SUPPLEMENTARY INFORMATION
CRUSTWATER MODEL OF SOLVATION

AJEET KUMAR YADAV, PRADIPTA BANDYOPADHYAY, EVANGELOS A. COUTSIAS, KEN A. DILL

The bulk pure water reference state: the Cage Water Model

We have used an analytical 3D version of a model, similar to the Mercedes-Benz model of water, developed by Urbic et. al. [1]. This model has been described in several studies [2, 3]. Here we are introducing the model briefly. In this model, the liquid state of water is taken as the perturbation from a hexagonal lattice of the ice. Each grid point of the underlying lattice is assigned to a single water molecule, i.e. no two waters can occupy the same grid point. Each water molecule is modelled as a sphere having four hydrogen bonding (H-bonding) arms which are tetrahedrally oriented. In this model, we consider the interaction of a water molecule (called as test water) to its clockwise neighboring molecule. The clockwise-like direction is chosen only for the bookkeeping purpose. This model considers that each water molecule can be in any of the three possible states with its clockwise-like neighbor: (a) Hydrogen-bonded (HB) state, (b) Lennard-Jones (LJ) state, and (c) Open (O) or non-interacting state as represented in the following Figure S1. Depending on the relative orientation and distance between the test water and its clockwise-like positioned water, the interaction between them can be characterized as one of the three states which are mentioned above. If a clockwise-like neighboring molecule points one of the H-bonding arms towards the test water and an H-bonding arm of test water forms an angle with the line joining the centers of these two waters within the range $\pi/3$, then the state is characterized as HB state. If the distance criterion between the test and clockwise-like neighbor is satisfied but not the orientational one then the two waters form LJ contact but no hydrogen bond. If neither distance nor orientational criteria is satisfied then the two waters are considered to be in open state i.e. two waters do not have interaction between themselves.

First, we have computed the isothermal-isobaric partition functions of each state, Δj, as a function of temperature, pressure and interaction energies, which is described below.

$$\Delta_j = c(T) \int\int\int\int\int dxdydzd\phi d\psi \int_0^{\frac{\pi}{3}} exp\left(\frac{-(u_j + pv_j/2)}{kT}\right) \sin\theta d\theta \qquad \text{(S1)}$$

where $c(T)$ is the kinetic contribution to the partition function, u_j is the interaction energy between two waters in the j^{th} state, and p, T, and k are pressure, temperature and Boltzmann constant, respectively. The triple integration $\int\int\int dxdydz$ represents the translational freedom over which the two waters can be in the j^{th}

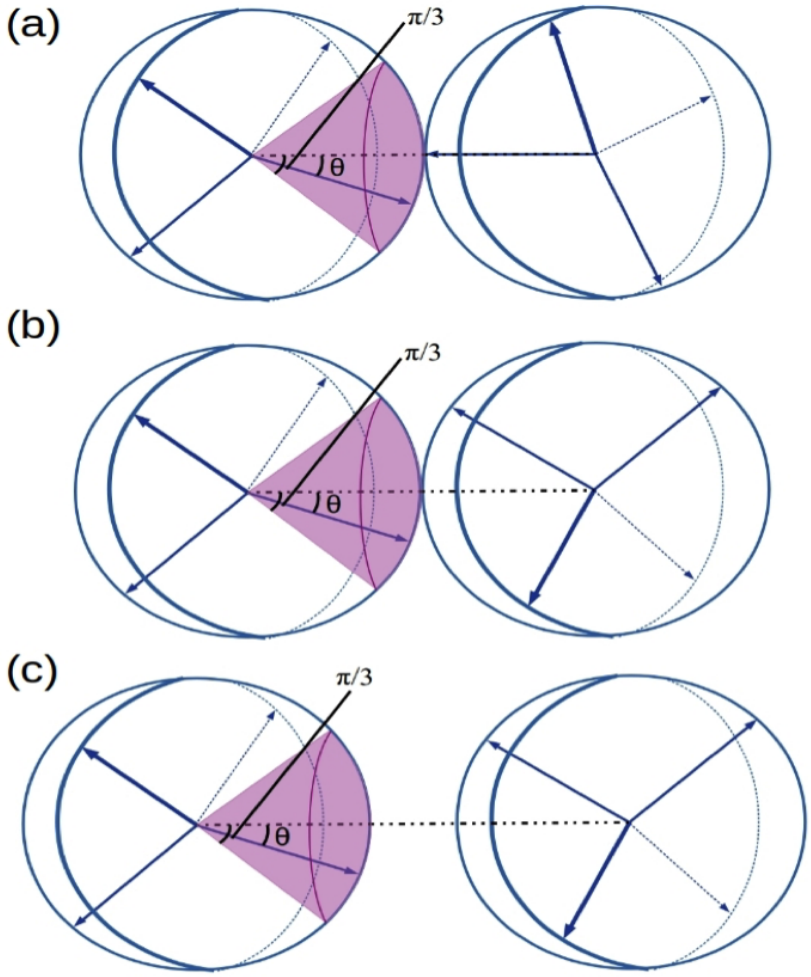

FIGURE S1. Pairwise water-water interactions. (a). Hydrogen bonded state (b). Lennard-Jones (LJ) state (c). Open state.

state and is defined as v_{eff}^{j} . The double integration $\int\int d\phi d\psi$ represents the orientational freedom over which test water can still be in the j^{th} state and is equal to $4\pi^2$ [2]. v_j is the volume per water molecule in the j^{th} state. θ is the angle between one HB arm of the test water molecule and the line joining the centers of the test and clockwise-like neighbor water molecule and the integration is upto $\pi/3$ because it corresponds to one fourth of the total solid angle.

From the partition function of the different states we can get the partition function of one hexagon as follows:

$$Q_1 = (\Delta_{HB} + \Delta_{LJ} + \Delta_O)^6 \tag{S2}$$

If all water molecules of the hexagon form hydrogen bonds between themselves then this state will have a higher co-operativity than the pairwise hydrogen bonds. After including the higher co-operativity, we get the partition function of the hexagon as :

$$Q_1 = (\Delta_{HB} + \Delta_{LJ} + \Delta_O)^6 - \Delta_{HB}^6 + exp(-\beta\epsilon_c)\Delta_{cage}^6 \tag{S3}$$

The last term of eq S3, $(exp(-\beta\epsilon_c)\Delta_{cage}^6)$, is used for higher co-operativity which replaces the corresponding weak pairwise term (Δ_{HB}^6), where ϵ_c is the co-operativity energy and $exp(-\beta\epsilon_c)$ is the corresponding Boltzmann factor and Δ_{cage} is termed as the isothermal-isobaric partition function of the cooperative state termed as cage. Δ_{cage} is same as Δ_{HB} with only difference that it uses a different volume per water molecule such that solid state has higher density. The interaction between the hexagons is treated by the mean-field attractive energy $-Na/v$ where N is the number of particles, a is van der Waals dispersion coefficient and v is the average molar volume. This is implemented in the same way as described in the work of Truskett et. al. [4, 5].

