, such as the density maximum at 4 °C, are related to the complex and highly dynamic network of hydrogen bonds it can form. Starting with a paper by Röntgen more than a century ago,[1] there has been a long-standing debate about the relative importance of local icelike structures, both in bulk water[2−5] and in the context of the hydrophobic effect.[6,7] The density maximum at 4 °C is a manifestation of the competition between two driving forces, i.e., the directionality of hydrogen bonding that favors a low-density, tetrahedral (icelike) form of water to minimize enthalpy, and maximizing entropy by nondirectional hydrogen bonds and disorder, resulting in a closer packing. Furthermore, recent small-angle X-ray scattering evidenced that water is structurally inhomogeneous with spatial correlations over ≈10 Å.[8] Even though these mesaurements were performed between room temperature and close to the boiling point, the experiments were discussed in the context of the liquid-to-liquid phase transition of water in the deeply supercooled regime, i.e., in the so-called "No Man's Land" of the (p,T) phase diagram that is, as of yet, not accessible experimentally.[9]

From a theoretical point of view, the tetrahedral order parameter,[10] pairwise and higher order orientational correlation functions,[11] and local hydrogen bonding[12] have been used to characterize correlations between neighboring water molecules. However, as with any complex molecular system, it is very difficult to physically capture the structural properties of water with the help of an a priori choice of one or a few geometric order parameters.

In the case of protein folding, the archetypal complex system, a new arsenal of graph-based theoretical tools have been recently developed to circumvent this limitation.[13−15] Network approaches have pointed out the presence of multiple stable structures in the unfolded state of a protein, which is a somewhat surprising result, giving the fact that folding often appears to be a two-state process.[13,16] Complex network analysis methods, which initially have been inspired by social and computer sciences (e.g., analysis of friendship networks or the World Wide Web[17,18]), have proven to be very useful to understand protein conformational changes in general.[19−22]

The essential idea of these methods is to map a molecular dynamics (MD) trajectory into a discrete set of *structurally* homogeneous microstates without assuming any a priori knowledge of the relevant degrees of freedom. In turn, the dynamics of the simulation is represented as a kinetic network (the conformation-space network) where the microstates and the transitions between them observed during the simulation represent the nodes (vertices) and the links (edges), respectively. Conformation-space networks represent a high resolution mapping of the underlying free-energy surface where highly interconnected groups of nodes indicate *kinetically* homogeneous free-energy valleys.[16,20]

The generalization of this approach to describe the thermodynamics as well as the kinetics of bulk water is difficult. In contrast to a protein, a missing definition of microstates and the intrinsically entropic nature of a liquid have prevented the application of the above graph-based methods. In this work, we present such a generalization by introducing the conformation-space network of water at room temperature. We found that bulk water is characterized by a number of substates that are normally hidden in conventional approaches. Our study provides a quantitive description of those substates in terms of their thermodynamics, kinetics, and structural properties.

II. Materials and Methods

The work is based on 100 ps long MD trajectories of SPC water (except where noted otherwise), using the Gromacs program package,[23] in the *NVT* ensemble at 300 K, 1019 water molecules at experimental density, 1 fs time step with a saving frequency of 4 fs, periodic boundary conditions, with the long-range electrostatic forces approximated by the particle mesh

* Corresponding authors. E-mail: P.H., p.hamm@pci.uzh.ch; R.F., francesco.rao@frias.uni-freiburg.de.
† University of Freiburg.
‡ Université de Strasbourg.
§ Universität Zürich.

10.1021/jp1060792
Published on Web 09/21/2010

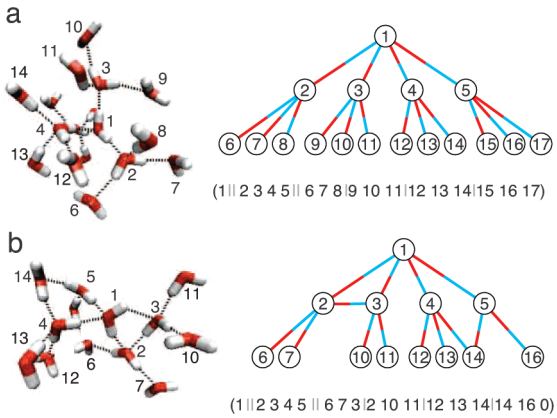

Figure 1. Water microstates. (a) Conformation in which all four hydrogen-bonding sites of each water molecule connect to new water molecules, and the corresponding microstate string. Water molecules are numbered according to their appearance in the tree search, and water molecules from subsequent generations are placed next to each other. The ||'s separate solvation shells; the |'s separate each hydrogen-bonded group. Hydrogen sites are depicted in blue; oxygen sites, in red. (b) If a hydrogen-bonding site is empty (e.g., molecule 5), it is labeled as 0, as are all subsequent entries down the tree. Small loops, such as 1−2−3, are included in a natural fashion.

