The formation of a cavity in water: changes of water distribution and prediction of the excess chemical potential of a hard-sphere solute under increasing pressure

Franca Maria Floris[1]

Dipartimento di Chimica e Chimica Industriale, Università di Pisa,
Via Giuseppe Moruzzi 13, 56124 Pisa, Italy

Abstract

This work deals with the formation of a spherical cavity in water along the isotherm at 298 K. A striking effect of increasing pressure was found on the radial distribution functions obtained by Monte Carlo simulations, with significant different behaviours observed when increasing the cavity radius at 8000 atm and 1 atm. Particular focus is on the value at contact, G(r), the central quantity in Scaled Particle Theory that is related to the derivative with respect to the radius of the work required to form the cavity. Within the limit of very small radii, exact conditions were applied to these two quantities. This allowed us to readily determine, at any pressure along the isotherm, the parameters of a simple model used to compute the excess chemical potential associated with the hydration of a hard sphere. This was made possible thanks to heuristic models used to describe how the number density of water changes along the isotherm and how the second moment of water distribution depends on the first moment. Use was also made of additional information on a cavity of molecular size. Apart from the dependence on pressure of hydrophobic solvation, this work also concerns calculation of the so-called cavitation contribution to the free energy of solvation when this is computed within implicit solvent models.

Email address: floris@dcci.unipi.it (Franca Maria Floris)

Keywords: cavity, SPT, simulations, pressure, hydration, distribution moments

1. Introduction

According to thermodynamics [1, 2], excess chemical potential or pseudochemical potential [1] expresses how the free energy of the system changes as a solute molecule is added to the solvent at a fixed position. For a hard-sphere solute-solvent potential, at constant T, this quantity is equal to the reversible work necessary to form a cavity in the solvent [3, 4, 5, 6]. The translation kinetic energy of the solute is added to obtain the chemical potential, from which other important thermodynamic quantities can be computed from derivatives with respect to state variables, such as pressure and temperature [1, 2].

This work studies the pressure dependence of the excess chemical potential associated with the insertion of hard-sphere solutes at infinite dilution in water along the isotherm of 298 K. A purely repulsive potential can be used to model interactions between a hydrophobic solute and water. However, it is also important for the computation of the so-called ”cavitation contribution” to the free energy of solvation within polarizable continuum models [7, 5] as the solute is enclosed in a molecular cavity defined by the union of spheres. When using these methods the focus is on the quantum treatment of the solute, and most of the computational time is spent for the electrostatic contribution [8], and, depending on the method used, for the dispersion contribution. [9]. It is in this context that it is useful to develop simple heuristic expressions in order to compute thermodynamic quantities related to the solvation process of these simple modeled solutes [5, 10].

At infinite dilution conditions, interactions between solute molecules can be ne-

glected and only solute-solvent interactions give a contribution to the excess chemical potential. This can be computed using a coupling parameter method [11], which requires information on how the solvent distribution function changes from the initial state of pure solvent to the final state in which the solute-solvent interaction is completely coupled. This method implies the study of "intermediate states" which are generally unrealistic. When applied to a hard-sphere solute-solvent interaction, the solute insertion process is equivalent to scaling the radius of a cavity from zero up to a final contact radius, r, as shown in the original paper of Scaled Particle Theory (SPT) [3]. Thus, our interest in the formation of a cavity in a solvent is extended to such small cavities that they cannot host any real solute.

Within SPT [3, 4, 6] , the excess chemical potential associated with cavity formation is related to the probability that no centers of the solvent molecules will be found in the spherical region defined by the contact radius. On the basis of statistical mechanics, this probability is expressed in terms of pure solvent quantities that can be computed from the number density and integrals involving many-particle correlation functions. These integrals define the second and higher moments of the pure solvent distribution functions and can be written in terms of probabilities that exactly n centers of the solvent molecules can be found in the cavity volume [3] . As specified in the literature [3, 12] , the excess chemical potential written in terms of moments is of general validity and can be applied to cavities of an arbitrary shape and using a realistic model potential for interactions between solvent molecules. However, as only the first two moments are easily available, its application is limited to very small cavities so that $n \leq 2$ or requires the computation of modeled probabilities. These can be computed within an information

theory based on the first two moments, as demonstrated by Hummer et. al. [12]. Here, we work within SPT and apply the exact relation only to such a small sized-cavity that the excess chemical potential is directly computed from the first two moments. Exact SPT conditions are then applied at an appropriate radius close to the extreme of the range where the contribution of higher moments vanish. These conditions involve derivatives of the excess chemical potential with respect to the cavity radius. In order to facilitate these calculations and apply SPT along the isotherm we describe the two first moments of water for spherical volumes as a function of radius and pressure. Such descriptions make use of a simple relation between the first moment and the water number density [12], which was modeled along the isotherm. This was also used in a heuristic expression proposed here to compute the second moment from the first moment using a modified Poisson distribution (see Appendix A).

Scaling the radius from small to larger sized-cavities so that a real solute can be hosted in, the approximate SPT expression [13, 14] or the more flexible expressions based on the thermodynamics of surfaces [15] are generally used [16, 5, 10]. The effect of pressure on parameters entering such expressions has been little explored until now. At fixed conditions of P and T, such parameters are preferably determined by fitting procedures. However, a complete study of the effect of pressure on the thermodynamics of cavity formation would require a great number of very expensive simulations.

Here, we test these simple models at a very high pressure by comparison with simulation results. We show that a less expensive parametrization based on exact relations gives results which are in agreement with those obtained by fitting. This

was made possible by the use of a relatively simple scheme formulated within the framework of SPT, with a limited number of conditions on the central function G(r), which is defined by the contact value of the cavity-solvent radial distribution function (rdf). As noticed above, these conditions were applied within the limit of a very small cavity, but some information on a molecular-sized cavity was employed in the parametrization. Differently from the approximate SPT expression, this is in fact necessary when using more flexible simple models.

