



OPEN

Water clusters and density fluctuations in liquid water based on extended hierarchical clustering methods

Yitian Gao, Hongwei Fang[✉], Ke Ni & Yixuan Feng

The microscopic structures of liquid water at ambient temperatures remain a hot debate, which relates with structural and density fluctuations in the hydrogen bond network. Here, we use molecular dynamics simulations of liquid water to study the properties of three-dimensional cage-like water clusters, which we investigate using extended graph-based hierarchical clustering methods. The water clusters can cover over 95% of hydrogen bond network, among which some clusters maximally encompass thousands of molecules extending beyond 3.0 nm. The clusters imply fractal behaviors forming percolating networks and the morphologies of small and large clusters show different scaling rules. The local favored clusters and the preferred connections between adjacent clusters correspond to lower energy and conformational entropy depending on cluster topologies. Temperature can destroy large clusters into small ones. We show further that the interior of clusters favors high-density patches. The water molecules in the small clusters, inside which are the void regarded as hydrophobic objects, have a preference for being more tetrahedral. Our results highlight the properties and changes of water clusters as the fundamental building blocks of hydrogen bond networks. In addition, the water clusters can elucidate structural and density fluctuations on different length scales in liquid water.

Water is the most ubiquitous liquid involved in multifarious physical, chemical, and biological processes on earth^{1,2}. Different from simple liquids, liquid water shows numerous anomalous thermodynamic and kinetic behaviors, which are of significant importance to our planet and living systems^{3,4}. At macroscopic temporal or spatial scale, liquid water can be regarded as a homogeneous substance. However, at microscopic scale, frequent spatial-temporal fluctuations emerge in the hydrogen bond network with local tetrahedral symmetry, elucidating the main reasons for anomalous properties of water⁵.

Water anomalies are strongly associated with locally favored structures corresponding to particular long-lived molecular arrangements with local minima of free energy^{6,7}. Based on hydrogen-bonded networks, liquid water is a mixture of two different types of molecular arrangements: tetrahedral-like and hydrogen-bond distorted structures related with low- and high-density patches^{8–12}, respectively. The presence of density fluctuations in hydrogen bond networks in supercooled or liquid water has been widely investigated according to experiments^{13–18} and molecular dynamics^{12,19–22}. The model of two-state water, which comes from the competition between a high-density liquid (HDL) of distorted hydrogen bond patches and a low-density liquid (LDL) of ordered locally tetrahedral patches, can reproduce the anomalous behavior of thermodynamic properties of water^{6,7,22}. Density heterogeneity occurs on a length scale of 10–15 Å at ambient temperature^{14,15}. The presence of an isobestic point at approximately 3450 cm⁻¹ of the O–H stretching mode proves the coexistence of two types of local structures in liquid water by Raman spectroscopy¹³. In addition, X-ray small-angle scattering¹⁴, X-ray absorption¹⁷ and emission spectroscopy¹⁸ can provide evidence for inhomogeneous structures of liquid water. Based on molecular dynamics simulations, many models have been proposed to describe heterogeneous hydrogen bond networks to explain physical anomalous properties of liquid water, such as flickering models²³ and percolation models^{19,24}. Many structural descriptors have been proposed to quantify local favored structures in liquid water under various conditions, such as local structure index (LSI)²⁵, the distance of five nearest neighbors (d_5)²⁶, and the sphericity of the Voronoi cell²⁷.

State Key Laboratory of Hydro-science and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, China. ✉email: fanghw@tsinghua.edu.cn

Liquid water is a dynamical mixture of tetrahedral-like and “ring-and-chain”-like structures with a slight bias toward the former^{19,28}, forming three-dimensional hydrogen bond network continually undergoing topological reformation³. Water clusters, an assembly of weakly-bound water molecules, may be formed in hydrogen-bonded networks by locally favored structures with local energy minima^{29,30}. However, the microscopic structures of liquid water have remained a hot debate despite a lot of intensive research¹⁵. Graph theory is widely used to define topological structures as the building blocks of hydrogen-bonded networks^{31–37}. Many investigations have introduced the definitions of rings³³, fragments³⁴ and clusters³⁵. The rings, the cyclic paths along hydrogen bonds, show a broad distribution, among which 6-membered rings are most favored at ambient temperatures^{31–33,38}. Fragments, three-dimensional cage-like building blocks constituted by rings, can elucidate the heterogeneity of hydrogen bond structures and rearrangements³⁴. Topologically, water clusters also comprise small cyclic water clusters such as tetramers, pentamers and hexamers^{39,40}. The topology of water clusters strongly affects the structure and chemistry of clusters^{41,42}. The composite patterns of how flickering water clusters form bulk water remain a mystery despite several explorations^{43–46}. A hierarchical clustering method⁴⁷ is proposed to search out hierarchical water clusters in the hydrogen bond network. Hydrogen bonds, rings and fragments are defined as 1st-, 2nd-, and 3rd-level structures, respectively. Water clusters can be successively attained by using Louvain algorithm based on the network of fragments.

The local topological structures are closely related with density heterogeneity in liquid water at microscopic scale⁴⁸. Many special structural patterns have been proposed to describe water structures and explain many anomalous properties^{44–46,48–52}. Based on hydrogen bonds, fluctuating water clusters containing 280 water molecules with local icosahedral symmetry were proposed by Chaplin⁴⁶. The networks contain a mixture of hexamer and pentamer substructures and convert between lower- and higher-density forms without breaking hydrogen bonds. A pair of liquid structures, molecular chains as high-density liquid structures and fused dodecahedra as tetrahedral fluctuations, are defined to reproduce physical properties⁴⁹. In addition, a mixture of two types of helical clusters are also regarded as building blocks⁴⁴. The empty spaces with hydrogen bond networks are strongly related with the nature of structural heterogeneities in liquid water^{48–52}. Density fluctuations form a diverse morphology of voids, which can allow for the identification of low- and high-density patches of the liquid⁴⁸.

In this paper, we aim to investigate the evolution of water clusters and the structural origin of density fluctuations using simulations with SPC/E. In our earlier work, we propose a graph-based hierarchical clustering methods⁴⁷ to identify three-dimensional cage-like water clusters jointed by small-sized rings. Here, we further improve the hierarchical clustering methods to define water clusters covering most of hydrogen bond networks. Using classical molecular dynamics simulations, we study on how the water clusters changes at ambient temperatures. The distributions, morphologies, energetic and topological properties of water clusters are analyzed in detail. The characteristic structures of water clusters are further regarded as fundamental building blocks of hydrogen-bonded networks. By examining local number density and structural descriptor, the density fluctuations in the vicinity of water clusters are detected, proving that the water clusters can examine density fluctuations in the hydrogen bond networks, especially for small clusters.

Results

Molecular dynamics simulations and hierarchical clustering method. Molecular dynamics simulations are performed using periodic boundaries with 181,914 water molecules in a cubic box interacting through SPC/E water potential⁵³. All of the simulations were run at temperatures of 278 K, 298 K and 318 K and a pressure of 1.0 atm with a Nose–Hoover thermostat and barostat. The time step was 1.0 fs. The systems were equilibrated for 2.5 ns in an NPT ensemble before outputting.

The hierarchical clustering methods⁴⁷ respectively regard hydrogen bonds, rings, fragments as first-, second- and third-level sub-structures in the hydrogen bond network. The ring size is limited to 8 and the fragments analyzed contain 3, 4, and 5 rings. The graph community is a structure in which the nodes have a higher density of connections within groups than between them. Using Louvain algorithm^{54,55}, the best graph communities are defined as water clusters, symbolizing the regions where the rings or specific structures are concentrated.

The alpha shape⁵⁶ can describe the shape for water cluster by exploring the polytope wrapped by outermost water molecules. The volume V and surface area A of water clusters are defined as the volume and surface area of the alpha shape of clusters.