The fractional populations (f_j) and the ensemble average energies ($\langle u_j \rangle$) of different states are calculated as described previously [2, 3]. After summing the individual state's average energies, we get the average energy of a water molecule in the bulk water as :

$$\langle \epsilon \rangle_b = 2[\, \langle u_{HB} \rangle f_{HB} + \langle u_{cage} \rangle f_{cage} + \langle u_{LJ} \rangle f_{LJ} \,] \tag{S4}$$

Inserting a solute causes a geometric restriction of waters

We have treated the insertion of a non-polar molecule into water in the same way as has been done by previous analytical models consisting of variants of MB water model as their pure water models [3, 6, 7]. The insertion of non-polar solute perturbs two aspects of the water molecules in the first solvation shell of the non-polar solute. The first is the geometric aspect and the second is the energetic aspect. Due to its geometric aspect, the insertion of the solute prevents water molecules of the first solvation shell from forming maximum number (four) of the hydrogen bonds with other water molecules. The geometric restrictions on the maximum number of hydrogen bonds depend on the size of the solute. For smaller solutes, the water molecules in the first solvation shell can form 3 or 4 hydrogen bonds depending on its orientation, while for larger solutes, this maximum reduces to only 1,2, or 3 hydrogen bonds. An angle ϕ is defined as the angle between the line joining the centers of a water molecule in the first solvation shell and the solute and one of the HB arms of the water molecule in the first shell. The value of this angle at which an HB arm points along the tangent to the solute is defined as the *critical angle* while the water is in LJ minimum with the solute. Thus, the HB arms, for which ϕ is less than the critical value, will lose its ability to form hydrogen bond with other water molecules. This is shown in the Figure S2.

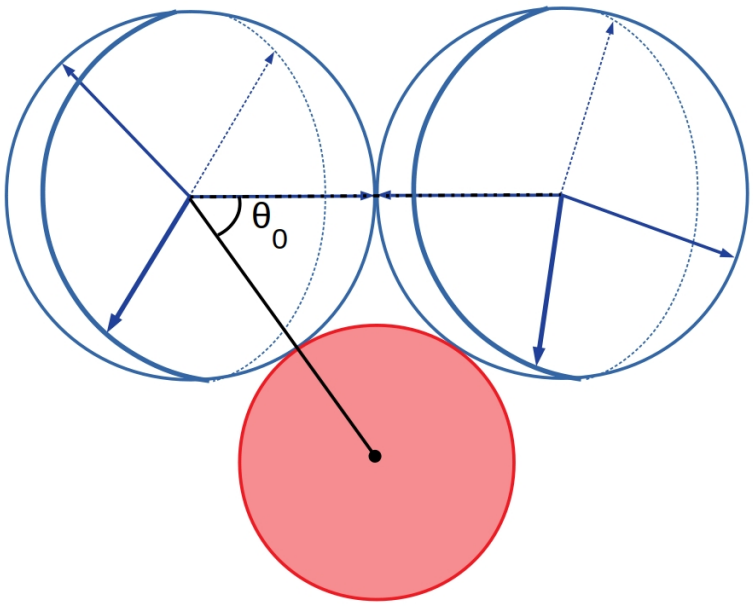

FIGURE S2. Definition of the critical angle. The shaded sphere represents the solute and other spheres represent water molecules

The expression for the critical angle is given by

$$\theta_0 = \arccos\left(\frac{r_{HB}}{(\sigma_{LJ} + 2r)2^{1/6}}\right) \tag{S5}$$

where r_{HB},σ_{LJ} and r are HB distance, LJ distance and solute radius respectively. Using this definition of the critical angle, we calculated the average number of the hydrogen bonds formed by a water molecule of the first solvation shell as explained in the following section.

EXACT GEOMETRIC FORMULAS FOR HOW MUCH A SOLUTE OCCLUDES FIRST-SHELL WATER-WATER HYDROGEN BONDING.

Basic definitions.

(1) Water molecule: radius 1 (scale to water radius).
(2) Water molecule: H-bonding arm mutual angle $\alpha = cos^{-1}(-1/3) = 109.4712^o$.
(3) Spherical solute (no H-bonding) radius ρ; however we represent the solute by the critical angle θ_0 of apperture of the spherical cap about an arm that it will occlude if it is tangent anywhere on that cap. To stress its role in occluding a particular arm from h-bond formation, we shall refer to θ_0 in the sequel as the *"occultation angle"*.
(4) Occultation angle (convex locally spherical boundary) $0 \leq \theta_0 \leq 90^o$ (S5).

The occultation cup. Water molecules are spheres on which the h-Bonding arms $\vec{a}_i, i = 1,2,3,4$ are located at coordinates (scaled $\times 3^{-1/2}$): $(1,1,1)$, $(1,-1,-1)$, $(-1,1,-1)$, $(-1,-1,1)$. We want to compute the total volumes $V_i, i = 0,1,2,3$ in $\mathcal{SO}(3)$ of states where 0,1,2, or 3 arms are occluded by solute sphere, assumed featureless. The computation is simplest in the body frame of the water molecule, with the solute sphere rolling on the water sphere. This reduces the integral of the relative rotation over the common axis to a constant (2π) and allows a simple geometrical interpretation of multiple arm occultation as the overlap of spherical caps, centered at each of the arm loci and having apperture equal to the solvent occultation angle θ_0. The occultation angle of the solute sphere is a function of the ratio r of its radius to that of water. It is defined in Figure S2. If the solute sphere is tangent to the water sphere at any point whose radius vector makes an angle less than θ_0 to an arm, that arm is occluded and cannot participate in H-bond formation with another water moledule. For values of θ_0 less than $\alpha/2$ only single arm occultation is possible while for greater values double and triple overlaps may occur. 4 arm occultation is impossible near a convex object, even in the limit of infinite radius, i.e. a planar hydrophobic surface. However it is possible at a concave surface (not considered here). In the rotated geometry considered the arm $\mathbf{a}_1$ points along the positive z-axis ; $\mathbf{a}_2$ is on the $x-z$ plane, and the remaining arms arranged counterclockwise around $\mathbf{a}_1$. Specifically, with

$$s\alpha = \frac{2\sqrt{2}}{3} \ , \ c\alpha = -\left(\frac{1}{3}\right) \ :$$

$$\begin{aligned}
\vec{a}_1 &= (0,0,1) \ , \\
\vec{a}_2 &= (s\alpha, 0, c\alpha) \ , \\
\vec{a}_3 &= \left(-s\alpha/2, \sqrt{3}s\alpha/2, c\alpha\right) \ , \\
\vec{a}_4 &= \left(-s\alpha/2, -\sqrt{3}s\alpha/2, c\alpha\right) \ .
\end{aligned}$$

We want to consider the volume of such states where different numbers of arms are occluded for a water molecule in the first layer around spherical solute. These are simply the areas spanned by the cups about each arm, provided we identify and properly account for areas of regions away from all cups (0-arm occultation), areas covered by only one, only two or only three cups. This means that we need to consider the arrangement of the cup bounding circles on the water sphere for values of θ_0 in the whole range, and identify their mutual intersections.