2. Calculation

2.1. Excess chemical potential and G(r)

According to statistical mechanics [3], the excess chemical potential of a hard-sphere solute, here denoted by μ^*, can be computed from the probability that an empty region can be found in the solvent and it is written in terms of moments of the solvent distribution. Within the limit of a very small cavity, this exact relation involves only the first two moments,

$$\mu^* = -k_B T \; ln\left[1 - < n > + \frac{1}{2} \; < n(n-1) > \right] \qquad (1)$$

where n is the instantaneous number of molecular centers in the pure solvent contained in a sphere of radius r, whose average $< n >$ can be obtained from the number density ρ [12]. The third term in brackets gives the average number of pairs in the same volume and can be obtained from integrals involving the pair distribution function [3, 12] . The expression above is valid for cavity volumes for which the contribution of higher moments vanish ($n \leq 2$).

For larger cavities, the excess chemical potential is computed with a simple model formulated within the thermodynamics of surfaces [15]. We adopt the most com-

mon division [4, 17] related to a cavity of radius r, which here defines the accessible surface and the exclusion volume to oxygen nuclei positions of water molecules. Namely,

$$\mu^*(r) = 4\pi\tilde{\gamma} f_c(r) r^2 + P(\frac{4\pi}{3} r^3) \tag{2}$$

where $\tilde{\gamma}$ has the dimension of a surface tension and $f_c(r)$ is a function describing curvature corrections, which equals 1 for a cavity in the limit of an infinite radius. The average density of solvent centers on the cavity surface is given by $\rho\, G(r)$, and G(r) is related to μ^* by the important relation derived from SPT [3, 16],

$$G(r) = \frac{1}{4\pi r^2 \rho k_B T} \frac{d\mu^*}{dr}. \tag{3}$$

As remarked by Reiss et al. [3], μ^* equals the reversible work which is expended in the formation of the cavity. Given that this process is equivalent to the coupling of a hard-sphere interaction potential with contact distance r, the value of the rdf for the solvent centers excluded from the spherical region coincides at r with G(r). Thus, as pair correlation functions are commonly calculated in simulations, this relation is very useful to validate simple models used to describe the radial dependence of f_c in Eq. (2). The general form of this model leads to the following simple expression for the derivative of μ^* with respect to r,

$$\frac{d\mu^*}{dr} = 4\pi\tilde{\gamma} \left(2r - 2\tilde{\delta} + \frac{\alpha}{r^2}\right) + P(4\pi r^2) \tag{4}$$

where $\tilde{\delta}$ and α are parameters entering the curvature correction $f_c(r)$ [16, 10].

2.2. *Solvent Compressibility*

The coefficient of isothermal compressibility can be derived from density in accordance with the thermodynamic definition:

$$k_T^\circ = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T = \frac{1}{\rho}\left(\frac{\partial \rho}{\partial P}\right)_T \tag{5}$$

or from volume fluctuations in the NPT ensemble, namely,

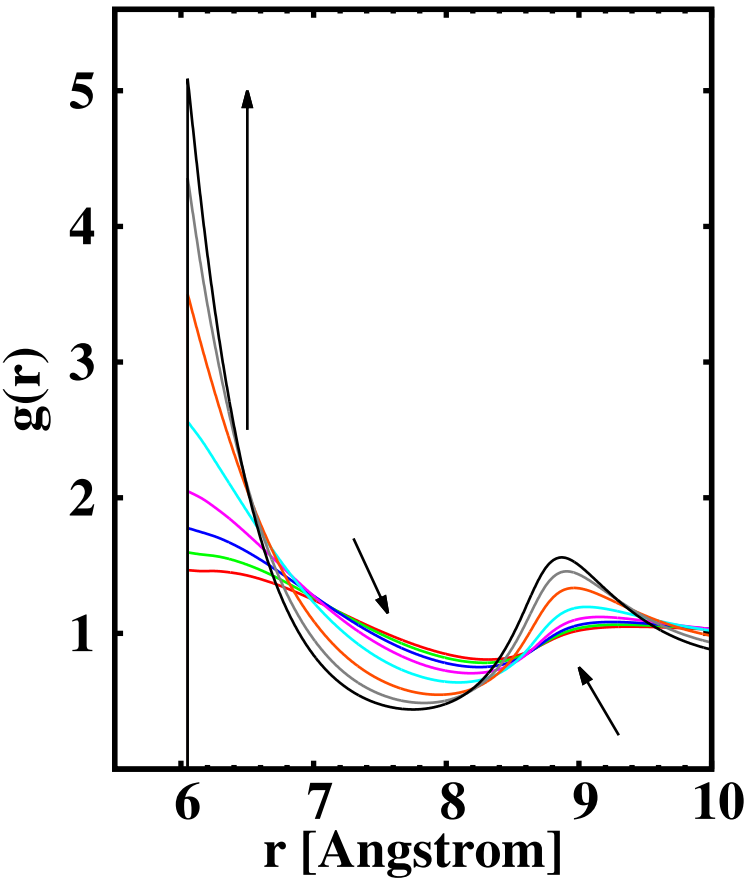
$$k_T^\circ = -\frac{< V^2 >_N - < V >_N^2}{k_B T < V >_N}. \tag{6}$$

3. RESULTS AND DISCUSSION

3.1. *Simulation Results*

NPT Monte Carlo (MC) simulations were run for hard-sphere cavities in 512 TIP4P [18] waters for contact radii up to about 6 Å. This was made by excluding the corresponding spherical volume to the motion of water-oxygen nuclei [19, 20]. Cavities in water at ambient conditions have been widely studied in our previous works [21, 10] and by other authors [4, 16, 6]. Since in this work these systems are used as representative of the low pressure limit, we briefly recall the main results. Concerning water distribution, for a cavity radius larger than approximately the most probable distance between the oxygen center with a methane-like solute, G(r) rapidly decreases, determining dewetting for nanometric cavities. At the same time, oscillations in the rdf present less pronounced deviations from 1, and cavity water correlations for similar r start to give positive contributions to the excess volume [21]. Fig. 1 shows an example of the striking effect of increasing pressure on the cavity center-O rdf. In addition to the increased rdf contact

Figure 1: Radial distribution functions vs r, the distance of the oxygen from the center of the cavity with contact radii of 6.05 Å, at T = 298.15 K and P from 1 to 8000 atm. Arrows indicate the direction of increasing pressure.



values with respect to those at 1 atm, at 8000 atm (Fig. 2) there is a remarkable difference regarding the persisting structure when increasing the cavity radius. In the range $4 - 6$ Å there is very little increase in the rdf contact value, and at the extreme of the range this is very close to the asymptotic limit at a large cavity radius ($P/(\rho K_B T)$).