Ewald approximation. The hydrogen-bond definition that we employ is based on the work of Skinner et al.[24,25] who found an empirical correlation between the occupancy N of the O−H σ^* orbital (electron density donated by the hydrogen-bond acceptor) from electronic structure calculations and the geometries observed in MD simulations. They parametrized the orbital occupancy in terms of the H···O intermolecular distance r and the angle ψ that the O−H ray makes with the out-of-plane unit vector of the acceptor molecule

$$N(r,\psi) = \exp(-r/0.343\ \text{Å})(7.1 - 0.050\psi + 0.00021\psi^2) \quad (1)$$

and consider a hydrogen bond to be formed if this occupancy exceeds a certain threshold ($N > 0.0085$).

Alternatively, we also tested a simple geometry criterion in which a hydrogen bond is said to be formed if the O···O distance is smaller than 3.3 Å and the angle between the O−H and the O···O rays is smaller than 30°. The essential layout of the conformation space network does not depend on the hydrogen-bond criterion (Supporting Information, Figure S4).

III. Conformation-Space Network of Water

Water is the classic associating liquid; the directionality of hydrogen bonding significantly changes the structure of the fluid, e.g., the number of nearest neighbors, relative to simple liquids like xenon. Hence, we adopted the view that it is the topology of hydrogen-bond network around a given water molecule that determines the structural and dynamical properties of the bulk. However, the binding partners to any central water molecule are not predefined but keep exchanging on a fast picosecond time scale. Therefore, any approach to define a microstate must be invariant to interchanging water molecules, as well as binding sites. Figure 1 shows two examples of hydrogen-bond networks, as they appear during a molecular dynamics (MD) simulation. In some cases, all four binding sites of each water molecule connect to new water molecules (Figure 1a), but in others there are loops and/or broken hydrogen bonds (Figure 1b). We describe each structure by a unique string that encodes the connectivity through hydrogen bonds. For each water molecule, we search for neighboring water molecules that hydrogen-bond to it.[24,25] From these waters of the first solvation shell, the search expands in a treelike manner (Figure 1 right). We assumed that each water molecule has a maximum of four binding sites, one to each of the two hydrogens (blue lines in Figure 1 right) and two to the oxygen (red). For a fully hydrogen-bonded network of a central water and two solvation shells, there are then a maximum of 17 possible waters. In the relatively rare occasions when two water molecules hydrogen-bond to one of the hydrogens, or three water molecules hydrogen-bond to the oxygen, we selected those with the larger occupancy numbers. We did, however, explicitly consider structures with fewer than four hydrogen-bond partners (e.g., molecule 5 in the example, Figure 1b). Each subsequent solvation shell is a new generation and follows, in order, in the microstate string, numbered by their position in the fully hydrogen-bonded tree (Figure 1a, right). The order of appearance of water molecules at the two oxygen binding sites is arbitrary. We resolved this symmetry by determining all possible permutations of interchanging the labeling of oxygen binding sites and uniquely chose one string to represent them all. For example, (1||0 0 4 5||0 0 0|0 0 0|12 13 14|0 0 0) and (1||0 0 4 5||0 0 0|0 0 0|0 0 0|15 16 17) encode identical conformations, and we choose the first. In contrast, we distinguished the two hydrogen sites, which could be achieved in a real experiment by isotope labeling. We had to truncate the tree after the second generation (second solvation shell), as we found that further generations make the number of microstates unmanageably large (>10^7 instead of 5×10^4 for a 100 ps trajectory).

Each hydrogen-bond-network string represents a *microstate* of a single water molecule. The analysis of the time evolution of those microstates allows a network description of the system, in analogy to refs 13 and 20. Two microstates (nodes) are linked if they appear subsequently in the trajectory; nodes and links are weighted by how many times a microstate appears and a transition occurs, respectively. This conformation-space network is a representation of a kinetic matrix. It was shown that nodes connected with each other by many edges interconvert rapidly and are kinetically homogeneous, while poorly connected nodes are kinetically separated.[20] Network visualization represents a successful approach to infer such kinetic properties, at least at a first qualitative level.[13] Here, to obtain a three-dimensional layout of the conformation-space network of water, we apply multidimensional scaling (MDS),[26−28] which is essentially the same as a principal coordinate analysis[29,30] (both would be identical if the metric is Eucledian[28]). In general, the MDS algorithm starts from a matrix of pairwise distances (restraints) between nodes. Consider four points in 3D space, then a matrix of mutual distances defines these points uniquely except for an overall translation and rotation. The same holds for N points in $(N - 1)D$ space (in case the distance matrix would be Euclidean), i.e., represents an object embedded in many dimensions (in our case several thousands). The aim of the MDS algorithm is then to reduce the effective dimensionality of the full set of restraints by spectral analysis, i.e., combining those dimensions that are redundant in the restraints space. In the most simple implementation of the MDS algorithm, distance would be the minimum number of links between two nodes.[26,27] In contrast, here we give distances the meaning of a traveling time, so the MDS algorithm will lump together those distances that correspond to fast interconversions (i.e., grouping kinetically homogeneous nodes) while separating slowly relaxing nodes.