We compute these volumes by integrating the measure

$$sin\theta d\theta d\phi d\psi$$

over appropriate limits. Due to symmetry, all ψ integrations will simply contribute a factor of 2π throughout. Integrating this measure over all possible orientations gives the "Total Volume"

$$V_0 = \int_{\phi=0}^{2\pi} d\phi \int_{\theta=0}^{\pi} d\theta \sin\theta \int_{\psi=0}^{2\pi} d\psi = 8\pi^2 \ . \tag{S6}$$

We proceed now to compute areas for various types of intersections. We see that there are three possibilities, depending on θ_0

(1) only 1-arm occultation: $0 \leq \theta_0 < \alpha/2 \approx 54.7356^o$

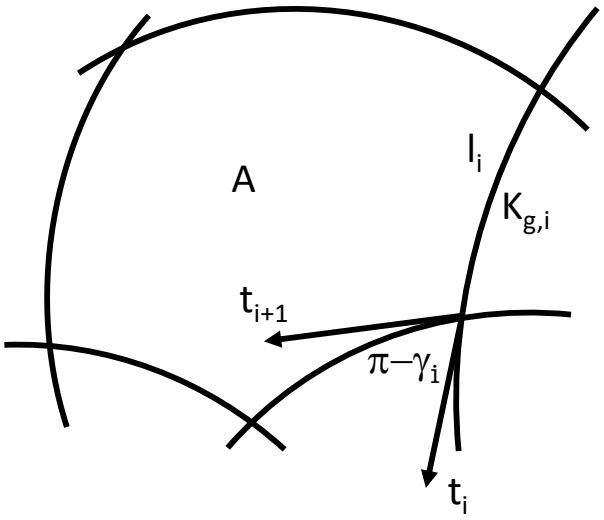

FIGURE S3. A circular polygon on the unit sphere

(2) 1- or 2-arm occultation: $\alpha/2 \leq \theta_0 \leq \pi - \alpha \approx 70.5288^o$
(3) 1-, 2- or 3-arm occultation: $\pi - \alpha \leq \theta_0 \leq \pi/2$

For reasons explained below ('Inclusion-Exclusion' argument and symmetry) we will only consider the following problems:

(1) Compute v_1, the total volume in $\mathcal{SO}(3)$ of states where the arm $\mathbf{a}_1$ is occluded, regardless of what other arms are doing.

$$v_1 = \int_{\phi=0}^{2\pi} d\phi \int_{\theta=0}^{\theta_0} d\theta \sin\theta \int_{\psi=0}^{2\pi} d\psi = 4\pi^2 \left(1 - c\theta_0\right) \ . \tag{S7}$$

(2) Compute v_{12}, the total volume in $\mathcal{SO}(3)$ of states where the arms $\mathbf{a}_1, \mathbf{a}_2$ are jointly occluded, regardless of what other arms are doing.
(3) Compute v_{123}, the total volume in $\mathcal{SO}(3)$ of states where the arms $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ are jointly occluded. Arm $\mathbf{a}_4$ will necessarily be free in this case.

The 2− and 3−arm occultation integrals are easiest done using the Gauss-Bonet formula, giving the area of a spherical region bounded by circular arcs Figure S3

$$A = 2\pi - \sum_i \left(\pi - \gamma_i\right) - \sum_j k_{g,j} l_j \ . \tag{S8}$$

Here the first summation involves the exterior angles of intersection between successive arcs (that is, the γ_i are the interior angles), while the second summation

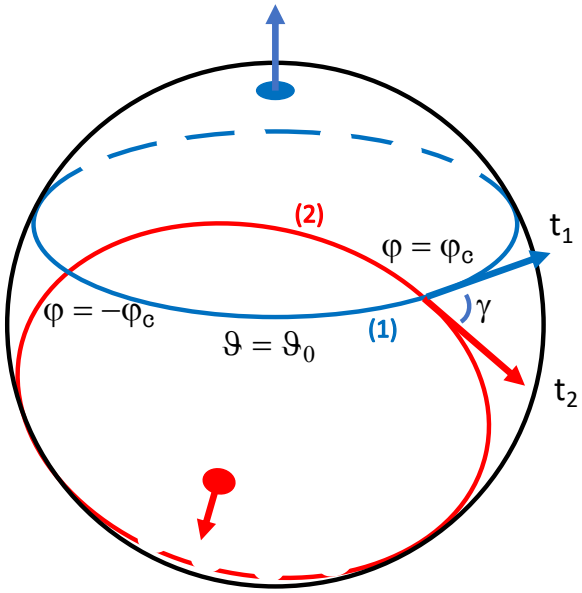

FIGURE S4. The meniscus u_{12} of intersection of the occultation caps associated with arms $\vec{a}_1$ and $\vec{a}_2$.

involves the length l_j of the $j-$th arc times its geodesic curvature, $k_{g,j}$. The geodesic curvature of a circle of radius r on a sphere of radius R is given by

$$k_g = \frac{\sqrt{R^2 - r^2}}{rR} \ .$$

From spherical trigonometry we know that the equation of a circle of apperture θ_0 on the unit sphere, with center on the radius with spherical coordinates (θ_1, ϕ_1) is given by

$$c(\phi - \phi_1) = \frac{c\theta_0 - c\theta c\theta_1}{s\theta s\theta_1} \ . \tag{S9}$$

The center of such a circle is located at distance $s\theta_0$ from the center of the unit sphere while its radius $r = c\theta_0$. Therefore the geodesic curvature is given by

$$k_g = \cot\theta_0 \ .$$

To compute the area v_{12} of intersection of the occultation cup of arm $\mathbf{a}_2$, a circle of apperture θ_0 centered at $(\theta_1, \phi_1) = (\alpha, 0)$ with the cup of arm $\mathbf{a}_1$, a circle of apperture θ_0 centered at the north pole Figure S4 we need their intersection angles and arc lengths. In equation (S9) we simply set $\theta = \theta_0$ to find ϕ_c, the ϕ coordinate(s), as

$$c\phi_c = cot\theta_0 \frac{1 - c\alpha}{s\alpha} = \sqrt{2} cot\theta_0 \ . \tag{S10}$$

Therefore the arclength of each circular arc bounding the meniscus of double intersection (nonzero only if $\theta_0 \geq \alpha_2$), is given by

$$l_j = l := 2r\phi_c = 2s\theta_0 \cos^{-1}\left(\sqrt{2}\cot\theta_0\right) \ , \ j = 1, 2 \ . \tag{S11}$$

The exterior angle of intersection of these two circles is found by considering the dot product of their tangent vectors at one of the points of intersection, say at (θ_0, ϕ_c). We have

$$\vec{t}_1 = (-s\phi_c, c\phi_c, 0) = \frac{1}{s\theta_0}\left(-\sqrt{1-3c^2\theta_0}, \sqrt{2}c\theta_0, 0\right)$$

while the tangent vector on second circle is found after algebra to be

$$\begin{aligned} \vec{t}_2 &= \frac{\vec{a}_2 \times \vec{d}}{||\vec{a}_2 \times \vec{d}||} \\ &= \left(-c\alpha\sqrt{1-3c^2\theta_0}, c\theta_0(\sqrt{2}c\alpha - s\alpha), s\alpha\sqrt{1-3c^2\theta_0}\right) \end{aligned}$$

so that

$$\vec{t}_1 \cdot \vec{t}_2 = c\gamma = \frac{1/3 + c^2\theta_0}{s^2\theta_0} \ .$$