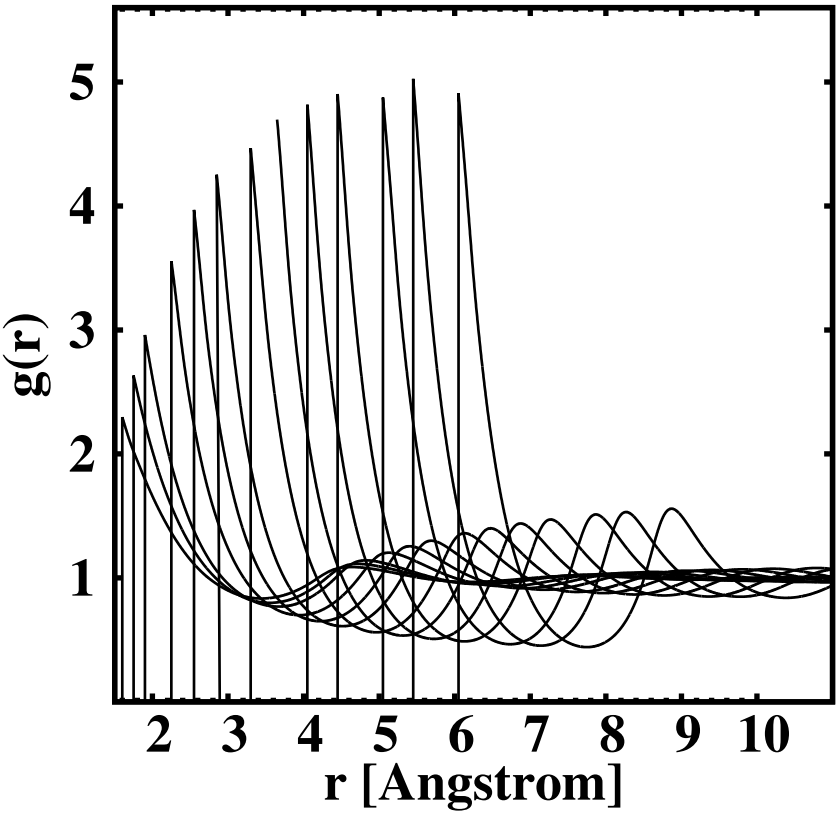


Figure 2: Cavity-water oxygen radial distribution functions for contact radii of 1.60, 1.75, 1.90, 2.25, 2.55, 2.85, 3.3, 3.65, 4.05, 4.45, 5.05, 5.45 and 6.05 Å, at T = 298.15 K and P=8000 atm. Results from NPT MC simulations of a cavity in 512 TIP4P waters using a modified version of the BOSS program[19].

In order to compute μ^* and G(r) from pure solvent quantities, NPT MC simulations of 512 TIP4P waters at T=298 K and several pressures between 1 atm and 10000 atm were run. The O-O rdfs are shown in Fig. 3, while the density and compressibility results are reported in Table 1. These are in good agreement with simulation results obtained with the same model [22, 23], but with a fewer number of water molecules (216 and 365). The comparison made in Fig. 4 with experimental data is in line with what has already been observed in the literature [22]. The same can be said about the comparison with simulation results with the TIP5P model [24].

Furthermore, expressions proposed in this work to describe the pressure dependence of density and the related quantities entering the exact relation (Eq. 1) of μ^* used for small cavities are validated by comparison with simulation results for radii of 1.6 Å , 1.75 Å and 1.90 Å . This validation is important in the parametrization of the μ^* expression used for larger cavities (Eq. 2). In order to establish how the surface tension parameter depends on pressure, additional information on a larger cavity is necessary. To this end, G(r) values were used from simulation results at various P along the isotherm, for a cavity radius of 6.05 Å . It can be noticed that this radius is small enough for the box size used so that systematic errors were avoided. Nevertheless, it is larger than the radius of a cavity appropriate to host a fullerene molecule. Since the rapid convergence of G(r) values observed when increasing the radius at 8000 atm and on the basis of the behaviour observed at ambient conditions, we think that, the procedure implemented to test simple models would give comparable results with fitting to data in a range up to approximately 10 $\mathring{A}$.

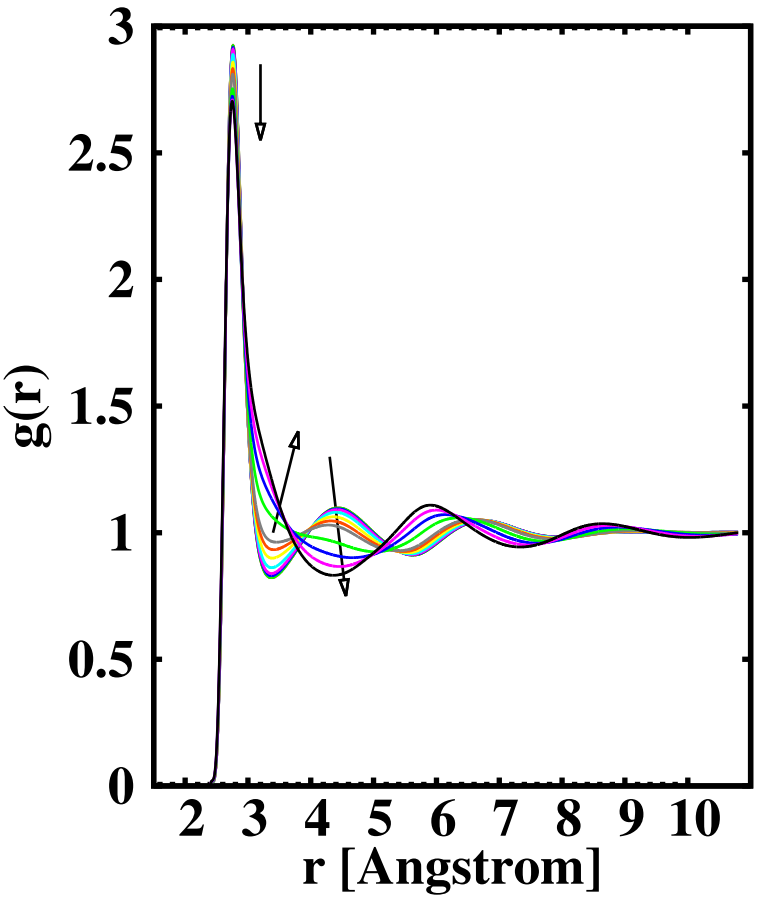


Figure 3: O-O radial distribution function for 512 TIP4P waters at T=298 K and P between 1 atm and 10000 atm. Arrows indicate the direction of increasing pressure.