Cluster statistics, morphologies and energy. The definition of ring-based clusters can intuitively describe the regions of cage-like local structures. In the hydrogen bond networks, over 95% of water molecules and over 60% of global box are involved in the formation of water clusters. Hence, the ring-based clusters can be regarded as specific microscopic structures in liquid water. It is essential to study the properties of water clusters to elucidate how the cage-like structures influence the hydrogen bond networks, and even physical properties of liquid water.

At ambient conditions, the hydrogen bond networks can form percolating networks which contain infinite hbond-based clusters spanning periodic cubic simulation box at least in one direction⁵⁷. But in our definition, the ring-based networks do not have infinite clusters. The largest clusters only comprise more than 1% molecules in the box. Most of clusters are rather small, among which 6- and 9-molecule clusters are most favored in networks, as shown in Fig. 1a. Small clusters are mainly single-fragment clusters. Similar with hbond-based cluster distributions⁵⁷, large ring-based clusters distribution also obeys a power law suggesting fractal behaviors. The distributions of $S > 100$ form a shoulder at low temperature.

Cluster morphologies, closely related with the voids inside clusters, allow for density fluctuations in the hydrogen bond network⁴⁸. Some examples in Fig. 1b show various cluster morphologies. Small clusters tend to form spherical shapes and large clusters easily shape irregularly. As illustrated in Fig. 1c,d, the peaks of V and

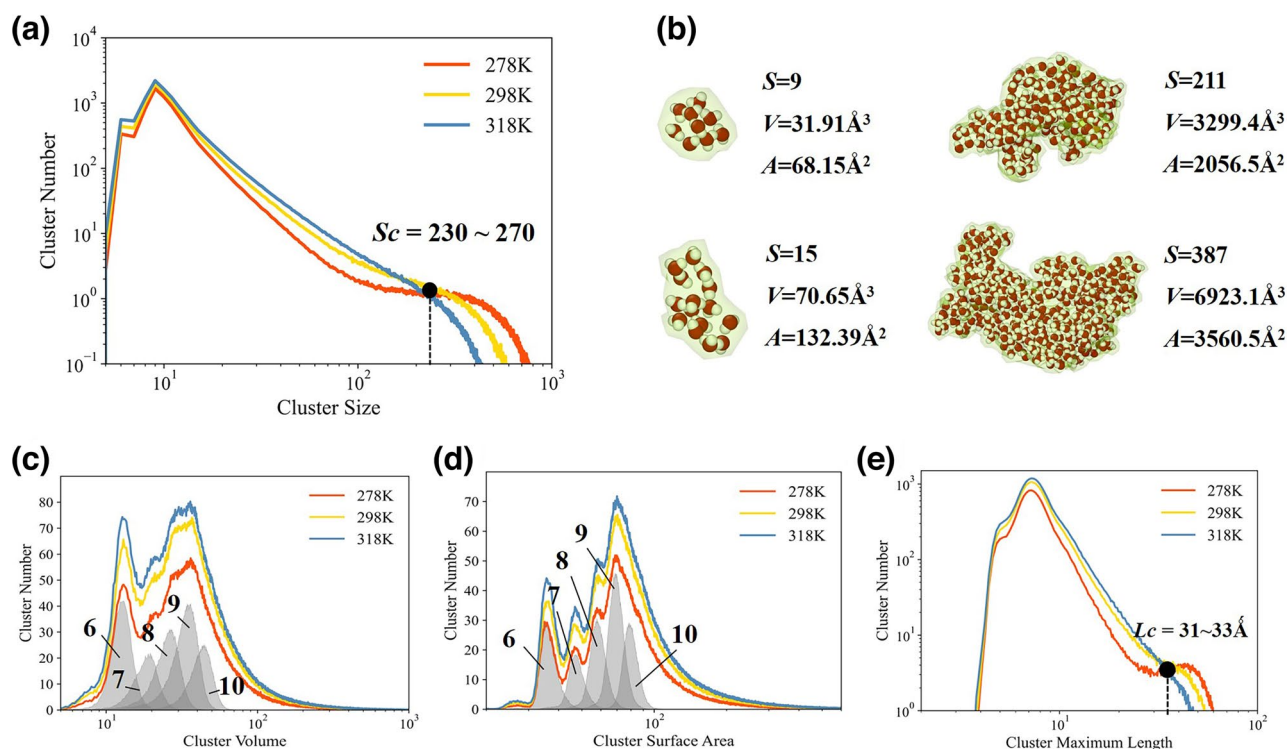


Figure 1. The cluster distributions and morphologies at various temperatures. **(a)** Water cluster distributions at 278 K, 298 K and 318 K. The black point denotes an isobestic point at $S_c = 230\text{--}270$. **(b)** Various water cluster morphologies at time step 0 at 278 K drawn by Ovito⁵⁸. S denotes cluster size. V and A denote the volume and surface area of the alpha shape of clusters. **(c)** The volume distributions of water cluster at 278 K, 298 K and 318 K. The gray zones represent volume distributions of clusters with 6–10 molecules. **(d)** The surface area distributions of water cluster at 278 K, 298 K and 318 K. The gray zones also represent surface area distributions of clusters with 6–10 molecules. **(e)** The maximum length distributions of clusters at 278 K, 298 K and 318 K. The black point denotes an isobestic point at $L_c = 31\text{--}33\text{Å}$.

A for small clusters imply optimal cluster shapes implying subtle balance between hydrogen bonds enthalpies and structural distortions. But large clusters have broader distributions of V and A without distinct peaks suggesting complex morphologies. Meanwhile, as shown in Fig. 1e, maximum diameters of clusters L , the longest distance between two molecules in a cluster, can indicate that structural heterogeneity occurs on a length scale of more than 3.0 nm.

Cluster energy plays an important role on the relationship between cluster structures and energy fluctuation in previous researches. The clusters have local energy minima based on quantum-chemical calculations and molecular dynamic simulations^{59,60}. The energy of water clusters E_S is defined as the sum of interaction energy between molecules in the clusters. Under our definition of ring-based clusters in bulk water, the cluster energy E_S is almost directly proportional to cluster size S in form of $E_S = a \times S + b$. The equation $E_S/S = a + b/S$ indicates that larger clusters may averagely have lower molecular energy E_S/S , as shown in Fig. 2a. At the same time, cluster energy is mainly determined by hydrogen bonds number and cluster topologies related with the *quality* of hydrogen bonds, as illustrated in Fig. 2b. The formation of low-energy structures suggests local favored topologies, explaining that energy fluctuations in the hydrogen bond networks may be greatly connected with structural arrangements and dynamics.

Temperature can greatly change the properties of ring-based clusters. Cluster distributions have an isobestic point $S_c = 230\text{--}270$ dividing cluster size with opposite temperature responses. With heating up, the cluster numbers decrease when $S > S_c$ but increase when $S < S_c$, suggesting that large clusters are destroyed into small ones and as a consequence average number of water clusters obviously increases. The isobestic points are determined by box sizes shown in part IV of supplemental material. Then, it also indicates that the agglomeration of cage-like structures also tends to break up in the ring-based networks. Low temperature can cause the formation of lower-energy structures. In terms of $E_S = a \times S + b$, at 278 K, 298 K and 318 K, the parameter a is equal to -12.92 eV, -12.37 eV and -11.81 eV, respectively, and b is equal to 65.49 eV, 60.60 eV and 54.12 eV.

Scaling behaviors of cluster distributions and morphologies. Fractals behaviors are the intrinsic nature of microscopic particles and networks to describe self-similarity. The infinite clusters at the percolating threshold of hydrogen bond networks belong to fractal object showing scaling behaviors of distributions. The morphologies of the cavities of Lennard–Jones fluid^{61,62} and the voids of hydrogen bond networks in liquid water^{48–52} also show obvious scaling behaviors suggesting underlying fractal structures.