Having this physical picture in mind, we choose for the mutual distances between two directly linked nodes, τ_{ij}, the time it takes for two isolated nodes to equilibrate (assuming first-order kinetics):

$$\tau_{ij} = \frac{2\Delta t}{w_{ij}}\left(\frac{1}{p_i} + \frac{1}{p_i}\right)^{-1} \quad (2)$$

where, p_i and p_j are the statistical weights of nodes i and j, respectively, w_{ij} is that of their link, and Δt is the saving time of trajectory snapshots (4 fs). If two nodes are not directly linked, the pathway with the shortest total distance τ_{ij} is chosen, taking the sum $\tau_{ij} = \tau_{ik} + \tau_{kj}$. A pathway over more than one link might be faster than a direct link. If that is the case, we replace the length of the direct link by that of the faster indirect pathway, thereby enforcing a triangle inequality with $\tau_{ij} \leq \tau_{ik} + \tau_{kj}$.

From an analysis of the largest MDS eigenvalues (Supporting Information, Figure S1a), we conclude that the first three dimensions carry most of the essential information, whereas subsequent eigenvalues decay exponentially and represent only Gaussian noise. This means that, in the present case, a three-dimensional layout of the network is sufficient to capture the most important properties of the dynamics. Note that the distance metric τ_{ij} is not Euclidian, as evidenced by negative MDS eignevalues (the largest negative eigenvalue is about 20% of the largest positive one).

IV. Results and Discussion

The three-dimensional MDS layout of the conformation-space network of SPC water at 300 K is highly modular with clusters of kinetically homogeneous microstates (Figure 2a,b). The number of hydrogen bonds at the hydrogen sites (encoded by the colors in Figure 2a,b: blue, two; red and green one; yellow, none) as well as at the oxygen site (encoded by the color saturation) determines to a certain extent which cluster a node belongs to, but that assignment is not complete. In particular, there is a substructuring within fully hydrogen-bonded states (blue nodes), i.e., the four fingers pointing down in Figure 2a. In other words, while the hydrogen bonding to the first solvation shell (represented by the color coding) plays an important role in determining the dynamic properties of a given water molecule, this information alone is not sufficient, and the topology of the hydrogen-bond network including more distant water molecules contributes as well. These results are obtained using the SPC model but we show in the Supporting Information, Figures S2−S4, that the essential structure of the conformation-space network does not depend on the water model (i.e., SPC, SPC/E, TIP4P, TIP5P), the length of the MD trajectory (from 10 ps to 1 ns), or the hydrogen-bond criterion (i.e., the one introduced by Skinner et al.[24,25] versus a simple geometrical cutoff criterion, see Material and Methods).

Toward a more quantitative analysis of the free-energy surface in terms of kinetically homogeneous regions, or substates, we use a recently introduced procedure to partition the landscape into free-energy basins[31] (which is a simpler version of the stochastic steepest descent algorithm introduced in ref 32). To that end, we build a transition-gradient network from the original one by keeping only one link per node, the one with the highest weight (excluding self-links). The physical justification for doing so is that nodes in the neighborhood of a barrier between two valleys will maintain only the link to the valley to which they relax the fastest. Following the most probable transition results in a "steepest descent" on the free-energy surface where microstates in the neighborhood of a barrier will connect to either one basin or another. We call the resulting basins of attraction *gradient-clusters* for convenience. Most of the conformation network can be represented by the $M = 12$ most populated gradient clusters, which account for the 88% of the total population (at 300 K; see Figure 3d). Subsequent gradient clusters have significantly smaller populations (Supporting Information, Figure S1b). Although this approach is completely independent from the MDS analysis, we find a remarkable mapping between the two, as shown in Figure 2c, where network nodes (i.e., water microstates) belonging to one of the 12 most populated gradient clusters are shown in different colors. The properties of these gradient clusters are summarized in Table 1.

The barrier between two gradient-clusters a and b, composed of nodes a_i and b_i, is computed as

$$\Delta G^{\ddagger}_{\mathrm{a}\to\mathrm{b}} = -k_{\mathrm{B}}T \log\left(\frac{W_{\mathrm{a}\to\mathrm{b}}}{\sum_{\mathrm{a}_i \neq \mathrm{a}_0} w_{\mathrm{a}_0\to\mathrm{a}_i}}\right) \quad (3)$$

where $W_{\mathrm{a}\to\mathrm{b}}$ represents the number of transitions between cluster a and b (i.e., $\sum_{\mathrm{a}_i}\sum_{\mathrm{b}_j} w_{\mathrm{a}_i\to\mathrm{b}_j}$), a_0 is the most populated node in cluster a, and $\sum_{\mathrm{a}_i \neq \mathrm{a}_0} w_{\mathrm{a}_0\to\mathrm{a}_i}$ is the sum of the outgoing transitions from node a_0, excluding the self-link, to all the other nodes of cluster a. This definition of a reaction barrier is the same used in a cut-based free-energy profile (cFEP)[33] and relates the reactive flux between basins to the flux around the minimum of a basin. With that definition, we find that the complete conformation-space-network can be traversed over only marginal barriers, which are on the order of $\lesssim k_{\mathrm{B}}T$ (Figure 3d). Nevertheless, certain gradient clusters are not directly connected. For example, gradient cluster 4, which is four-coordinated, has to react through an intermediate state (gradient cluster 2 or 3) to reach a hydrogen-bond broken cluster (gradient cluster 8 or 9, respectively; see Figure 3d).