Collecting results, we find for $v_{12} = 2\pi A$:

$$v_{12}/(2\pi) = 2\cos^{-1}\left\{\frac{1/3 + c^2\theta_0}{s^2\theta_0}\right\} - 4c\theta_0 \cos^{-1}\left(\sqrt{2}\cot\theta_0\right) \ .$$

We now compute the area of a triple intersection, v_{123}, which can occur in the range $\pi - \alpha \leq \theta_0 \leq \pi/2$. Since cup 3 is cup 2 rotated about the z-axis by $2\pi/3$ degrees, that displaces the second intersection point by $2\pi/3$ Figure S5 giving a common arc of

$$\Delta\phi = 2\phi_c - \frac{2\pi}{3}$$

with arc length $r\Delta\phi$. Then, since now we have three arcs and three intersections, while angles of intersection and geodesic curvature are as before, we find

$$\begin{aligned} v_{123}/(2\pi) &= 3\gamma - 3lk_g \\ &= 3\cos^{-1}\left\{\frac{1/3 + c^2\theta_0}{s^2\theta_0}\right\} - \dots \\ &\quad -6\cos\theta_0\left(\cos^{-1}\left(\sqrt{2}\cot\theta_0\right) - \frac{\pi}{3}\right) \end{aligned}$$

The total volumes computed through inclusion-exclusion. We can now find the total volumes of 0,1,2 and 3 arm occultation as

(1) 3-arm occultations: Simply add together all possible combinations of 3 arms, since they represent disjoint states (can't have all four occluded):

$$V_3 = 4v_{123} \tag{S12}$$

(2) 2-arm occultations: if we add up all 6 alternatives, we will be adding in also pieces of triple intersections: e.g. the states where arms $\mathbf{a}_1, \mathbf{a}_2$ are occluded, whose volume is v_{12}, will not be pure 2-arm occultations but will also include the (mutually disjoint) volumes v_{123} and v_{124}. So, considering all combinations, we have

$$V_2 = (v_{12} - v_{123} - v_{124}) + \dots + (v_{34} - v_{134} - v_{234}) = 6v_{12} - 12v_{123} \tag{S13}$$

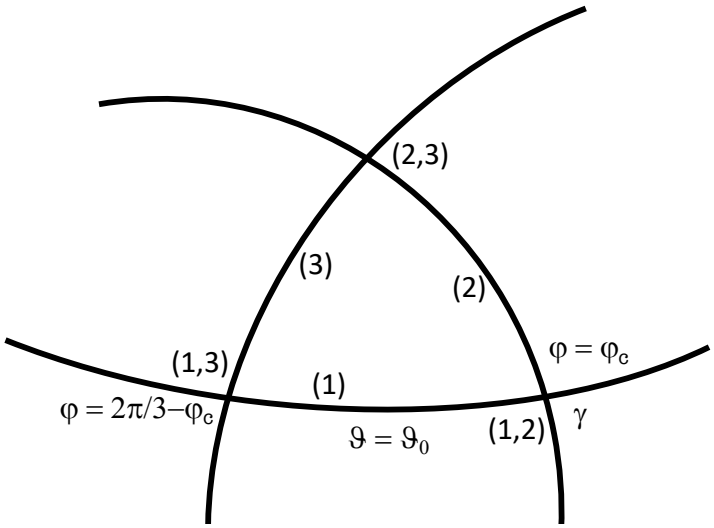

FIGURE S5. The spherical triangle u_{123} of intersection of the occultation caps associated with arms $\vec{a}_1$, $\vec{a}_2$ and $\vec{a}_3$.

(3) 1-arm occultations: if we add up all 4 alternatives, we will be adding in also pieces of double and triple intersections: e.g. the states where arm $\mathbf{a}_1$ is occluded include the states where arm pairs $(\mathbf{a}_1, \mathbf{a}_2)$, $(\mathbf{a}_1, \mathbf{a}_3)$ and $(\mathbf{a}_1, \mathbf{a}_4)$ are occluded and in there one can also find all triple intersections including $\mathbf{a}_1$. So, if we subtract double intersections (v_{12}, v_{13} and v_{14}) we will need to subtract triple intersections v_{123}, v_{124}, v_{134} (which will then need to be also reconciled so the proper amounts are subtracted or added back in), as follows:

$$\begin{aligned} V_1 &= v_1 - \{v_{12} - v_{123} - v_{124}\} - \{v_{13} - v_{123} - v_{134}\} - \dots \\ &\quad \dots - \{v_{14} - v_{124} - v_{134}\} - \{v_{123} + v_{124} + v_{134}\} \\ &= v_1 - \{v_{12} + v_{13} + v_{14}\} + \{v_{123} + v_{124} + v_{134}\} \end{aligned}$$

and adding up all arms

$$V_1 = 4v_1 - 12v_{12} + 12v_{123} \ . \tag{S14}$$

Computing the integrals. Collecting results from above, we arrive at the following sets of equations, applicable at the various ranges of the parameter θ_0. We have

(1) only 1-arm occultation: $0 \le \theta_0 < \alpha/2 \approx 54.7356^o$

$$v_1/(2\pi) = 2\pi(1 - c\theta_0) \tag{S15}$$

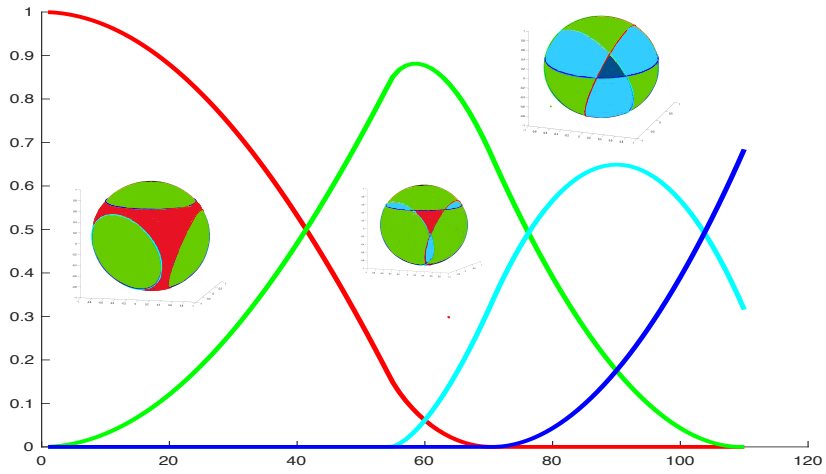

FIGURE S6. Relative volumes of Hb arms loss, given by the areas of intersection of the corresponding occultation caps, plotted as a function of the critical angle, θ_0. Here 0, 1, 2 or 3 arms loss areas are shown resp. in red, green, cyan and blue. Insets show the corresponding spherical sectors on the unit sphere. $\theta_0 = \pi/2$ corresponds to an infinite planar interface while larger values would correspond to a concave surface (not considered here).