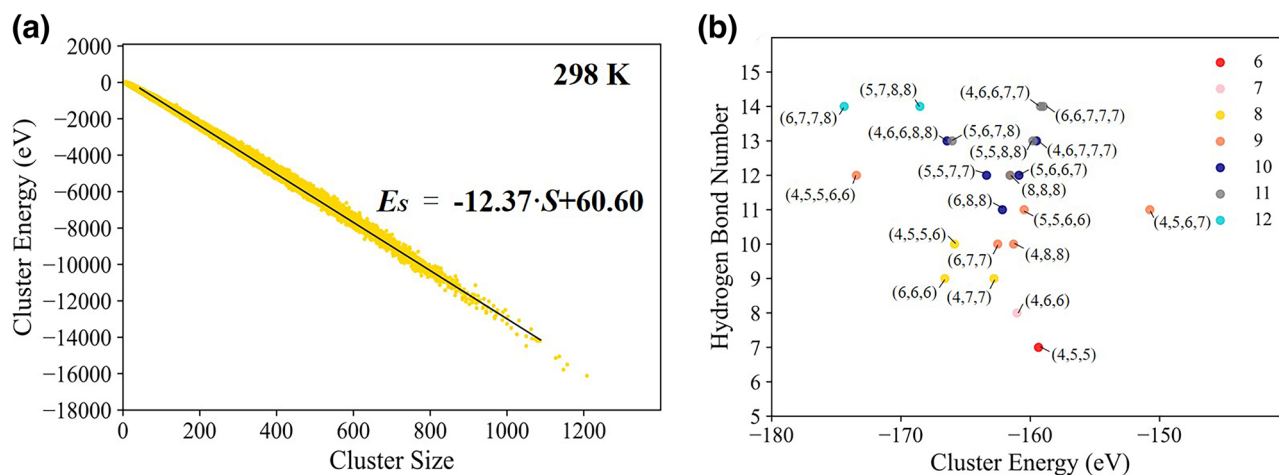


Figure 2. Scatter plots of cluster energy and cluster size and the relationship between cluster energy and cluster structures. **(a)** Scatter plots of cluster energy and cluster size at 298 K. The black line is the fit slopes by $E_S = a \times S + b$, where E_S denotes cluster energy and S is cluster size. **(b)** The scatter of average cluster energy with certain cluster size and cluster structures. The colors represent cluster size. The labels of points represent ring components in a cluster.

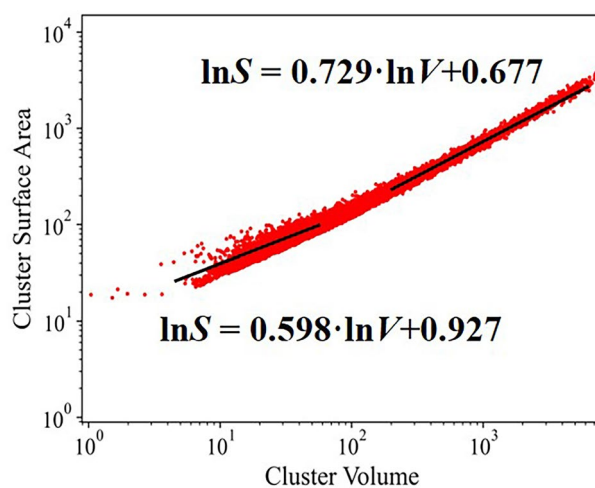


Figure 3. The log–log scatters of cluster volume and surface area. The black line is the fit slopes by $\ln S \sim a \ln V + b$.

The percolating threshold demonstrates the existence of percolating transition forming infinite hydrogen-bonded clusters, which closely relates with phase transition, especially liquid–liquid transition in supercooled water^{52,57}. At ambient conditions, the ring-based networks do not reach percolating threshold forming infinite cage-like clusters. The ice can form infinite ring-based clusters because the networks are tessellated by tetrahedral structures. It implies that percolating transition of ring-based networks may cause the formation of infinite cluster at lower temperature than hydrogen bond networks. The infinite clusters can represent specific chain- or sheet-like structure of spanning hydrogen-bonded network at mesoscopic scale. Formation of giant cluster structures may affect the properties of liquid water which are sensitive to mesoscopic structures⁵⁷. The ring-based clusters also showing scaling behaviors of distributions and large cluster distributions also obeys a approximate power law $n_S \sim S^{-\tau}$, where n_S denotes cluster number and S denotes molecular number in a cluster. The fractal dimensions τ are roughly equal to 2.44, similar to exponents $\tau = 2.20$ in the case of hydrogen bond networks⁵⁷.

The complex morphologies of water clusters also show underlying fractal structures. Both cluster volume V and surface area A scale well with cluster radius $r = L/2$ illustrated in Fig. 3, giving $d_V = 2.89$ and $d_S = 2.01$. It indicates that the cluster shapes are self-similar. The surface-to-volume scaling of clusters also show scaling behaviors in form of a power law $\ln A \sim a \ln V + b$ ⁶³. Similar with voids in hydrogen bond networks, small and large clusters show different scaling characteristics. For small clusters with $V < 100 \text{ \AA}^3$, $a = 0.598$, $b = 0.927$; for large clusters with $V > 100 \text{ \AA}^3$, $a = 0.729$, $b = 0.677$. The small clusters are non-fractal three-dimensional objects

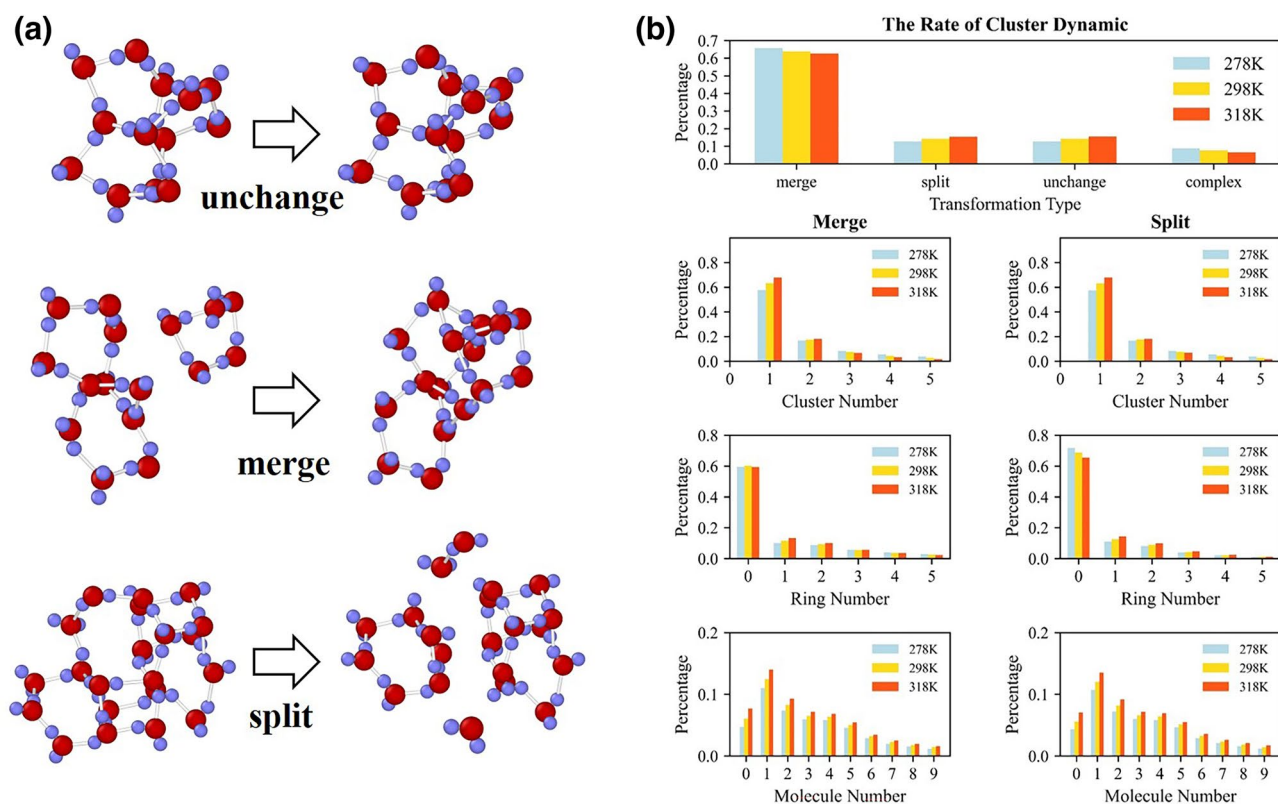


Figure 4. Example of cluster transformation and the distribution of cluster transformation at various temperatures. (a) Some example of cluster transformation involving unchanged, merged and split at 298 K drawn by Ovito⁵⁸. (b) The rate of cluster transformation and the products or reactant distributions of cluster splitting and merging at 278 K, 298 K and 318 K.

because the parameter a is near $2/3$. Then, large clusters slightly display surface fractals because the parameter a lie between $2/3-1$. The shape factor b depends on the shape, e.g. for spheres $b = 1.57$, for cubic cavities $b = 1.79$, for tetrahedra $b = 1.97$.