The correlation function $\langle c_n(t)\, c_n(0)\rangle$ is calculated to investigate the lifetimes of those substates. In this formula, $c_n(t) = 1$ when the trajectory populates a gradient cluster n at time t, and $c_n(t) = 0$ otherwise (Supporting Information, Figure S5). The correlation functions decay biexponentially with a fast component in the range $\approx$30 −40 fs, which we attribute to recrossing noise at the edges of our gradient clusters, and a slower component between $\approx$200 and $\approx$400 fs, which reflects the actual lifetime of the cluster. Within classical transition state theory, the cluster lifetime is expected to be $k = k_0 \mathrm{e}^{-\Delta F^{\ddagger}/k_{\mathrm{B}}T}$. With a pre-exponential factor of $k_0 = k_{\mathrm{B}}T/h = (160\ \mathrm{fs})^{-1}$, the observed time scales are consistent with barriers $\Delta F^{\ddagger}$ on the order of $\lesssim k_{\mathrm{B}}T$.

In Figure 4a,b tetrahedral order parameter distributions[34] and radial distribution functions of the 12 most populated gradient clusters are shown. They group into three types, which can be classified as four-, three-, and two-coordinated water. For example, waters in gradient clusters 1−4 are all four-coordinated and hence more ordered than on average, causing a higher tetrahedral order parameter (Figure 4a), as well as a higher peak at 2.8 Å and a more pronounced minimum between first and second solvation shells in the radial distribution function (Figure 4b). Nevertheless, the 3D hydrogen and oxygen distributions of clusters with the same coordination number can be distinctively different; see, e.g., clusters 1−4 in Figure 5a. This emphasizes that the radial distribution functions are not decisive