$v_{12} = v_{123} = 0$

(2) 1- or 2-arm occultation: $\alpha/2 \leq \theta_0 \leq \pi - \alpha \approx 70.5288^o$ $v_{123} = 0$ while v_1 as above and

$$v_{12}/(2\pi) = 2\cos^{-1}\left\{\frac{1/3 + c^2\theta_0}{s^2\theta_0}\right\} - \ldots \quad (S16)$$
$$-4\cos\theta_0\cos^{-1}\left(\sqrt{2}\cot\theta_0\right)$$

(3) 1-, 2- or 3-arm occultation: $\pi - \alpha \leq \theta_0 \leq \pi/2$ v_1 , v_2 as above, while

$$v_{123}/(2\pi) = 3\gamma - 3k_g l \quad (S17)$$
$$= 3\cos^{-1}\left\{\frac{1/3 + c^2\theta_0}{s^2\theta_0}\right\} - \ldots$$
$$-6\cos\theta_0\left(\cos^{-1}\left(\sqrt{2}\cot\theta_0\right) - \frac{\pi}{3}\right)$$

Implementation code for the geometric calculations.

```
function  Qsphere(  )
alpha = pi-acos(1/3);
alphad = alpha*180/pi
calpha = -1/3; sa = 2*sqrt(2)/3;
k=0;
for i = 1:110
    k=k+1;
  theta0_d(k) = i;
  theta0 = (i/180)*pi;
  ct0 = cos(theta0);st0 = sin(theta0);
  v1=(1-ct0)/2;
  V3(k) = 0; V2(k) = 0;
  V1(k) = 4*v1;
  if (theta0 <= alpha/2 )
     v12 = 0;v123=0;
  else
    v12 =(.5/pi)*acos((1/3+ct0^2)/st0^2)-...
          (1/pi)*ct0*acos(sqrt(2)*ct0/st0);
      if (theta0 <= pi-alpha)
          v123 = 0;
      else
          v123=(.75/pi)*acos((1/3+ct0^2)/st0^2)-...
          (1.5/pi)*ct0*(acos(sqrt(2)*ct0/st0)-pi/3)-.25;
      end
  end
          V3(k) = 4*v123;
          V2(k) = 6*v12-12*v123;
          V1(k) = 4*v1-12*v12+12*v123;
          if (theta0 <= pi - alpha)
          V0(k) = 1-V1(k)-V2(k);
          else
          V0(k) = 0;
          end
end
plot(theta0_d,V0,'b-',theta0_d,V1,'g-',...
     theta0_d,V2,'r-',theta0_d,V3,'c-')
end
```

The average number of hydrogen bonds one water molecule of the first shell can form is given by (see Figure S6):

$$\langle \zeta(\theta_0) \rangle = 4V_0 + 3V_1 + 2V_2 + V_3 \tag{S18}$$

Water-solute interaction energetics

Apart from the geometric perturbation, the solute also perturbs the energetics of the water-water interactions in the first solvation shell. After inserting the solute, the average energy of a water of the first shell can be written as:

$$\langle \epsilon(\theta_0)\rangle_h = \frac{1}{2}[\langle \zeta(\theta_0)\rangle\{\langle u_{HB}\rangle f_{HB} + \langle u_{cage}\rangle f_{cage}\} + 4\langle u_{LJ}\rangle f_{LJ} + \epsilon_{SW}] \qquad \text{(S19)}$$

In the above equation, ϵ_{SW} represents the energy of interaction between solute and water whose expression is given in the eq. (9) of the main text. We have fitted ϵ_{SW} used in the equation (S19) for the four inert gases namely Ne, Ar, Kr, Xe & four molecular solutes methane, benzene, naphthalene and C_{60} to get free energy (ΔG) of transferring a non-polar solute (calculated from the eq. (4), of the main text) from the model which reproduces corresponding values of the experimental data. From the dependence of ϵ_{SW} on the solute radius and temperature for the four inert gases, ϵ_{SW} of the other two inert gases (Ne and Rn) were determined. It is to be noted that for calculating the density perturbation of water we have used a generic solute-water interaction term, which is same for all gases. On the other hand, for calculating the average energy of an average water, the best value of the solute-water interaction term has been optimized for each gas to reproduce experimental hydration free energy.

Procedure for parameterizing ϵ_{SW}. One practical problem of fitting our calculated thermodynamic quantities to the experimental ones is that our model uses a reduced unit. Hence, it is necessary to scale our results to have a fair comparison between theory and experiments. To scale the temperature, we have taken the model of Urbic et al. [3] where it was reported that the reduced temperature range 0.1 to 0.2 represents the liquid water range, therefore we scaled this range as 0 to 100 $^\circ C$.

To compare theoretical and experimental hydration thermodynamic results on equal footing, they must have same units. For this, we have taken the ratio of theoretical and experimental ΔG for hydration at the highest temperature for the liquid range i.e. at $T^* = 0.2$ for theory and $T = 100^\circ C$ for Xe as the scaling factor for the ΔG for hydration as shown in the following equation. Then ΔG of hydration at all temperatures from experiment are divided by this factor to make the experimental ΔG of hydration in reduced unit.

$$\text{scaling factor} = \frac{\text{Experiment}[\Delta G^{Xe}(T = 100^\circ C)]}{\text{Theory}[\Delta G^{Xe}(T^* = 0.2)]} = 5.58 \qquad \text{(S20)}$$

The solute-water interaction, ϵ_{SW} is optimized for each system so that it reproduces the experimental hydration free energy expressed in reduced units.

For optimizing the parameters, a1, a2, b1, b2, f1,f2, and f3 as given in eq. (9) of the main text, we have used simulated annealing (SA) as the optimization technique and the objective function (Obj) used in SA is given by the following equation:

$$Obj = \sum_{i=1}^{4}\sum_{j=1}^{n}(\Delta G_{ij}^{\text{Experiment}} - \Delta G_{ij}^{\text{Theory}})^2 \qquad \text{(S21)}$$

The indices i and j represent the number of molecules under consideration and the experimental data points, respectively.

We minimize Obj by varying ϵ_{SW}. Usually at least ten trajectories starting with different values of ϵ_{SW} are used in the SA calculation. The starting temperature was taken as 10000 in SA and the cooling rate used was 0.98. At each temperature 100000 steps were run. The values of the optimized parameters for all the four inert gases and the four molecular solutes taken in this study are given in the Table S1.

The atomic non-polar solutes (Ne, Ar, Kr, Xe) and molecular solutes (CH_4, C_6H_6, $C_{10}H_8$, C_{60}) are parameterized separately since they do not have the same interaction energy but we have kept the functional form of the interaction energy for the both types of the solutes same.

Values of the parameters. All the parameters for pure water are taken to be same as used in the ref [1]. The parameters are : $\epsilon_{HB} = 1.0$, $r_{HB} = 1.0$, $\epsilon_{LJ} = 0.1$, $\sigma_{LJ} = 0.7$, $a = 0.045$, $k_s = 78$ and $\epsilon_c = 0.18$. All calculations are performed in the reduced units e.g. temperature, $T^* = k_B T/|\epsilon_{HB}|$, distances are scaled by r_{HB} e.g. $\sigma_{LJ} = 0.7$ will imply that $\sigma_{LJ} = 0.7 r_{HB}$ and then scaled to the real units. The scaling from reduced units to the real units is done by multiplying the reduced hydration free energy with the *scaling factor* i.e. 5.58.