Table 1: NPT MC results for the density and the coefficient of isothermal compressibility obtained for 512 TIP4P waters at 298.15 K and various P. The numbers in parentheses are the statistical uncertainties in the last digit.

$P\ (atm)$	$\rho\ (g/cm^3)$	$10^6 k_T^{\circ}(atm^{-1})$
1	0.99754(2)	51.93(1)
100	1.00276(5)	50.57(3)
200	1.00779(5)	44.84(6)
500	1.02294(5)	45.25(3)
1000	1.04587(4)	40.11(2)
1500	1.06675(5)	35.96(2)
2000	1.08560(5)	32.56(2)
4000	1.14779(6)	22.80(2)
6000	1.19479(6)	18.30(7)
8000	1.2349(2)	14.50(2)
10000	1.26846(6)	12.05(1)

3.2. P dependence of water density and compressibility

In this section we present models to describe how the water number density, ρ, changes along the isotherm at T 298 K. In agreement with the observed P dependence of the average volume of a fixed number of water molecules a good description was obtained by the following expression:

$$\frac{1}{\rho(P)} = t_0 + t_1 P + t_2 P^2 ln(P/P_0) + t_3 P^{2.5} + t_4 P^3 \tag{7}$$

where P_0 is the unity used for pressure, here 1 atm, t_0, t_1 , t_2, t_3 and t_4 are constant parameters. As shown in Fig. 4, this equation performs very well in fitting TIP4P simulation results as well as experimental data [25] (see Supporting Information(SI) for parameters). Less satisfactory fitting was instead found with the

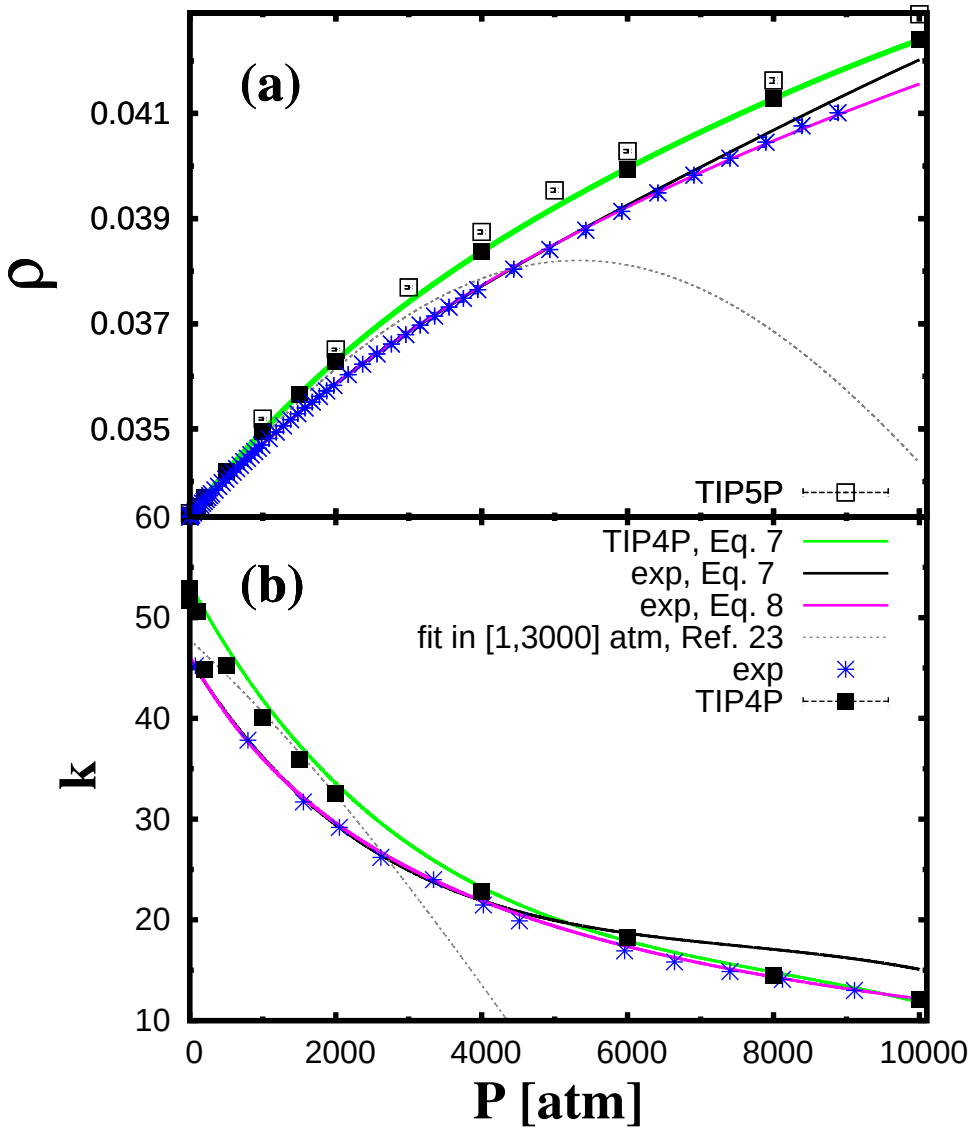


Figure 4: Pressure dependence of the water number density (a) and the isothermal compressibility (b) at T = 298.15 K. Simulation results of this work computed for boxes of 512 TIP4P waters are indicated by filled squares (for statistical uncertainties and units see Table 2). Comparison is shown with other simulation results (empty squares) obtained for boxes of 512 TIP5P waters [24], and with experimental results (stars). Lines indicate results obtained using least-square fits to simulation and experimental data of the number density with Eqs. 7 and 8 from which the corresponding isothermal compressibility was obtained using Eq.5. Short dashed lines refer to literature data [23] obtained from the quadratic fit of $< V >$ of 365 TIP4P waters.

semi-empirical Tait equation and even the modified Tait equation [26]. Instead, a quadratic fit of the average volume was found acceptable only in a limited pressure range up to approximately 3000 atm. This simple model has been used in this range to fit data of a smaller box of TIP4P waters [23]. The curve is reported in the figure and clearly shows disagreement with our data for pressure greater than 3000 atm.