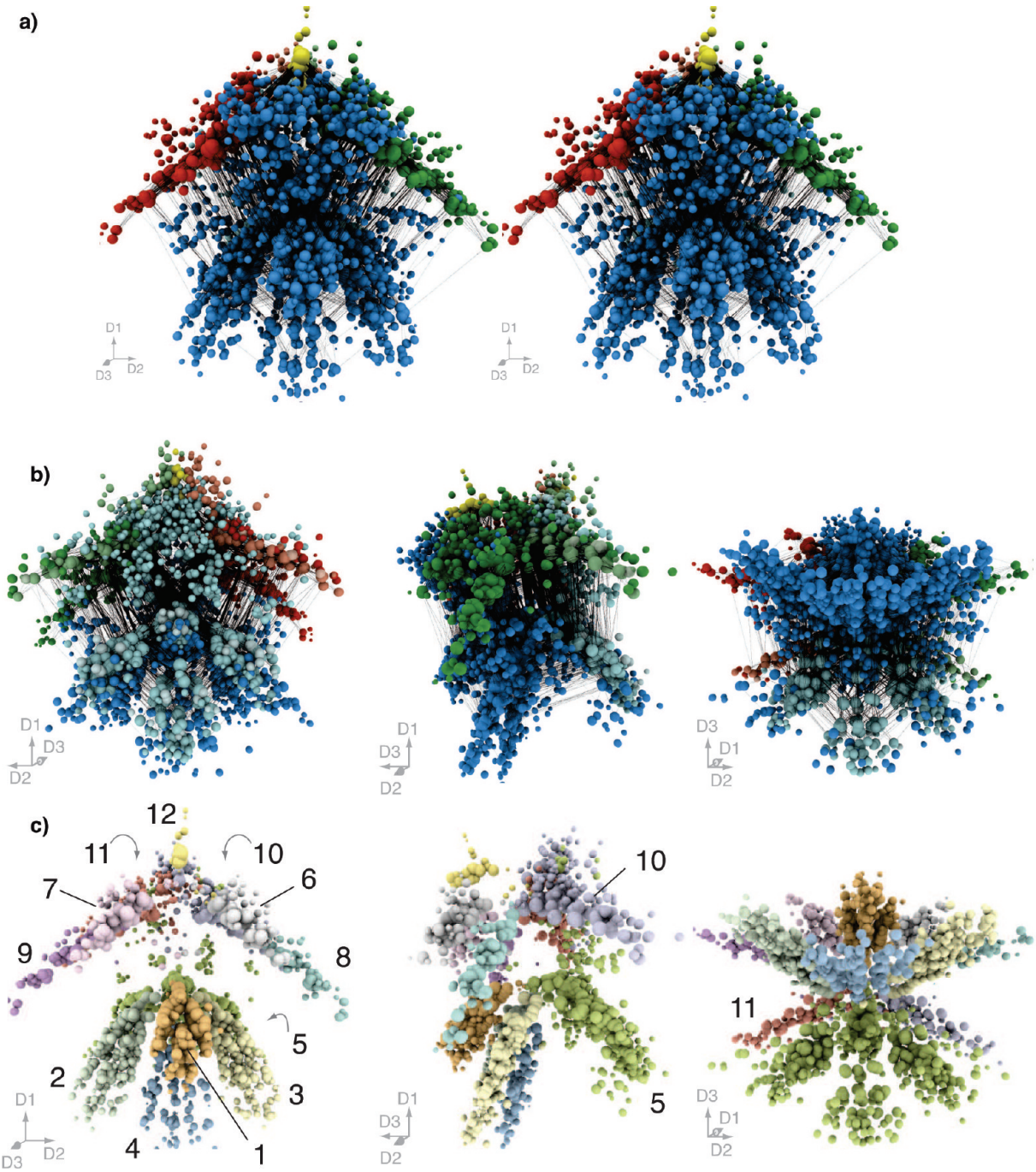

Figure 2. Conformation-space network of SPC water at 300 K from a 100 ps MD trajectory. To avoid overcrowding, we only considered nodes and links with a statistical weight above a certain threshold, >200 and >30, respectively. These two threshold values reduce the number of nodes from ~50 000 to ~4000, covering 95% of the total population, and reducing the number of links to typically 15 000−20 000. (a) Stereoview and (b) three orthogonal views of the conformation-space network embedded in three dimensions by the MDS algorithm. Each microstate is represented by a node; distances represent the transition times between microstates; the volume of each node represents the logarithm of its statistical weight (i.e., its free energy); the color represents the number of hydrogen bonds donated: yellow, none; green and red, one (by either one or the other hydrogen); blue, two (one per hydrogen). The color saturation encodes the number of hydrogen bonds accepted: full colors, two; light colors, one; very light colors, none. (c) Mapping of the 12 most important gradient clusters (Figure 3) onto the MDS representation, all in different colors. Nodes that do not belong to any of the 12 most populated gradient clusters are omitted.

enough to characterize the 3D structure of hydrogen-bonded water. Note that clusters 6 and 7, 8 and 9, as well as 10 and 11, are mirror images of each other with the left and right hydrogen-bond donation site interchanged (see Table 1). In some cases, the important information is hidden for the view angle chosen in the representation of Figure 5a. For example, clusters 5 and 10 look very similar in this representation, but when cluster 11 is inspected instead of cluster 10, the difference is very obvious.

In Figure 4c we investigate the conditional radial distribution functions $g_{O_nO_n}(r)$. That is, we plot the probability of finding a water from the cluster n next to a water from the same cluster n (whereas Figure 4b plots the probability $g_{O_nO}(r)$ of finding *any* water next to a water molecule from cluster n). With regard to this property, gradient clusters 1 and 4 deviate the strongest from the average. The probability of finding two neighboring molecules in gradient cluster 1 is significantly enhanced, while

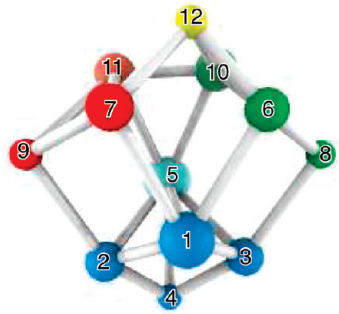

Figure 3. Coarse grained gradient clusters, labeled as in Table 1. The coarse grained gradient clusters are positioned according to the center of mass of the corresponding nodes in Figure 2c. Links represent barriers less than k_BT.

TABLE 1: Characterization and Relative Population (%) of the 12 Most Populated Gradient Clusters at Three Temperatures (at Constant Volume)

cluster:	1	2	3	4	5	6	7	8	9	10	11	12
h-bond:[a]	22	22	22	22	21	r2	l2	r2	l2	r1	l1	02
270 K	19.7	6.4	6.4	1.8	37[b]	5.4	5.4	1.7	1.7	*b*	*b*	1.3
300 K	13.2	5.9	5.9	2.0	23.4	5.3	5.3	2.2	2.2	10.1	10.1	1.9
330 K	9.5	5.3	5.3	2.6	23.7	5.1	5.1	2.6	2.6	10.4	10.4	2.4

[a] The first digit in the hydrogen-bond characterization refers to the number and site of donating hydrogen bonds (2, from the (l)eft or (r)ight hydrogen, 0); the second, to the number of accepting hydrogen bonds. [b] At 270 K, clusters 10 and 11 merge with cluster 5.

the opposite is true for gradient cluster 4. The local oxygen distribution shown in Figure 5 explains the mechanism. In cluster 1, neighboring waters are also four-coordinated (3.61 bonds on average; see arrows in Figure 5a for cluster 1), forming an extended tetrahedral hydrogen-bond network. This correlation manifests the onset of forming translational symmetry or, in simple words, the onset of ice forming. In accord with this view, gradient cluster 1 has the longest of all lifetimes (400 fs; see Figure S5, Supporting Information). In cluster 4, in contrast, neighboring waters are three-coordinated (2.55 bonds on average). This undercoordination creates a void or *groove* around the central water that is then filled by waters more distant in the hydrogen-bond network (see arrow in Figure 5a for cluster 4).