Units of the parameters in ϵ_{SW}. Eq. 9 of the main text as in reduced units is

$$\epsilon_{SW}(r^*, T^*) = (a_1 + a_2 r^{*2}) + (b_1 + b_2 r^{*2})\, T^* - (f_1 + f_2 r^* + f_3 r^{*2})\, T^* \ln T^* \quad \text{(S22)}$$

The parameters for the above eq. S22 is given in the following table.

TABLE S1. The values of the parameters for the inert gases, Ne, Ar, Kr, and Xe and molecular solutes CH_4, C_6H_6, $C_{10}H_8$, C_{60}

	a_1	a_2	b_1	b_2	f_1	f_2	f_3
Inert gases	0.05	-2.84	6.09	10.42	2.97	-22.85	19.39
CH_4,C_6H_6,$C_{10}H_8$,C_{60}	-0.55	-0.39	7.37	7.22	0.66	-9.99	1.77

The above eq. for ϵ_{SW} after multiplying with 5.58 *(scaling factor)* can be written in real units as follows :

$$\epsilon_{SW}(r,T) = 5.58\Big[\left(a_1 + \frac{a_2}{r_{HB}^2}(r_{HB}r^*)^2\right) + \left(b_1 + \frac{b_2}{r_{HB}^2}(r_{HB}r^*)^2\right)\frac{(T-173.15)}{1000} - \left(f_1 + \frac{f_2}{r_{HB}}(r_{HB}r^*) + \frac{f_3}{r_{HB}^2}(r_{HB}r^*)^2\right)\frac{(T-173.15)}{1000}\ln\frac{(T-173.15)}{1000}\Big]$$

$$\epsilon_{SW}(r,T) = \left(A_1 + A_2 r^2\right) + \left(B_1 + B_2 r^2\right)\frac{(T-173.15)}{1000} - \left(F_1 + F_2 r + F_3 r^2\right)\frac{(T-173.15)}{1000}\ln\frac{(T-173.15)}{1000} \quad \text{(S23)}$$

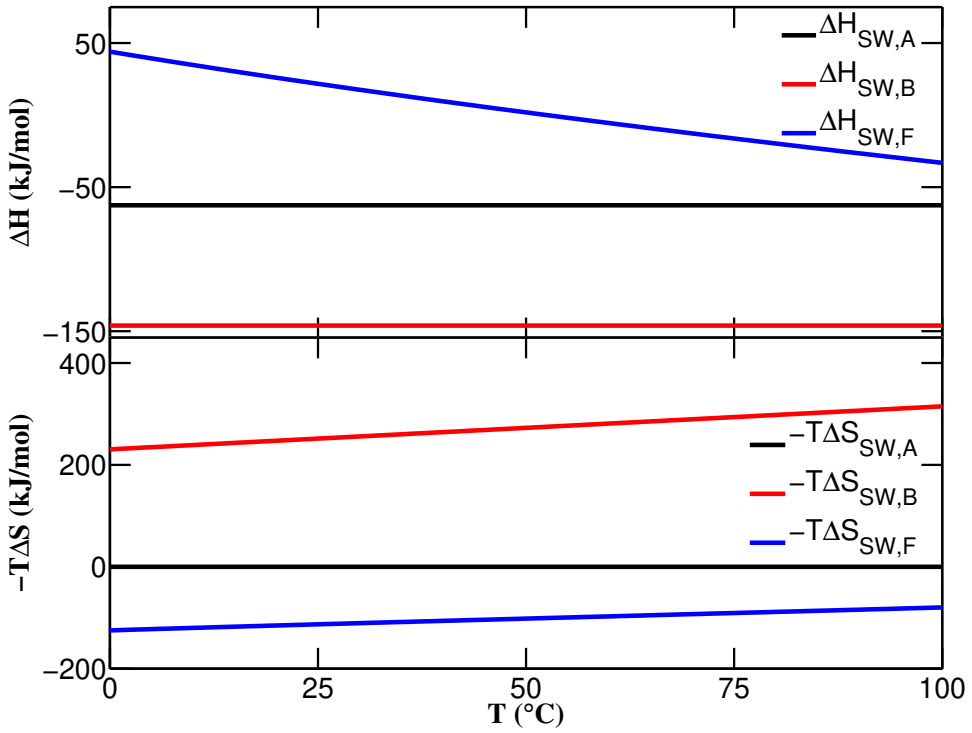

FIGURE S7. **The solute-water interaction components as given in eq. (9) of main text** for (a) enthalpy (b) $-T\Delta S$ vs temperature, for Ar. A, B, and F denote first, second, and the third terms of equation (S23).

where $A_1 = 5.58a_1 \;\; kJ/mol$; $A_2 = a_2 \left(\frac{5.58}{r_{HB}^2}\right) \;\; kJ/(mol.\mathring{A}^2)$; $B_1 = 5.58b_1 \;\; kJ/(mol.K)$; $B_2 = b_2 \left(\frac{5.58}{r_{HB}^2}\right) \;\; kJ/(mol.K.\mathring{A}^2)$; $F_1 = 5.58f_1 \;\; kJ/(mol.K)$; $F_2 = f_2 \left(\frac{5.58}{r_{HB}}\right) \;\; kJ/(mol.K.\mathring{A})$; $F_3 = f_3 \left(\frac{5.58}{r_{HB}^2}\right) \;\; kJ/(mol.K.\mathring{A}^2)$. The values of the parameters of eq. S23 are given in the Table 1 of the main text in lower case letters to be consistent with eq. 9.

The relative contributions of each of the terms of eq. (S23) are shown in Figure S7. The first term (the *'A'* term) of eq. (S23) has only enthalpy component, while the other two terms have both enthalpy and entropy parts. The second (*'B'*) term (S23) is the primary reason for the favorable solute-water interaction. In the entropy part, both second and third terms contribute with the second term being dominant.

ADDITIONAL PLOTS

Temperature Dependences for Nonpolar Inert Gases. Here, in the figure S8, we have compared the hydration thermodynamic data for the remaining three inert gases, He, Ne, Rn. Our results differ from experimental values for Rn but our model gives correct trends in accordance with experimental results of the first

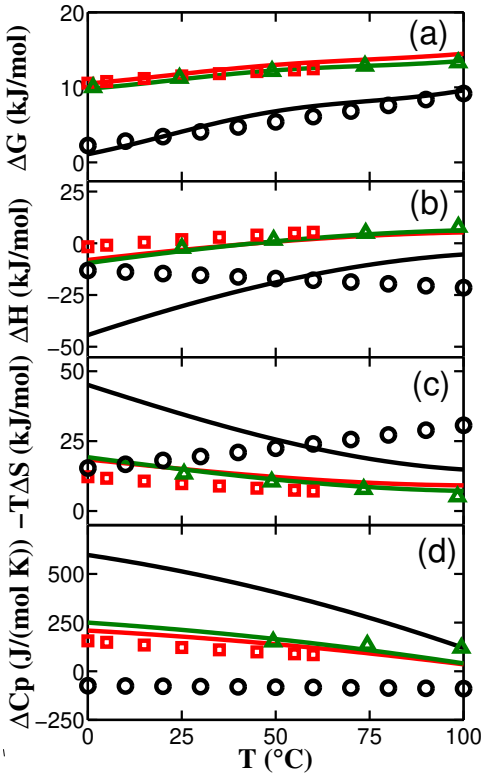

FIGURE S8. Comparison of thermodynamic quantities with temperature from our *Crustwater* theory (lines) for other three inert gases (Color code: Red (He), Green (Ne), and Black (Rn) with experiment (symbols).).[8, 9, 10]

five inert gases. The main reason of the deviation in the case for Rn is that the experimental $\Delta G(T)$ for Rn has downward curvature while other five have upword curvature.