The dynamics and transformations of water clusters. The thermal fluctuation cause frequent arrangement of hydrogen bond network and then the recombination of water clusters. Accordingly, the cluster lifetime is very transient and can averagely last for ~ 10 fs, compared with the hydrogen bond lifetime of approximately 11.79 ps, 7.99 ps and 5.84 ps at 278 K, 298 K and 318 K. Without any assumptions of lifetime⁶⁴, average lifetime of clusters with $S < 10$ is approximately $4.0\Delta t$, where Δt denotes the timestep $\Delta t = 10.0$ fs. The average lifetime of clusters with $S > 50$ is only $0.39\Delta t$. Temperature also can shorten the lifetime of the clusters with $S < 10$, $4.0\Delta t$ at 278 K, $3.7\Delta t$ at 298 K and $3.4\Delta t$ at 318 K.

Cluster transformations have four patterns: changeless, merge, split and complex fusion and fission, as illustrated in Fig. 4a. Assume that a cluster, a ring and a molecule that belongs to the debris of rings or clusters at timestep t are defined as C_t , R_t and m_t .

1. Changeless: the cluster remains changeless in neighboring configurations.

$$1C_t \Rightarrow 1C_{t+1} \quad (1)$$

2. Merge: A cluster is constituted by at least one cluster or one ring at previous configurations.

$$i_c C_t + i_R R_t + i_m m_t \Rightarrow 1C_{t+1} \quad (i_c \geq 1 \text{ or } i_R \geq 1) \quad (2)$$

3. Split: A cluster is transformed into at least one cluster or one ring in the next configurations.

$$1C_t \Rightarrow i_c C_{t+1} + i_R R_{t+1} + i_m m_{t+1} \quad (i_c \geq 1 \text{ or } i_R \geq 1) \quad (3)$$

4. Complex fusion and fission: None of the transformations belong to the above types.

$$i_c C_t + i_R R_t + i_m m_t \Rightarrow i'_c C_{t+1} + i'_R R_{t+1} + i'_m m_{t+1} \quad (4)$$

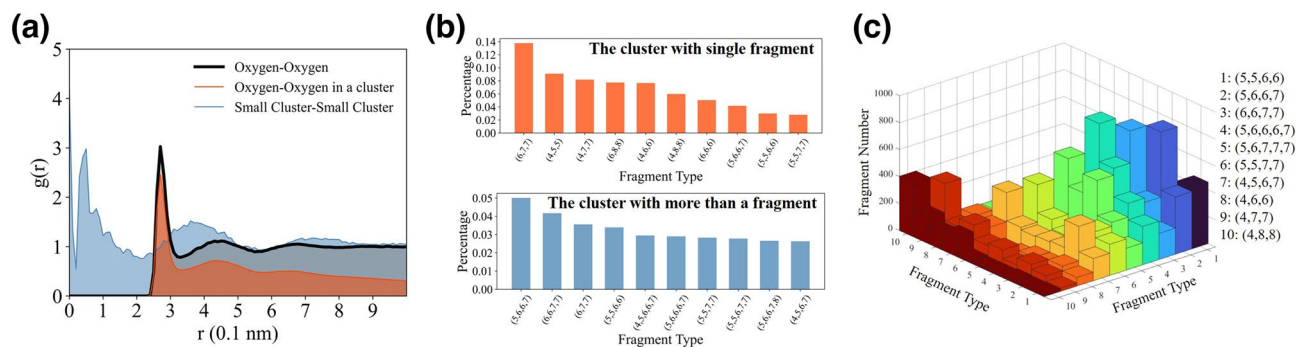


Figure 5. The oxygen–oxygen and cluster–cluster radial distance functions and topological properties of clusters. **(a)** The radial distance functions of oxygen–oxygen in global system, oxygen–oxygen in large clusters and small cluster–small cluster. The cluster RDF is calculated by the mass point of clusters with no more than 10 molecules. **(b)** The top 10 favored fragments in the clusters with a fragment and more than one fragment at 278 K. **(c)** The distributions of adjacent fragments in the clusters. Labels 1–10 represent most local favored fragments in the hydrogen bond network.

In the 10-fs intervals, over 60% of clusters remain unchanged as shown in Fig. 4b. The rates of merging and splitting are comparatively equivalent about 12.0–16.0%. The reactants of cluster merging are mostly 1 cluster and 1–5 molecules. Only 30% merging contain isolated rings. The products of split are also 1 cluster and several molecules. High temperature can decrease the rate of unchange clusters and increase the rate of merging and splitting from 12.7% at 278 K to 15.4% at 318 K. The cluster transformations are very frequent and prefer to merge or split the debris of other clusters, instead of ring structures, suggesting randomness of cluster dynamics. The cumulative effects of clusters cause the change of local structures and then complete rearrangements of hydrogen bond networks.

Structural heterogeneity of water clusters. In the hydrogen bonded networks, the water clusters represent ring-concentrated regions where different constitution of ring structures implies structural heterogeneity closely related with the arrangement of hydrogen bond networks.

The spatial distributions of clusters suggest microscopic local structures between clusters. The RDFs (radial distribution functions) of oxygen–oxygen atoms in the clusters have several peaks at 2.7 Å, 4.3 Å and 6.5 Å similar with that of all oxygen atoms and gradually approach zero, as illustrated in Fig. 5a. For small clusters, the RDF peaks are located at 0.5 Å, 3.6 Å and 7.2 Å in Fig. 4d, respectively representing adjacent water clusters. It indicates that structural inhomogeneity around ring-concentrated regions reveal weak structuring up to ~ 1 nm.

The cluster structures have several favored topologies. Several fragments are locally favored because of large quantities and long lifetimes in part II of supplemental material. Several favored patterns of fragments involve (6,7,7), (6,8,8), (4,7,7), (4,6,6), (6,6,7,7), (5,6,6,7), (5,6,6,7,8) and so on. The more 5- and 6-membered rings in the clusters gain in popularity with local structures. Most clusters have only one fragment, 59.1% of which contain 3 rings. Over 86% clusters with more than one fragments have 4 and 5 rings. The 3-ring fragments are easily isolated. The 6-molecule clusters belong to (4,5,5) fragments. Most 9-molecule clusters were (6,7,7) and (4,8,8) fragments. The 4- or 5-ring fragments tend to aggregate into large clusters including (5,6,6,7) and (6,6,7,7), as shown in Fig. 5b.

Figure 5c shows the probability of different clusters adjacent to each other. It reveals that the fragments prefer to adjoin same type of fragments, for instance 3-ring fragments tend to combine 3-ring fragments. The top 5 favored connection modes between cluster with 2–6 fragments at 298 K are also charted in part III of supplemental material. Strong tendency of certain pairs of fragments being adjacent in the clusters accounts for lower conformational entropy of connected structures in liquid water.