In describing the density dependence on pressure, a stringent test on the validity of models is provided by the examination of their performances on isothermal compressibility, which is related to the pressure derivative of the density by Eq. (5). As shown in Fig. 4 (b), Eq. (7) is able to predict generally quite well both simulation and experimental data. However, systematic errors shown at P greater than 6000 atm when fitting experimental data suggest that in this range results obtained with this equation should be interpreted with caution. Thus, to improve compressibility results, an alternative expression was also considered. The quantity $-k_T^{\circ}V$ obtained from the simulation results gives an estimate of the slope of the volume curve plotted against P. The fitting of this curve suggested the following equation for P dependence of density

$$\frac{1}{\rho(P)} = \frac{1}{\rho_0} + \frac{(ab-c)ln(bP+1)+bcP}{b^2}. \tag{8}$$

(See Supporting Information(SI) for parameters a, b, c). This equation can be seen as a modified Tait equation and gives good density fitting with improved slope at higher pressures when comparison is made with experimental data. However, on simulation results, no significant improvement with respect to Eq. (7) was observed and in this case we show in the figure only the curves that fit TIP4P data

with this equation.

3.3. Small cavities ($n \leq 2$): P dependence of μ^* and G(r)

Here we show results obtained from the exact relation written in terms of the first two moments, which is valid for small cavities with a radius so that no more than two centers of the solvent can be found in the spherical volume. At each value of P over the range from 1 to 10000 atm, the average number of oxygen pairs in a specific spherical region was computed by numerical integration of the O-O rdf in pure water (see Fig. 3). Once this quantity was obtained, μ^* and G(r) were computed using the exact relations, Eq. (1) and Eq. (3) respectively [10]. As expected, the cost of cavity formation increases as the pressure increases when radii are sufficiently large, while for cavity radii less than 0.6 Å there is no significant effect of the pressure (see Fig. 5(a)). Namely, with the "radius" of a water molecule of around 1.35 Å, these cavities are inappropriate to host a real solute. Following Reiss et al., in this case, one can think that a point solute has been added to the solvent. However, when excluding a volume to the motion of solvent centers it is not always necessary to associate this with the addition of a solute molecule. These small cavities show, nonetheless, a significant effect of increasing pressure.

In the examined range, G(r) increases with the radius with larger slopes at greater P. As shown in Fig. 5(b), comparison with values directly computed by simulations at 1 atm and 8000 atm is good. The sudden drop in the curves indicates the failure of Eq. (1) for larger cavity radii so that $n > 2$ [4]. This equation is valid up to a radius that slightly decreases as the pressure increases, passing from 1.83 Å at 1 atm to 1.77 Å at 10000 atm. At a first glance, this decrease appears