Pressure Dependences for Nonpolar Inert Gases. Figure S9 shows the pressure dependence of the remaining five inert gases, He, Ne, Kr, Xe, and Rn. The Figure S10 shows the enthalpy and entropy components of SW and WW contribution to the solvation for Ar.

Pressure dependence for Methane. Figure S11 shows pressure effects on methane solvation results from the simulations of Chan [11] for comparison with prediction from our theory (Figure 8 of main text). The variation of ΔG with p from our theory has a similar trend to the extensive simulations of Chan. This positive slope is also consistent with that found by Koga [12]. The variation of both ΔH and ΔS with p from our model differs from the trend obtained from the simulation although it must be pointed out that the simulated values of enthalpy and entropy have large standard deviations. We can conclude that both hydration enthalpy and entropy is likely to be less variable as compared to their temperature dependence. This is main conclusion from our calculation and Chan's simulated results. This amounts

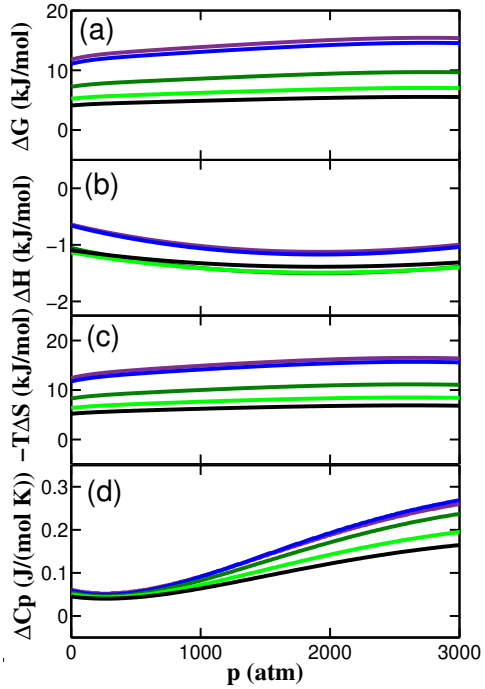

FIGURE S9. Variation of thermodynamic quantities with pressure for remaining five inert gases (Color code: Purple (He), Blue (Ne), Dark green (Kr), Light green (Xe), and Black (Rn)).

to small change in heat capacity as shown in panel (d) of Figure 8 in the main text as found in Chan's work and also in the experimental work of Hnedkovsky et al. [13] .

Size Dependences.

Reference comparison: size Dependence with $\epsilon_{sw} = \epsilon_{LJ}$. We looked at the size dependence of different hydration thermodynamic quantities using different types of solute-water interaction (ϵ_{sw}). Figure S12 shows the results with a constant value of ϵ_{sw} taken as ϵ_{LJ} at $T = 298$ K and $p = 1$ atm. As expected, with the increase of size values of all thermodynamic quantities increase. In that regard, this is similar to hard-sphere system.

Fractional population plots. Figure S13 shows the microscopic interpretations of pressure effects in terms of the model state populations. In our model the populations remain the same for pure water and water with a solute. The cages melt with the increase of pressure and converted to pairwise Hbond (f_{HB}). The variation of f_{HB} is more gradual with pressure due to lesser rate of cage melting. However, as higher pressure makes the system more 'tight', the open state population is actually decreasing with p. In other words, breaking of Hbonds lead to more LJ states with p. Generally speaking, the changes in populations are more gradual in the case of

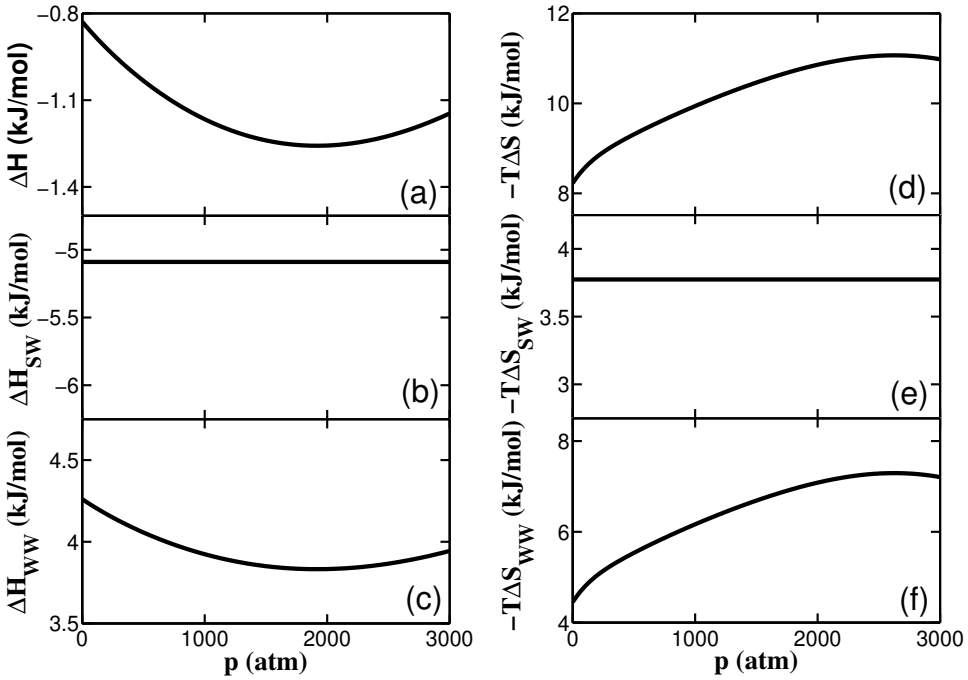

FIGURE S10. **The WW and SW components** of the enthalpy and entropy components of solvation free energy vs pressure, for Ar. The top line just shows the totals given in Fig 5 of the main text for Ar.

Solute	ΔG (kJ/mol)	ΔH (kJ/mol)	$T\Delta S$ (kJ/mol)	ΔCp (J/mol/K)
He	11.79(12.10)	-4.59(0.62)	-16.38(-35.87)	8.55
Ne	11.08(11.16)	-5.46(-02.48)	-16.54(-44.52)	9.59
Ar	8.42(8.20)	-11.99(-11.01)	-20.42(-65.03)	14.19
Kr	7.23(7.05)	-16.40(-14.02)	-23.64(-70.27)	16.95
Xe	5.22(5.38)	-26.49(-19.23)	-31.71(-82.24)	23.27
Rn	4.09(3.74)	-34.17(-15.05)	-38.27(-18.79)	28.15
CH_4	7.84(8.21)	-14.77(-10.98)	-22.52(-19.35)	188.59
C_6H_6	-0.53(-3.6)	-38.95(-29.6)	-33.69(-26)	168.31
$C_{10}H_8$	-10.3(-2.39)	-129.40	-22.61	53.5
C_{60}	-12.18(-17.4)	-215.11	-38.43	-111.24

TABLE S2. Comparison between experimental and theory for all the solutes studied here at 25°C and 1 atm. Values in the parenthesis are for experimental data.

pressure than that with temperature.