The surface rings among adjacent clusters have significant effects on the compatibility among adjacent fragments. The rings connecting two fragments in a cluster are defined as body rings, and the others are defined as surface rings. The 55% of 4-membered rings and 45% of 6-membered rings bridge adjacent fragments. The surface rings, where most hydrogen bond arrangements occur, dominate at high temperature. When the clusters are formed, the defects are transformed from their interior to surface area. The dynamics of surface area can cause the growth or decay of ring-concentrated regions³⁴.

Water clusters and density fluctuations in the hydrogen bond network. The water clusters can represent ring-concentrated cage-like structures. We explore how the water clusters are related with density fluctuations and then prove that water clusters can be regarded as characteristic structures to describe inhomogeneous patches of hydrogen bond networks on different length scales.

First, local density (ρ_{local}) near the center of clusters is a significant parameter to quantify density fluctuations around cluster structures. The local density (ρ_{local}) around the center of a cluster is defined as the number of oxygen atoms inside a probe sphere with a radius of 4.6 Å⁴⁸. The bulk number density at 298 K fluctuates at approximately 0.0334 molecule/Å³. The small and large water clusters respectively have $S \leq 10$ and $S \geq 100$. The average maximum cluster radius is approximately 3.85 Å and 19.40 Å.

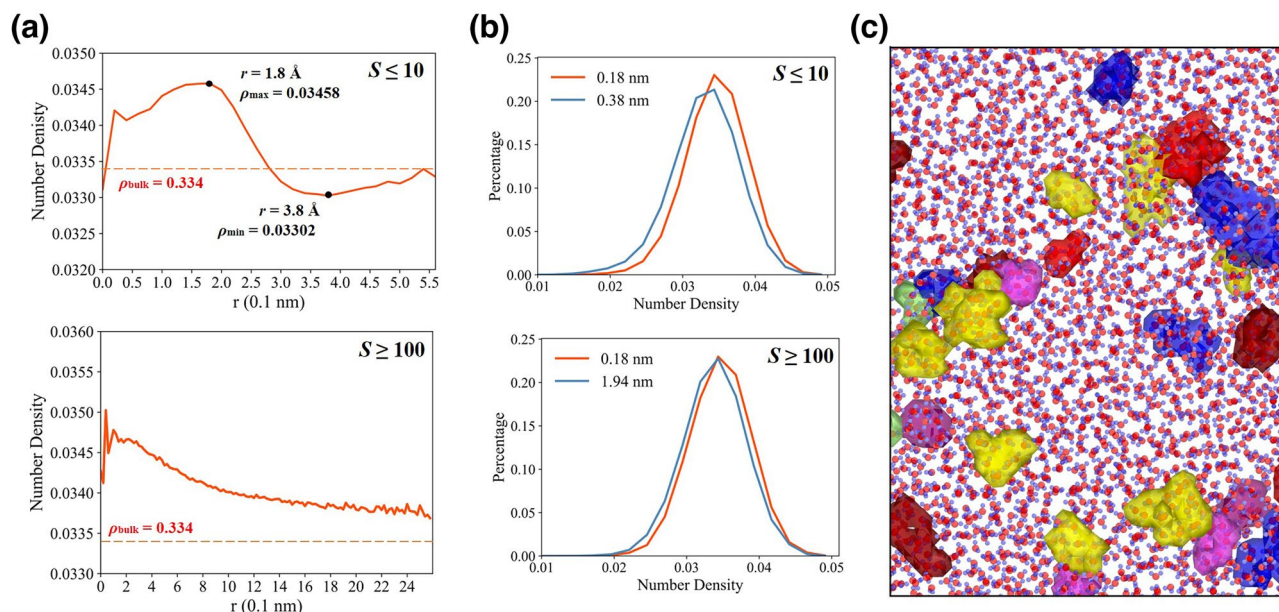


Figure 6. High- and low-density patches around the geometrical center of clusters at 278 K. **(a)** The number density around the geometrical center of clusters with $S \leq 10$ and $S \geq 100$, which denotes molecular number. The number density is calculated as the number of oxygen atoms inside a probe sphere located at certain oxygen atoms in the clusters with a radius of 4.6 Å. **(b)** The distributions of number density located at maximum density and minimum density. **(c)** Snapshots of several clusters selected in local region drawn by Ovito⁵⁸.

As shown in Fig. 6a, high-density patches exist in the interior of clusters. For small clusters, the distribution of densities reaches the peak at 1.8 Å and the minima near clusters surface. High-density patches is ~ 3% larger than bulk density. The regions outside the interior of large clusters are still larger than bulk density. In Fig. 6b, for small and large clusters, 64.8% and 64.7% local densities are larger than the bulk density at density peak compared with 51.3% and 57.4% at minima, which explain the discrepancy of density at different regions. Figure 6c shows an instantaneous snapshot of large cluster at 278 K, illustrating spatial inhomogeneities of different types of clusters.

The formations of structuring of water molecules have an implication on hydrophobic effect in liquid water⁶⁵. The hydrophobic solvation may create an empty cavity in the hydrogen bond network and form stronger inter-hydrogen bonds⁶⁶. To elucidate how water molecules behave inside the clusters, the local structure index (LSI)²⁵ has been introduced to characterize local structures in liquid water^{7,12,22,48}, measuring the extent to which the gap opens between water molecules in the first shell. LSI is defined as,

$$LSI = \frac{1}{n} \sum_{i=1}^n (\Delta(i) - \bar{\Delta})^2 \quad (5)$$

$\Delta(i) = r_{i+1} - r_i$ and $\bar{\Delta}$ is the average over all molecules in the first shell. The molecules are numbered in order of distance from the center molecule, $r_1 < r_2 < \dots < r_i < 3.7 \text{ \AA} < r_{i+1}$. Large LSI values represent tetrahedral water molecules, whereas small LSI values correspond to disordered first shells²¹.

In Fig. 7a, the tetrahedrality of water molecules decreases with increasing temperature. The isosbestic point at $LSI \approx 0.03 \text{ \AA}^2$ can divide two features of structures, representing the bimodality of LSI distributions having high-density- and low-density-like local structural environments²¹. The average tetrahedrality for water molecules in small clusters is slightly larger than that in large clusters. The voids inside small clusters, can be regarded as hydrophobic objects and hence the water molecules on small cluster surface have a preference for being more tetrahedral. The local structures of water molecules in the clusters show spatial nonuniformity and the tetrahedrality is negatively correlated with densities along the radius of clusters. The high-density patch in the clusters correspond to higher tetrahedrality, implying a more ordered first shell environment near non-polar hydrophobic solutes. It is also consistent with the results that long-range structural fluctuations of patches of four-coordinated molecules form from the liquid to supercooled water⁶⁷.

There is compositional similarity between the water clusters and voids. The water molecules in the clusters can encircle the voids wrapped in the closed rings. The small clusters may form spherical voids; moreover, the large clusters may involve some void clusters. Both structures show similar properties of distribution and morphology. Similar with water molecules around the voids⁴⁸, small clusters reveal similar correlations with density fluctuations. The deeper relationship between water clusters and voids needs to be explored further.