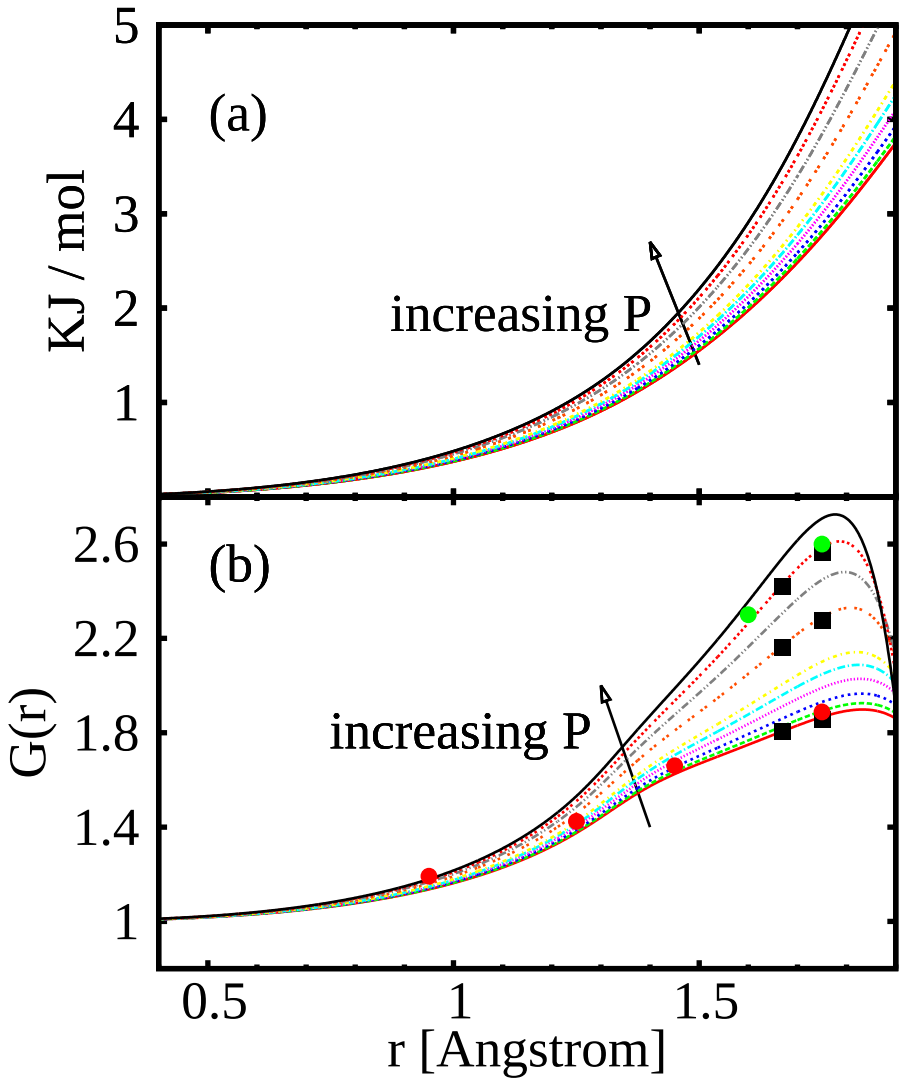


Figure 5: Small cavities ($n \leq 2$) at T=298 K. Pressure effect on the radial dependence of (a) μ^* (Eq. (1)) and G(r) (b) (Eq. (3)) over the range from 1 to 10000 atm. In Eq. (1) $< n >$ and $< n(n-1) >$ were obtained from simulation results of ρ and O-O rdfs (Fig. 3) [3, 4, 10]. Filled squares represent G(r) results from the same equations but with $< n >$ and $< n(n-1) >$ from Eq. (7) and Eq. (9). G(r) values from the contact values of the cavity-water oxygen rdfs computed by MC simulations of cavities in TIP4P waters are also shown for some radii at 1 atm (filled red circles) and at 8000 atm (filled green circles).

to be related to the slight decrease in the most probable O-O distance observed at greater P (Fig. 3). However, it certainly arises from the increased number of oxygen pairs obtained by integration of O-O rdf in a range of distances up to twice the cavity radius (see Eq. (5) of Ref. [10]). In this regard, we can notice that rdf values increase with an increase of pressure at distances of the first peak tail and around the first minimum. As a consequence, the probability of occurrence of oxygen triplets in a sphere of fixed radius increases with P, so determining the observed smaller range of applicability for Eq. (1).

For a cavity of radius r, a quadratic fit of μ^* against pressure can be used. Alternatively, using the exact relation written in terms of the first two moments, it is possible to exploit the well established dependence on pressure of ρ (see Section 3.2). To follow this route it is convenient to describe indirectly the P dependence of the average number of oxygen pairs observed in a spherical region by its dependence on $< n >$. It was found that the natural logarithm of this quantity can be expressed as

$$log\Big[\frac{< n(n-1) >}{2}\Big] = B \times \Big[log\Big(\frac{< n >^2}{2}\Big) - < n > \Big] \tag{9}$$

where

$$< n >= \frac{4}{3}\pi r^3 \rho(P) \tag{10}$$

and the factor B is a function of the cavity radius and of $< n > (P)$ as detailed in Appendix A, where a justification of Eq. (9) is given [27]. We recall that we apply Eq. (9) to small radii such that no more than two water-oxygen centers can occur inside the spherical volume. In this case $< n(n-1) > /2$ is equal to the

probability that exactly 2 centers can be found in this volume [3]. Therefore, this equation is consistent with the probability expression found within information theory when a Poisson default model is used (see for example Hummer et al. [12]). However, in the present work, differently from solutions coming from this theory, the B factor is sought as function of $< n >$ instead of n. This implies that parameters are here independent of $< n >$. We notice that the form proposed to describe how factor B depends on r and $< n >$ correctly gives a number of pairs equal to zero when $< n >$ goes to zero and fits the data very well ($n \leq 2$). Hence, using ρ from Eq. (7) in Eq. (10) and introducing $< n >$ in Eq. (9), μ^* can be computed from the first two moments by Eq. (1) and, in this manner, its derivatives with respect to r and P can be readily evaluated along the isotherm. In particular, the first and second derivatives with respect to r were computed at any P to obtain G(r) and $G'(r)$. Comparison with results obtained using $< n >$ and $< n(n-1) >$ (via O-O rdfs integrals) from simulations is very good over the range of pressures investigated. For simplicity, in Fig. 5 (b) comparison of G(r) is shown for some selected values of P.

3.4. Parametrization of the approximate SPT expression and simple models

In this section, the approximate SPT expression and simple models derived from the thermodynamic of surfaces [15] are used to estimate μ^* and G(r) for cavity radii larger than 1.7 Å. As already cited in the literature [28, 5, 10, 6, 29], similarly to these simple models, the approximate SPT expression can be written as the sum of a surface and volume terms. Nevertheless, a fundamental distinction between them concerns the limiting radius where they have been formulated, this being subnanometric for the approximate SPT and macroscopic for simple

models. Such a distinction is important because the number of independent parameters is different and these correspond to different quantities related to one other. This implies that, despite their similarities, these expressions can present different problems when parametrized at very high pressure.

3.4.1. Approximate SPT: the change of a_w with increasing P

The simplest and most commonly used SPT expression needs only the solvent number density and the definition of the size parameter of the solvent molecule (a_w), which can be identified with its hard-sphere diameter. Inadequacies of this expression are well known from studies at ambient pressure [16, 5, 21, 10], but undoubtedly it is very attractive due to its simplicity. As a first approach, for water solvent, this parameter could be fixed at the distance corresponding to the first maximum of O-O rdf (Fig. 3). This is very slightly influenced by P, changing by only 0.01 Å in passing from ambient pressure to 10000 atm. At ambient conditions, its value is 2.76 Å which is very close to the value proposed for a_w in the earlier work of Pierotti [13]. However, the comparison with simulation results of μ^* and G(r) for hard-sphere solutes in TIP4P water has suggested that the optimal value for this parameter is around 2.9 Å [5, 10].

On the basis of the above, only a slight dependence on P is predictable for this parameter in the case that it is assumed to be related to the most probable O-O distance. Hence, once the functions $a_w(P)$ and $\rho(P)$ are established, the approximate SPT expression can be used to obtain an estimate of the surface tension parameter [28, 5, 29]. With the assumptions made above, $\tilde{\gamma}$ would increase with increasing P because of the dominant effect of increasing density. Is this correct?

In order to answer this question, we compare the G(r) values calculated at 8000 atm using the approximate SPT expression with those obtained from Monte Carlo simulations.