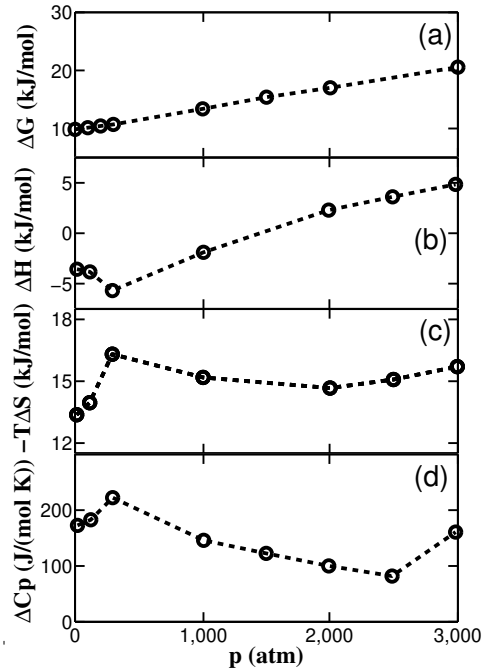

FIGURE S11. Variation of thermodynamic quantities with pressure for methane from ref [11] (circles), the dotted lines are drawn to only connect the data points.

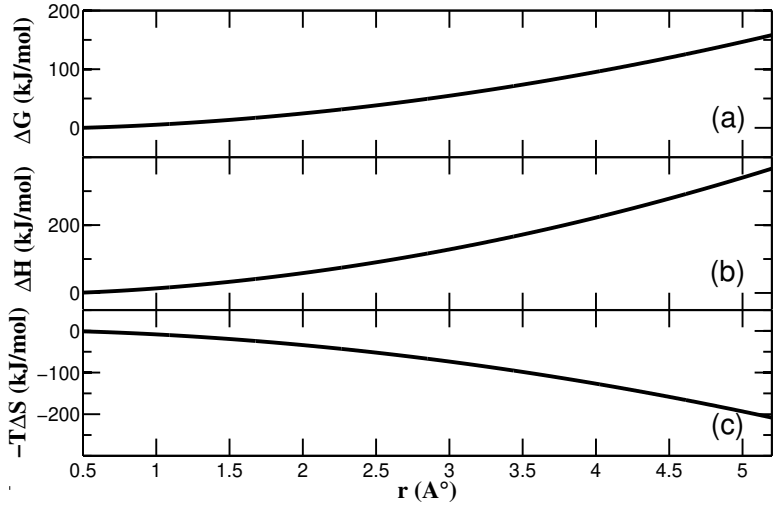

FIGURE S12. Variation of thermodynamic quantities with the size of a particle. Solute-water interaction is taken as ϵ_{LJ}.

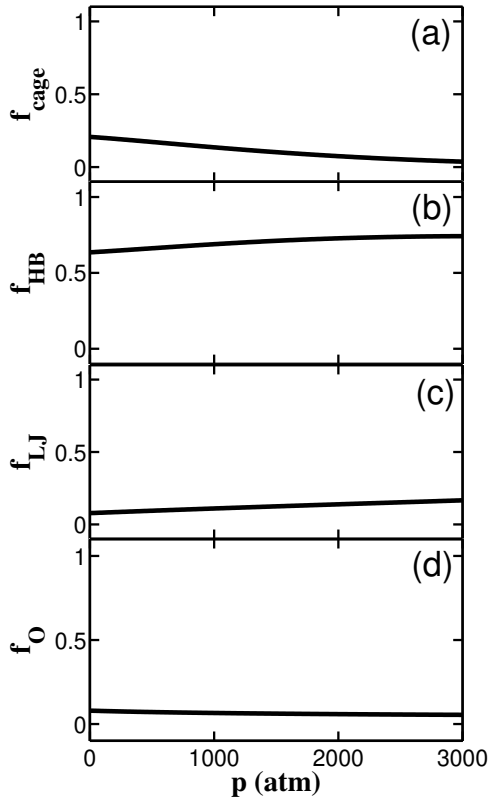

FIGURE S13. **How the model state populations in bulk water depend on p.** f_{HB} and f_{cage} are the fractional populations for the pairwise and cage hydrogen bonds states, respectively. f_{LJ} and f_O represent the fractional populations for the LJ and Open states, respectively.

References

[1] T Urbic. Analytical model for three-dimensional mercedes-benz water molecules. *Phys. Rev. E*, 85(6):061503, 2012.

[2] T Urbic. Liquid-liquid critical point in a simple analytical model of water. *Phys. Rev. E*, 94(4):1–5, 2016.

[3] T Urbic and K. A. Dill. Analytical theory of the hydrophobic effect of solutes in water. *Phys. Rev. E*, 96(3):1–12, 2017.

[4] Thomas M. Truskett and Ken A. Dill. Predicting water's phase diagram and liquid-state anomalies. *The Journal of Chemical Physics*, 117(11):5101–5104, 2002.

[5] T. M. Truskett and K. A. Dill. A simple statistical mechanical model of water. *J. Phys. Chem. B*, 106(45):11829–11842, 2002.

[6] Miha Lukšič, Tomaz Urbic, Barbara Hribar-Lee, and Ken A. Dill. Simple model of hydrophobic hydration. *The Journal of Physical Chemistry B*, 116(21):6177–6186, 2012.

[7] Ajeet Kumar Yadav, Pradipta Bandyopadhyay, Tomaz Urbic, and Ken A. Dill. Analytical 2-dimensional model of nonpolar and ionic solvation in water. *The Journal of Physical Chemistry B*, 125(7):1861–1873, 2021.

[8] R. Crovetto, R. Fernández-Prini, and M. L. Japas. Solubilities of inert gases and methane in H_2O and in D_2O in the temperature range of 300 to 600 k. *J. Chem. Phys*, 76(2):1077–1086, 1982.

[9] D. Krause and B. B. Benson. The solubility and isotopic fractionation of gases in dilute aqueous solution. iia. solubilities of the noble gases. *J. Solution Chem.*, 18(9):823–873, 1989.

[10] C. Lewis, P. K. Hopke, and J. J. Stukel. Solubility of radon in selected perfluorocarbon compounds and water. *Ind. Eng. Chem. Res.*, 26(2):356–359, 1987.

[11] M. S. Moghaddam and H. S. Chan. Pressure and temperature dependence of hydrophobic hydration: Volumetric, compressibility, and thermodynamic signatures. *J. Chem. Phys.*, 126(11), 2007.

[12] K. Koga and N. Yamamoto. Hydrophobicity varying with temperature, pressure, and salt concentration. *The Journal of Physical Chemistry B*, 122(13):3655–3665, 2018.

[13] Lubomir Hnedkovsky and Robert H. Wood. Apparent molar heat capacities of aqueous solutions of CH_4, CO_2, H_2S, and NH_3 at temperatures from 304 k to 704 k at a pressure of 28 mpa. *The Journal of Chemical Thermodynamics*, 29(7):731–747, 1997.