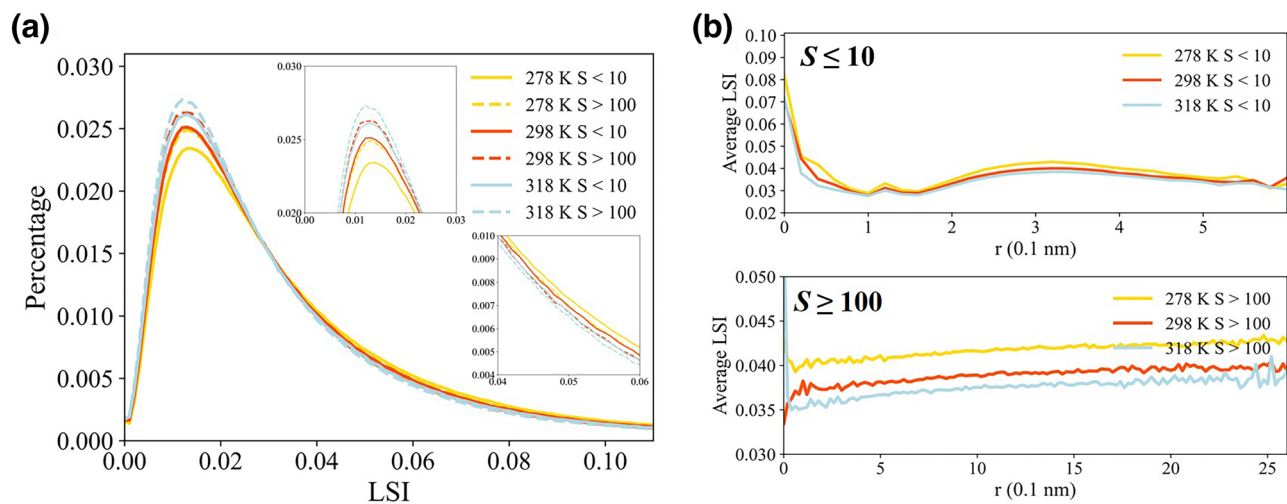


Figure 7. The distribution of LSI for waters in the vicinity of the clusters at various temperatures. **(a)** The distribution of the LSI of water molecules in the clusters at 278 K, 298 K and 318 K. The solid line denotes the clusters with $S \leq 10$, and the dashed line denotes the clusters with $S \geq 100$. **(b)** The average LSI around the geometrical center of clusters with $S \leq 10$ and $S \geq 100$.

Conclusions and discussions

The microscopic structures in liquid water remain a hot debate in recent researches. We search out water clusters in the hydrogen bond networks of liquid water using 100-thousand-particle molecular dynamic simulations by hierarchical clustering methods.

In the present paper, we show that there is underlying correlation between hydrogen bond networks and the properties of ring-based clusters within it. We demonstrate that the water clusters can intuitively identify density and structural fluctuations in the hydrogen bond networks, which leads to microscopic interpretations of water anomalies. In particular, we show that the regions inside the clusters correspond to high-density patches with lower tetrahedrality, especially for small clusters. The small clusters have a preference for having more tetrahedral molecules. It implies a more ordered first shell environment near non-polar hydrophobic objects which is equivalent with the voids in small clusters. And we also show that some clusters are favored by hydrogen bond networks depending on hydrogen bond numbers and cluster topologies. The compatibility of adjacent clusters depends on conformation and size of surface rings bridging them. The structural preference for water clusters also explain the arrangement and dynamics of hydrogen bond networks.

The ring-based clusters in percolating hydrogen bond networks show scaling behaviors of cluster distributions and morphologies forming underlying fractal structures and self-similarity. We show that small clusters are non-fractal three-dimensional objects and large clusters slightly display surface fractals. And large clusters distribution obeys power law with fractal dimension equal to 2.44. It also implies that percolating transition of ring-based networks may cause the formation of infinite cluster at lower temperature than hydrogen bond networks. Also, temperatures largely influence cluster properties and dynamics in the hydrogen bond network. We demonstrate that heating up can destroy large, low-energy clusters into small, high-energy one because of increasing defects of topological structures.

One concern is a lack of mathematical rigor in defining water structures. The simplicial complex in complex network may provide a tool to propose unified mathematical definition of water clusters. Other concern is that the quantitative relationship between the dynamics of water clusters and water properties need to be explored further.

As a fundamental building unit of hydrogen bond networks, water cluster can detect structural and density fluctuations on different length scales in liquid water and also provide a new insight to explain the nucleation process of ice, liquid-liquid coexistence in supercooled water and hydrogen bond structures in the vicinity of a hydrophobic solute. Furthermore, it will be interesting in the future to discretize continuous liquid water as an assembly of large clusters, hierarchically forming a bridge between isolated molecules and bulk matter.

Methods

Molecular dynamics simulations. The 100-thousand molecular dynamics simulations are conducted by LAMMPS software package to explore the properties of water clusters. The water model is modeled using SPC/E rigid water model⁵³, which is widely utilized to simulate macroscopic properties and microscopic structures of liquid water. The bond length and the bond angle of water are fixed with SHAKE algorithm. The cutoff of non-bonded interactions is 1.0 nm, and the long-range Coulomb interactions are computed using a particle-particle mesh solver (PPPM). The system contains 181,914 water molecules in a box with periodic boundary conditions. The initial box size is 28.0 nm along the x-direction, 10.0 nm along the y-direction and 20.0 nm along the z-direction. The equations of motion are integrated using velocity-Verlet algorithm⁶⁸ with a time step of 1.0 fs. A Nose-Hoover thermostat and barostat are used to control the temperature and pressure, respectively. The equilibration runs using an NPT ensemble are 2.5 ns. The temperatures of the system are kept constant at

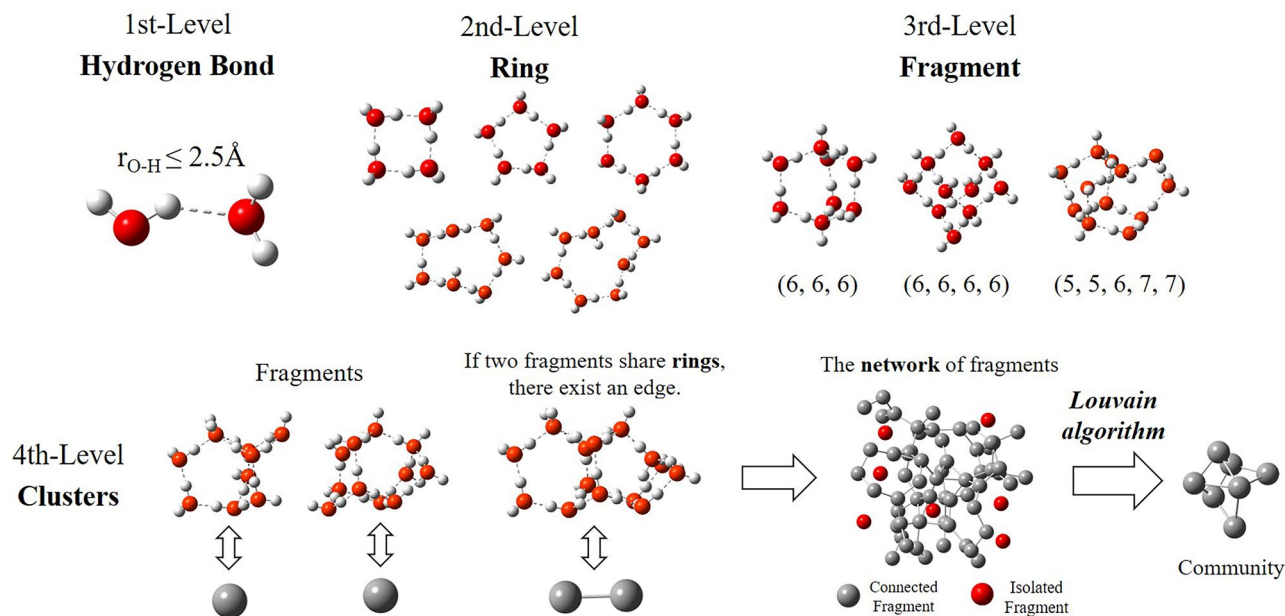


Figure 8. Schematics of the extended hierarchical clustering method. Hydrogen bonds, rings and fragments are considered 1st-, 2nd-, and 3rd-level structures, which are ball-and-stick models from a chemical perspective, and the red and white balls denote oxygen and hydrogen atoms, respectively. The full and dotted sticks denote O–H and hydrogen bonds, respectively. The clusters are illustrated from a topological perspective. The balls represent the graph community as water clusters. Note that the structures in the figure are only a selection among the considered structures by the clustering algorithm. The structures are drawn by GaussView 5.0.8⁶⁹.