The population of cluster 1, which is the most ordered icelike structure, decreases by more than a factor of 2 from 270 to 330 K (see Table 1). Thus, this structure is stabilized enthalpically, as expected. From the 11 other structures, some slightly increase in population with temperature and some decrease, but the overall effect is that all increase to compensate for the population loss of cluster 1. The temperature dependence of the radial distribution function of clusters 1 and 4 (i.e., the most enthalpic and the most entropic cluster) is shown in Figure 6. The profiles are essentially conserved, becoming slightly more structured as the temperature is lowered. Consequently, the structural organization of the substates is conserved for different temperatures. This result suggests that the overall properties of water are mostly determined by the change of the relative populations of the various substates (Table 1).

Water is often discussed on the level of two states,[5] i.e., a low-enthalpy, low-density form, and a high-entropy, high-density form of water, which ultimately leads to the suspected liquid-to-liquid phase transition at low enough temperatures. Nevertheless, there has been experimental evidence that these two forms still exist at ambient temperatures or even up to the

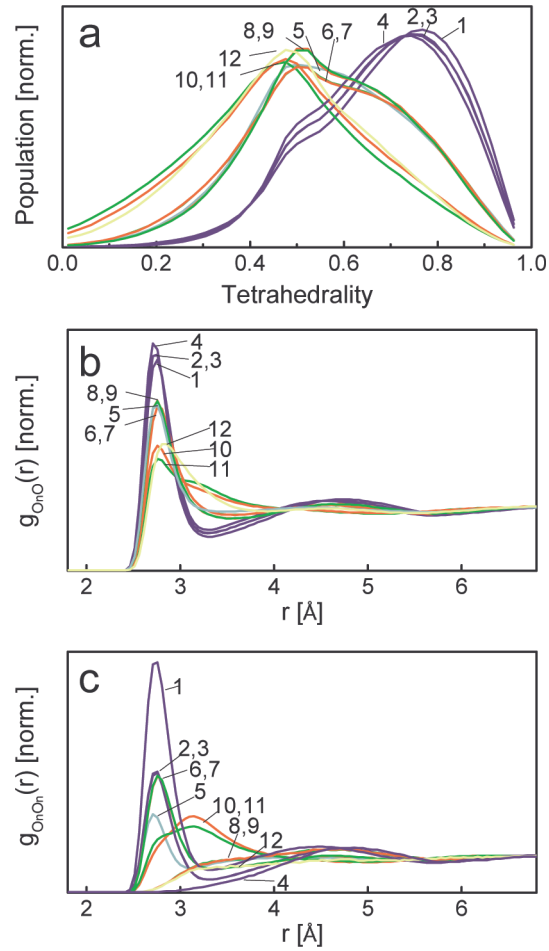

Figure 4. Structural characterization of the 12 most populated clusters at 300 K. (a) Oxygen−oxygen radial distribution functions $g_{O_nO}(r)$ and (b) the corresponding tetrahedral order parameter distributions[34] (same color code as in Figure 2a,b). (c) Conditional oxygen−oxygen radial distribution functions $g_{O_nO_h}(r)$, normalized by the cluster density.

boiling point, in the sense of fluctuating and quickly interconverting domains.[8] In light of this discussion, cluster 1 would be the low-enthalpy, low-density form, as judged form its temperature dependence (Table 1) and its structural properties (Figure 5), whereas most of the rest constitutes the entropically stabilized state. Our network analysis, however, indicates that the entropically stabilized state contains additional substructure. In future work we will investigate how this state evolved toward high-density water at lower temperatures. In ref 8 it has been argued that water domains have a spatial correlation length of ≈10 Å, but it has also been shown that SPC water cannot reproduce this large size. Our study nevertheless indicates that correlations exist up to at least the second solvation shell (≈5 Å). Due to the cutoff after the second generation in the tree search of hydrogen bonds, this method cannot describe the possible role of longer-distance correlations in bulk water.

V. Conclusion

In this study, we have investigated the conformation-space network of water at ambient conditions. Both the visual inspection of the network and a more quantitative analysis, based on transition gradients, have pointed out the presence of multiple coexisting metastable states in bulk water. A kinetic and thermodynamic description of the substates and their structural characterization have been provided. The latter has elucidated a critical difference in the second solvation shell between two