Fig. 6 (a) demonstrates that the approximate SPT expression completely fails in describing this quantity at high pressure. The main problem with this expression is its incapability to scale correctly from small to large cavity radii. Differently to what has been observed at ambient conditions [10], comparison with simulation results shows a qualitative disagreement. This happens for values of a_w in a quite large range ($1.50 - 2.94$ Å). Reducing this parameter, the convergence to the asymptotic value of G(r) is as rapid as that observed for simulation results, even if with an incorrect sign of the curve's slope. This indicates that a simple reduction of the value of this parameter is not sufficient to significantly improve the performance of the approximate SPT expression. However, it would seem reasonable to presume a value for $\tilde{\gamma}$ which at 8000 atm is lower than that at 1 atm, in contrast with what is found if a_w is chosen on the basis of the most probable O-O distance. Namely, this is valid within the usual approach for which the excluded volume times P is included in μ^*, as in Eq. (2).

3.4.2. Fitting with the simple model at 8000 atm

Working within the framework of the thermodynamics of surfaces (Eq. (2) and Eq. (4)), with the simplest model ($\alpha = 0$) for the cavity surface term, one has to fix the value of two independent parameters, $\tilde{\gamma}$ and $\tilde{\delta}$. These can be determined from the linear fitting of the surface contribution to the derivative of μ^* with respect to r. Linear fitting of data at 8000 atm compared with that at 1 atm shows a generally

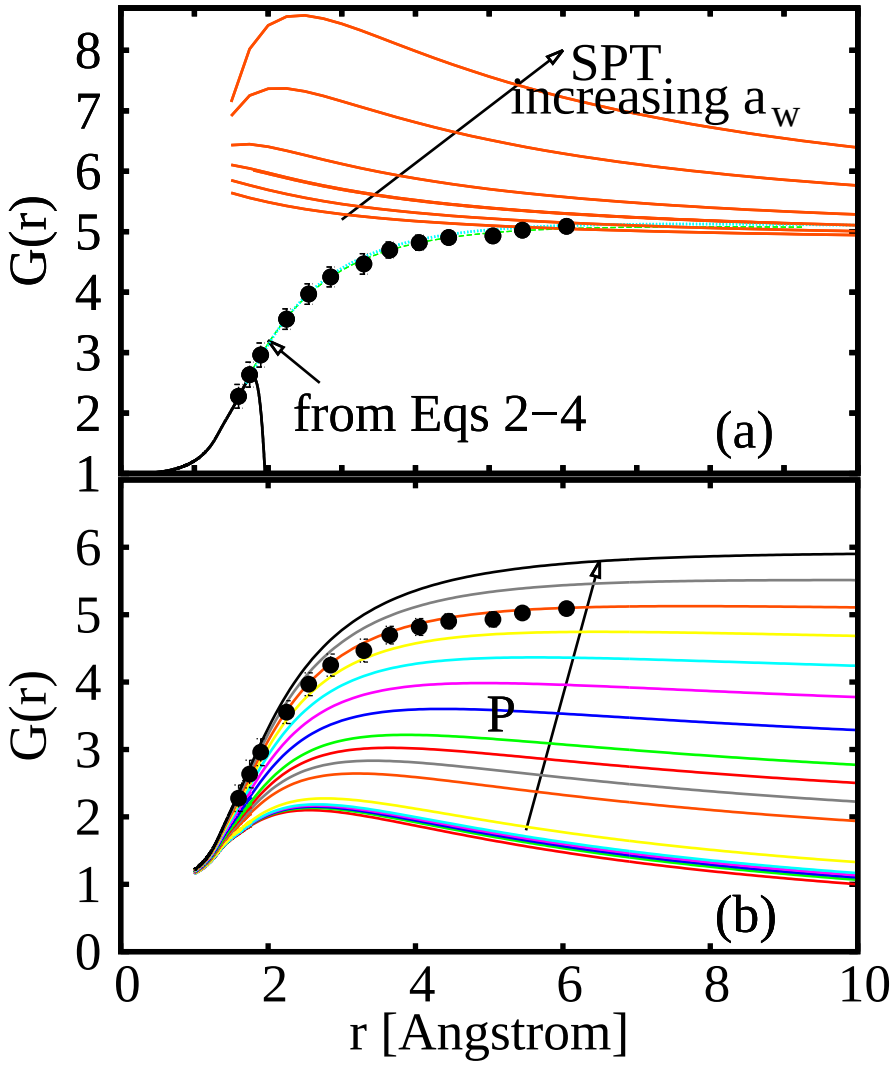


Figure 6: Contact values of the cavity-water oxygen rdfs for cavities in TIP4P water at 298 K. (a) Results at 8000 atm: from MC NPT simulations (circles with error bars); from approximate SPT expression [13, 14, 10] with values for the water diameter parameter (a_w) between 1.5 Å and 2.94 Å (orange lines); from fitting with the simple model, i.e. using Eqs. 2 - 4 as indicated by the arrow (read text); results for small cavities from Eq. 1 (solid black line). (b)Effect of increasing pressure, as indicated by the arrow, over the range from 1 to 10000 atm for G(r) results obtained from Eqs. 3 and 4 using exact SPT conditions. For the parameters dependence on P see Fig. 7. As in (a), circles with error bars are used for MC NPT simulation results at 8000 atm.

better performance together with a striking reduction of $\tilde{\gamma}$ and the increase of $\tilde{\delta}$ [10] ($\tilde{\gamma}$ = 35.4 $dynes/cm$ and $\tilde{\delta}$= 3.8 Å when radii less than 2 Å are not included in the data). In contrast to the approximate SPT expression, this model gives a good description of G(r) (Fig. 6 (a)), while the addition of the term depending on α is necessary to improve the fit if radii less than 2 Å are included in the data. In this case $\tilde{\gamma}$ = 43 $dynes/cm$, $\tilde{\delta}$= 4.1 Å and α = 6.4 Å^3 were obtained.

3.4.3. Using exact SPT conditions to parametrize the simple model along the isotherm

Alternatively to fitting, once $\tilde{\gamma}$ is fixed, parameters $\tilde{\delta}$ and α can be obtained by imposing continuity for G(r) and its derivative at a cavity radius value at the extreme of validity of Eq. (1). Results are shown in Fig. 6 (b) and compared with simulation results at 8000 atm. By using this procedure, dependence on P of both parameters is obtained by exploiting Eqs. (7)- (10) for an assumed expression $\tilde{\gamma}(P)$. Results presented in Sections 3.4.1 and 3.4.2 suggest a decrease in this parameter with increasing pressure when very far from ambient conditions. In fact, reasonable agreement with G(r) simulation data at 8000 atm is obtained for $\tilde{\gamma}$ in the range of 35-48 dynes/cm. [10].