278 K, 298 K and 318 K, and the pressure is 1.0 atm. The production runs output 500 configurations every 10.0 fs for structural analysis.

The density, diffusion coefficient, viscosity and structural properties of simulated water agree well with the results of experiments and other molecular dynamic simulations, and the simulation data are shown in Table 1 and Fig. 1 of the supplementary material.

Extended hierarchical cluster methods. To detect specific topological structures in hydrogen-bonded networks, hydrogen bonds, rings and fragments, clusters are constructed as the first-, second-, third- and fourth-level structures of the water network using hierarchical cluster methods⁴⁷, among which the clusters are an assembly of three-dimensional cage-like structures. To ensure that water clusters can largely cover hydrogen bond networks, we modify the definition of hierarchical cluster structures and propose extended hierarchical cluster method as illustrated in Fig. 8.

Hydrogen bonds, the first-level structures, are defined using a popular geometric standard in which two water molecules are regarded as hydrogen bonded when the distance between oxygen atoms of one molecule and the hydrogen atoms of another is less than 0.25 nm⁷⁰. The rings, the second-level structures, are cyclic paths along hydrogen bonds investigated by “shortest-path” (SP) criterion⁷¹. The maximum ring size is limited to 8. The fragments are three-dimensional cage-like building blocks using topological definition proposed by Matsumoto³⁴. The fragments combined by 3, 4, and 5 rings are taken into consideration. If two fragments share common rings, they have an edge, forming the network of fragments. The water clusters are identified as the best graph communities using Louvain algorithm^{54,55}.

Alpha shape of water clusters. The alpha shape is a formalization of the intuitive notion of shape for spatial point sets. An alpha shape is a concrete geometric model which is mathematically well-defined and unique for a given point set. The parameter alpha dictates the level of refinement, with a larger alpha resulting in coarser fits and a smaller alpha in finer fits. The volume and surface area of water clusters are based on the alpha shapes⁵⁶ of water molecules in the clusters. The alpha shapes of clusters only take oxygen atoms into consideration. The alpha radius is chosen as 3.5 Å. The formation of alpha shape is implemented in MATLAB⁷².

Data availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Received: 22 January 2022; Accepted: 25 April 2022

Published online: 16 May 2022

References

- Eisenberg, D. & Kauzmann, W. *The Structure and Properties of Water* (Oxford University Press on Demand, 2005).
- Bellissent-Funel, M. C. *et al.* Water determines the structure and dynamics of proteins. *Chem. Rev.* **116**, 7673–7697 (2016).
- Ludwig, R. Water: From clusters to the bulk. *Angew. Chem. Int. Ed.* **40**, 1808–1827 (2001).
- Gallo, P. *et al.* Water: a tale of two liquids. *Chem. Rev.* **116**, 7463–7500 (2016).
- Malenkov, G. Liquid water and ices: Understanding the structure and physical properties. *J. Phys. Condens. Matter* **21**, 283101 (2009).
- Russo, J. & Tanaka, H. Understanding water's anomalies with locally favoured structures. *Nat. Commun.* **5**, 4–6 (2014).
- Nilsson, A. & Pettersson, L. G. M. The structural origin of anomalous properties of liquid water. *Nat. Commun.* **6**, 1–11 (2015).
- Mishima, O. & Stanley, H. E. The relationship between liquid, supercooled and glassy water. *Nature* **396**, 329–335 (1998).
- Geiger, A. & Stanley, H. E. Low-density 'patches' in the hydrogen-bond network of liquid water: Evidence from molecular-dynamics computer simulations. *Phys. Rev. Lett.* **49**, 1749–1752 (1982).
- Harrington, S., Zhang, R., Poole, P. H., Sciortino, F. & Stanley, E. H. Liquid-liquid phase transition: Evidence from simulations. *Phys. Rev. Lett.* **78**, 2409–2412 (1997).
- Poole, P. H., Sciortino, F., Grande, T., Stanley, H. E. & Angell, C. A. Effect of hydrogen bonds on the thermodynamic behavior of liquid water. *Phys. Rev. Lett.* **73**, 1632–1635 (1994).
- Shi, R. & Tanaka, H. Microscopic structural descriptor of liquid water. *J. Chem. Phys.* **148**, 124503 (2018).
- Walrafen, G. E. Raman spectral studies of the effects of temperature on water and electrolyte solutions. *J. Chem. Phys.* **44**, 1546–1558 (1966).
- Huang, C. *et al.* The inhomogeneous structure of water at ambient conditions. *Proc. Nat. Acad. Sci. USA* **106**, 15214–15218 (2009).
- Nilsson, A. & Pettersson, L. G. M. Perspective on the structure of liquid water. *Chem. Phys.* **389**, 1–34 (2011).
- Pettersson, L. G. M., Henschman, R. H. & Nilsson, A. Water—The most anomalous liquid. *Chem. Rev.* **116**, 7459–7462 (2016).
- Wernet, Ph. The structure of the first coordination shell in liquid water. *Science* **304**, 995–999 (2004).
- Tokushima, T. *et al.* High resolution X-ray emission spectroscopy of liquid water: The observation of two structural motifs. *Chem. Phys. Lett.* **460**, 387–400 (2008).
- Montes de Oca, J. M., Rodriguez Fris, J. A., Accordini, S. R., Malaspina, D. C. & Appignanesi, G. A. Structure and dynamics of high- and low-density water molecules in the liquid and supercooled regimes. *Eur. Phys. J. E* **39**, 1–8 (2016).
- Martelli, F. Unravelling the contribution of local structures to the anomalies of water: The synergistic action of several factors. *J. Chem. Phys.* **150**, 094506 (2019).
- Santra, B., DiStasio, R. A., Martelli, F. & Car, R. Local structure analysis in ab initio liquid water. *Mol. Phys.* **113**, 2829–2841 (2015).
- Shi, R., Russo, J. & Tanaka, H. Common microscopic structural origin for water's thermodynamic and dynamic anomalies. *J. Chem. Phys.* **149**, 224502 (2018).
- Frank, H. S. & Wen, W. Y. Ion-solvent interaction. Structural aspects of ion-solvent interaction in aqueous solutions: A suggested picture of water structure. *Discuss. Faraday Soc.* **24**, 133–140 (1957).
- Geiger, A., Stillinger, F. H. & Rahman, A. Aspects of the percolation process for hydrogen-bond networks in water. *J. Chem. Phys.* **70**, 4185–4193 (1979).
- Shiratani, E. & Sasai, M. Molecular scale precursor of the liquid-liquid phase transition of water. *J. Chem. Phys.* **108**, 3264–3276 (1998).
- Cuthbertson, M. J. & Poole, P. H. Mixturelike behavior near a liquid-liquid phase transition in simulations of supercooled water. *Phys. Rev. Lett.* **106**, 1–4 (2011).
- Ruocco, G., Sampoli, M. & Vallauri, R. Analysis of the network topology in liquid water and hydrogen sulphide by computer simulation. *J. Chem. Phys.* **96**, 6167–6176 (1992).
- Liu, J., He, X. & Zhang, J. Z. H. Structure of liquid water—a dynamical mixture of tetrahedral and 'ring-and-chain' like structures. *Phys. Chem. Chem. Phys.* **19**, 11931–11936 (2017).
- Maheshwary, S., Patel, N., Sathyamurthy, N., Kulkarni, A. D. & Gadre, S. R. Structure and stability of water clusters (H₂O)_n, n = 8–20: An ab initio investigation. *J. Phys. Chem. A* **105**, 10525–10537 (2001).
- Ignatov, I. & Mosin, O. Structural mathematical models describing water clusters. *Nanotechnol. Res. Pract.* **3**, 141–158 (2014).
- Bakó, I. *et al.* Hydrogen bond network topology in liquid water and methanol: A graph theory approach. *Phys. Chem. Chem. Phys.* **15**, 15163 (2013).
- Ruocco, G., Sampoli, M., Torcini, A. & Vallauri, R. Molecular dynamics results for stretched water. *J. Chem. Phys.* **99**, 8095–8104 (1993).
- Belch, A. C. & Rice, S. A. The distribution of rings of hydrogen-bonded molecules in a model of liquid water. *J. Chem. Phys.* **86**, 5676–5682 (1987).
- Matsumoto, M., Baba, A. & Ohmine, I. Topological building blocks of hydrogen bond network in water. *J. Chem. Phys.* **127**, 134504 (2007).
- Radhakrishnan, T. P. & Herndon, W. C. Graph theoretical analysis of water clusters. *J. Phys. Chem.* **95**, 10609–10617 (1991).
- Marques Leite Dos Santos, V., Brady Moreira, F. G. & Longo, R. L. Topology of the hydrogen bond networks in liquid water at room and supercritical conditions: a small-world structure. *Chem. Phys. Lett.* **390**, 157–161 (2004).
- Choi, J., Lee, H., Choi, H. R. & Cho, M. Graph theory and ion and molecular aggregation in aqueous solutions. *Annu. Rev. Phys. Chem.* **69**, 125–149 (2018).
- Bakó, I., Megyes, T. & Bálint, S. Water–methanol mixtures: Topology of hydrogen bonded network. *Phys. Chem. Chem. Phys.* **10**, 5004–5011 (2008).
- Yang, L., Ji, H., Liu, X. & Lu, W. Ring-stacking water clusters: Morphology and stabilities. *ChemistryOpen* **8**, 210–218 (2019).
- Kazimirski, J. K. & Buch, V. Search for low energy structures of water clusters (H₂O)_n, n = 20–22, 48, 123, and 293. *J. Phys. Chem. A* **107**, 9762–9775 (2003).
- Sediki, A., Lebsir, F., Martiny, L., Dauchez, M. & Krallafa, A. Ab initio investigation of the topology and properties of three-dimensional clusters of water (H₂O)_n. *Food Chem.* **106**, 1476–1484 (2008).
- Kalinichev, A. G. & Churakov, S. V. Size and topology of molecular clusters in supercritical water: A molecular dynamics simulation. *Chem. Phys. Lett.* **302**, 411–417 (1999).
- Monserrat, B., Brandenburg, J. G., Engel, E. A. & Cheng, B. Liquid water contains the building blocks of diverse ice phases. *Nat. Commun.* **11**, 1–8 (2020).
- Lozynski, M. Liquid water: The helical perspective of structure. *Chem. Phys.* **455**, 1–6 (2015).
- Giovambattista, N., Buldyrev, S. V., Starr, F. W. & Stanley, H. E. Dynamic heterogeneities in liquid water. *AIP Conf. Proc.* **708**, 483–490 (2004).
- Chaplin, M. F. A proposal for the structuring of water. *Biophys. Chem.* **83**, 211–221 (2000).
- Gao, Y., Fang, H. & Ni, K. A hierarchical clustering method of hydrogen bond networks in liquid water undergoing shear flow. *Sci. Rep.* **11**, 1–14 (2021).
- Ansari, N., Dandekar, R., Caravati, S., Sosso, G. C. & Hassanali, A. High and low density patches in simulated liquid water. *J. Chem. Phys.* **149**, 204507 (2018).