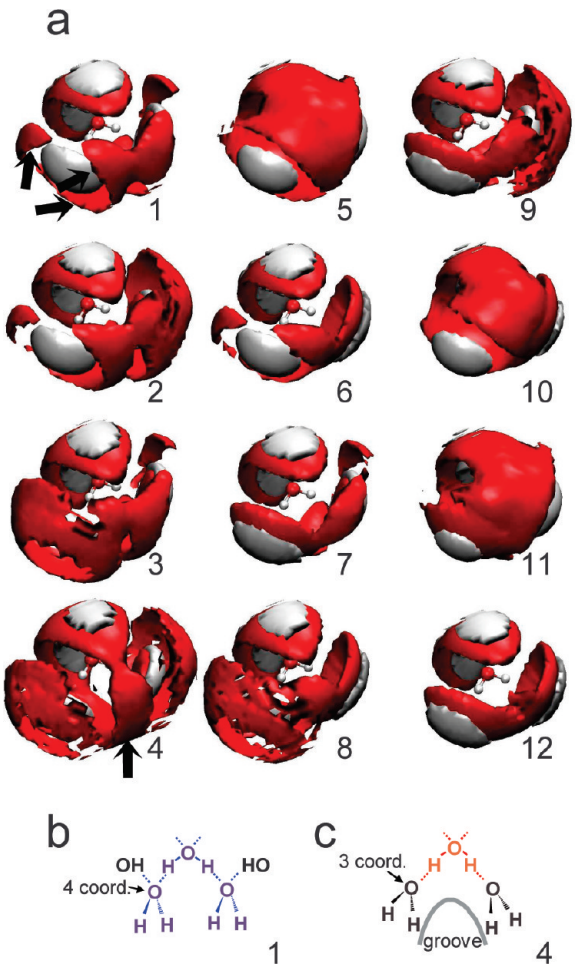

Figure 5. Water structures for the 12 most populated gradient clusters. (a) Oxygen and hydrogen densities. The arrows point to features that are discussed in the text. (b, c) Schematic hydrogen-bond networks in clusters 1 and 4, respectively.

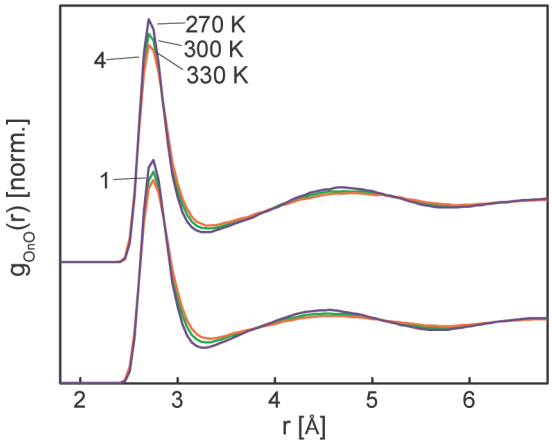

Figure 6. Normalized oxygen−oxygen radial distribution functions $g_{O_nO}(r)$ of cluster 1 (bottom) and 4 (top) as a function of temperature.

four-coordinated water clusters that may be reminiscent of a postulated liquid-to-liquid phase transition at low temperatures. The typical time scales of interconversion of the found metastable states is 200−400 fs.

We conclude that liquid water is structurally inhomogeneous, with spatial correlations that extend over at least two solvation shells. These observations are compatible with the known structure forming properties of water, which are responsible for the many forms of ice at lower temperatures (15 are now known[35]). In contrast to previous MD studies on the hydrogen-bond dynamics of water, which focused on the hydrogen bonds of a water molecule to its immediate partners solely,[25,36−38] we include here the influence of at least the second solvation shell, without having to invoke a predetermined order parameter. This achievement is made possible by introducing a network description of the dynamics.

The conformation-space-network structure is intrinsic to all water models we tested (SPC, SPC/E, TIP4P, TIP5P) (Supporting Information, Figure S2), suggesting evidence that it exists also in real water. It is interesting to note that an analysis of conformational entropy[39] provides evidence that SPC and TIP4P agree the best with experiment, and that these two models give the most structured conformation-space network (Supporting Information, Figure S2). 3D-IR spectroscopy, which is able to disentangle inhomogeneities of conformational degrees of freedom of water,[40] will allow us to test these ideas experimentally. Finally, we propose here a framework to characterize the inhomogeneities of water in terms of structure and dynamics in more complex environments, such as confined water, water around ions, and water at protein surfaces, as well as hydrogen-bond chains that are important in the context of the Grotthus mechanism of proton transport and the autoionization of water.[41]

Acknowledgment. This work was financially supported by Swiss National Science Foundation (SNF) through the NCCR MUST as well as by a fellowship to F.R. from the European Molecular Biology Organization and the Excellence Initiative of the German Federal and State Governments.

Supporting Information Available: Figures S1−S5 showing MDS eigenvalues and gradient cluster data, MDS conformation-space networks, and cluster lifetime autocorrelation function, as well as a movie rotating the 3D conformation-space-network of Figure 2a,b. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Röntgen, W. C. *Anal. Phys. Chem.* **1892**, *281*, 91–97.
(2) Senior, W. A.; Verrall, R. E. *J. Phys. Chem.* **1969**, *73*, 4242–4249.
(3) Woutersen, S.; Emmerichs, U.; Bakker, H. J. *Science* **1997**, *278*, 658–660.
(4) Laenen, R.; Rauscher, C.; Laubereau, A. *Phys. Rev. Lett.* **1998**, *80*, 2622–2625.
(5) Wernet, P.; Nordlund, D.; Bergmann, U.; Cavalleri, M.; Odelius, M.; Ogasawara, H.; Näslund, L. A.; Hirsch, T. K.; Ojamäe, L.; Glatzel, P.; Pettersson, L. G. M.; Nilsson, A. *Science* **2004**, *304*, 995.
(6) Frank, H. S.; Evans, M. W. *J. Chem. Phys.* **1945**, *13*, 507.
(7) Rezus, Y. L. A.; Bakker, H. J. *Phys. Rev. Lett.* **2007**, *99*, 148301.
(8) Huang, C.; Wikfeldt, K. T.; Tokushima, T.; Norlund, D.; Harada, Y.; Bergmann, U.; Niebuhr, M.; Weiss, T. M.; Horikawas, Y.; Leetmaa, M.; Ljundberg, M. P.; Takahashi, O.; Lenz, A.; Ojamäe, L.; Lubartsev, A. P.; Shin, S.; Pettersson, L. G. M.; Nilsson, A. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 15214–15218.
(9) Mishima, O.; Stanley, H. E. *Nature* **1998**, *396*, 329–335.
(10) Kumar, P.; Buldyrev, S. V.; Stanley, H. E. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 22130–22134.
(11) Lazaridis, T.; Karplus, M. *J. Chem. Phys.* **1996**, *105*, 4294–4316.
(12) Luzar, A.; Chandler, D. *Phys. Rev. Lett.* **1996**, *76*, 928–931.
(13) Rao, F.; Caflisch, A. *J. Mol. Biol.* **2004**, *342*, 299–306.
(14) Krivov, S. V.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 14766–14770.
(15) Caflisch, A. *Curr. Opin. Struc. Biol.* **2006**, *16*, 71–78.
(16) Muff, S.; Caflisch, A. *Proteins* **2008**, *70* (4), 1185–1195.
(17) Barabasi, A.-L.; Albert, R. *Science* **1999**, *286*, 509–512.
(18) See contributions in: *Science* **2009**, *325*, 405−432.
(19) Evans, D. A.; Wales, D. J. *J. Chem. Phys.* **2004**, *121*, 1080.
(20) Gfeller, D.; De Los Rios, P.; Caflisch, A.; Rao, F. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1817–1822.
(21) Noé, F.; Horenko, I.; Schütte, C.; Smith, J. C. *J. Chem. Phys.* **2007**, *126*, 155102.

(22) Rao, F.; Karplus, M. *Proc. Natl. Acad. Sci.* **2010**, *107*, 9152–9157.

(23) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26*, 1701–1718.

(24) Kumar, R.; Schmidt, J. R.; Skinner, J. L. *J. Chem. Phys.* **2007**, *126*, 204107.

(25) Auer, B.; Schmidt, J. R.; Skinner, J. L. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 14215–14220.

(26) Torgerson, W. S. *Psychometrika* **1952**, *14*, 401–419.

(27) Pich, C. Applications of Multidimensional Scaling to Graph Drawing. *Ph.D. thesis*, University Konstanz, see http://kops.ub.uni-konstanz.de/volltexte/2009/8399/, 2009.

(28) Cherkassky, V.; Mulier, F. M. *Learning from Data: Concepts, Theory, and Methods*; Wiley: New York, 2007.

(29) Becker, O. M. *Proteins* **1997**, *27*, 213–226.

(30) Gower, J. C. *Biometrika* **1966**, *53*, 325–338.

(31) Rao, F. *J. Phys. Chem. Lett.* **2010**, *1*, 1580–1583.

(32) Prada-Gracia, D.; Gomez-Gardenes, J.; Echenique, P.; Falo, F. *PLoS Comput. Biol.* **2009**, *5*, e1000415.

(33) Krivov, S. V.; Karplus, M. *J. Phys. Chem. B* **2006**, *110*, 12689–12698.

(34) Errington, J. R.; Debenedetti, P. G. *Nature* **2001**, *409*, 318–321.

(35) Salzmann, C. G.; Radelli, P. G.; Mayer, E.; Finney, J. L. *Phys. Rev. Lett.* **2009**, *103*, 105701.

(36) Luzar, A.; Chandler, D. *Nature* **1996**, *379*, 55–57.

(37) Laage, D.; Hynes, J. T. *Science* **2006**, *311*, 832–835.

(38) Eaves, J. D.; Loparo, J. J.; Fecko, C. J.; Roberts, S. T.; Tokmakoff, A.; Geissler, P. L. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13019–13022.

(39) Zielkiewicz, J. *J. Chem. Phys.* **2005**, *123*, 104501.

(40) Garrett-Roe, S.; Hamm, P. *Acc. Chem. Res.* **2009**, *42*, 1412–1422.

(41) Geissler, P. L.; Dellago, C.; Chandler, D.; Hutter, J.; Parrinello, M. *Science* **2001**, *291*, 2121.

JP1060792