49. Camisasca, G., Schlesinger, D., Zhovtobriukh, I., Pitsevich, G. & Pettersson, L. G. M. A proposal for the structure of high- and low-density fluctuations in liquid water. *J. Chem. Phys.* **151**, 034508 (2019).
50. Shih, J. P., Sheu, S. Y. & Mou, C. Y. A Voronoi polyhedra analysis of structures of liquid water. *J. Chem. Phys.* **100**, 2202–2212 (1994).
51. Ansari, N., Laio, A. & Hassanali, A. Spontaneously forming dendritic voids in liquid water can host small polymers. *J. Phys. Chem. Lett.* **10**, 5585–5591 (2019).
52. Ansari, N., Onat, B., Sosso, G. C. & Hassanali, A. Insights into the emerging networks of voids in simulated supercooled water. *J. Phys. Chem. B* **124**, 2180–2190 (2020).
53. Berendsen, H., Grigera, J. R. & Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **91**, 6269–6271 (1987).
54. Barbedo, A. S. & Falcao De Freitas, A. Finding community structure in very large networks. *Aaron. Bol. da Soc. Port. Cardiol.* **15**, 81–90 (2003).
55. Blondel, V. D., Guillaume, J. L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, 1–12 (2008).
56. Akkiraju, N. *et al.* Alpha shapes: definition and software. In *Proceedings of the 1st International Computational Geometry Software Workshop* 63–66 (1995).
57. Oleinikova, A. & Brovchenko, I. Percolating networks and liquid-liquid transitions in supercooled water. *J. Phys. Condens. Matter* **18**, S2247 (2006).
58. Stukowski, A. Visualization and analysis of atomistic simulation data with OVITO—The Open Visualization Tool. *Model. Simul. Mater. Sci. Eng.* **18**, 2154–2162 (2010).
59. Lenz, A. & Ojamäe, L. A theoretical study of water clusters: The relation between hydrogen-bond topology and interaction energy from quantum-chemical computations for clusters with up to 22 molecules. *Phys. Chem. Chem. Phys.* **7**, 1905–1911 (2005).
60. Lenz, A. & Ojamäe, L. A theoretical study of water equilibria: The cluster distribution versus temperature and pressure for (H₂O)_n, n=1–60, and ice. *J. Chem. Phys.* **131**, 134302 (2009).
61. Hirunsit, P. & Balbuena, P. B. Effects of confinement on water structure and dynamics: A molecular simulation study. *J. Phys. Chem. C* **111**, 1709–1715 (2007).
62. Stanley, H. E. Application of fractal concepts to polymer statistics and to anomalous transport in randomly porous media. *J. Stat. Phys.* **36**, 843–860 (1984).
63. Vishnyakov, A., Debenedetti, P. G. & Neimark, A. V. Statistical geometry of cavities in a metastable confined fluid. *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.* **62**, 538–544 (2000).
64. Voloshin, V. P. & Naberukhin, Y. I. Hydrogen bond lifetime distributions in computer-simulated water. *J. Struct. Chem.* **50**, 78–89 (2009).
65. Wu, X., Lu, W., Streacker, L. M., Ashbaugh, H. S. & Ben-Amotz, D. Temperature-dependent hydrophobic crossover length scale and water tetrahedral order. *J. Phys. Chem. Lett.* **9**, 1012–1017 (2018).
66. Grdadolnik, J., Merzel, F. & Avbelj, F. Origin of hydrophobicity and enhanced water hydrogen bond strength near purely hydrophobic solutes. *Proc. Natl. Acad. Sci. USA* **114**, 322–327 (2017).
67. Moore, E. B. & Molinero, V. Growing correlation length in supercooled water. *J. Chem. Phys.* **130**, 244505 (2009).
68. Martyts, N. S. & Mountain, R. D. Velocity Verlet algorithm for dissipative-particle-dynamics-based models of suspensions. *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.* **59**, 3733–3736 (1999).
69. GaussView, Version 5, Roy Dennington, Todd Keith and John Millam (Semichem Inc., 2009).
70. Kumar, R., Schmidt, J. R. & Skinner, J. L. Hydrogen bonding definitions and dynamics in liquid water. *J. Chem. Phys.* **126**, 1–13 (2007).
71. Franzblau, D. S. Computation of ring statistics for network models of solids. *Phys. Rev. B* **44**, 4925–4930 (1991).
72. MATLAB and statistics toolbox Release 2020b (The Mathworks, Inc., 2020).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. U2040214) and the 111 Project (No. B18031).

Author contributions

H.F., Y.G., K.N. and Y. F designed the research. Y.G. performed simulations and analyzed the data. H.F., Y.G., K.N. and Y.F. discussed the manuscript. Y.G. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-11947-6>.

Correspondence and requests for materials should be addressed to H.F.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022