However, despite the failure to predict correct results for radii less than 6 Å , at large radii the approximate SPT expression can be used to compute $\tilde{\gamma}(\rho)$. Using the ρ value from simulations, values of $\tilde{\gamma}$ in the range above were obtained from the approximate SPT expression for a_w in the range between 2.2-2.4 Å. Thus, it could appear justifiable to fix the high pressure asympototic value of this parameter at 2.4 Å , which is close to the O-O minimum contact distance in water. Dependence on P of this parameter might be modeled by assuming a transition

between the values of the low and high pressure limits. An example is given in Fig. 7 (a)(blue curve). We notice that the corresponding $\tilde{\gamma}$ (P) profile (green curve) predicts reasonable values at 1 and 8000 atm where it shows a positive and negative slope respectively.

However, this approach can result somewhat arbitary without some additional information on the pressure at which the slope sign changes. For this reason the pressure profile of parameter $\tilde{\gamma}$ was preferred based on simulation results of G(r) obtained for the largest cavity studied in this work (6.05 Å). When these values were introduced into the approximate SPT expression, the corresponding a_w were found to be in good agreement up to 8000 atm with the hypotized pressure profile for this parameter (in Fig. 7 (a) you can compare squares with the blue line).

On the contrary, when using Eq. (3) and Eq. (4), $\tilde{\gamma}$ was varied until G(r) values reached agreement with simulation results. At the same time, parameters $\tilde{\delta}$ and α were determined by the requirement of continuity for G(r) and $G'(r)$ at a cavity radius value of 1.67 Å with the values obtained from Eq. (1). This was readily done at any possible value of P by exploiting Eq. (7) and Eq. (9). Fig. 7 shows the pressure dependence for all parameters entering Eq. (4), including w_0, which was finally fixed by imposing continuity for μ^*. Results confirm that the slope sign for $\tilde{\gamma}$ is positive at 1 atm and negative at 8000 atm, even if the profile in comparison to that obtained from the approximate SPT presents a less pronounced curvature with a different position for the maximum value. Nevertheless, such disagreement is within statistical uncertainties on $\tilde{\gamma}$, which are unfortunately very sensitive to statistical uncertainties on G(r).

In contrast to $\tilde{\gamma}$, all the other parameters show a pressure dependence with positive

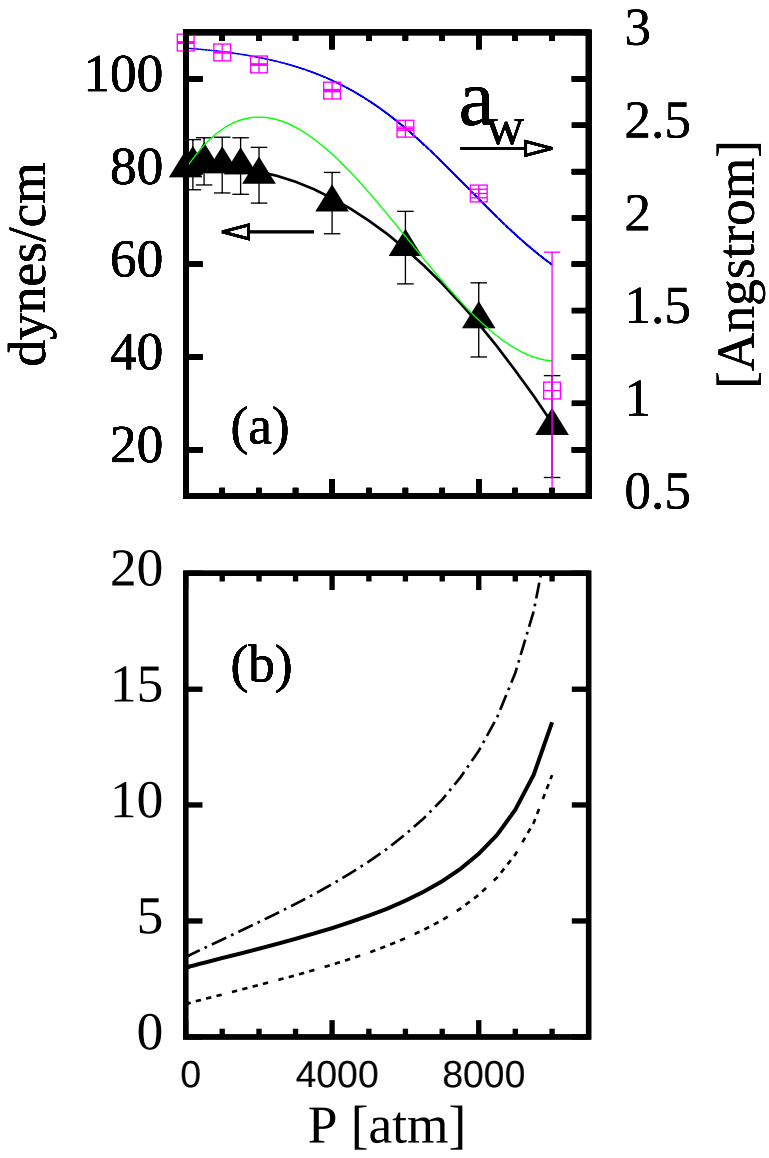


Figure 7: P dependence at T=298 K for parameters entering μ^* computed by using the simple model (Eq. (2). (a) $\tilde{\gamma}$ obtained from G(r) for r = 6.05 Å (triangles, left scale) fitted by a quadratic function (solid black line). The squares (right scale) are for the corresponding values of SPT parameter a_w. The green line represents $\tilde{\gamma}$ values derived by approximate SPT from a_w(P) described by the blue line (right scale). (b) Parameters entering $f_c(r)$ for $\tilde{\gamma}$ described by the solid black line in (a) : $\tilde{\delta}$ (solid line) in Å, $\tilde{\alpha}$ (dotted line) in Å^3, $w_0/(4\pi\tilde{\gamma})$ (solid dotted dashed line) in Å^2 (see Eq. (3) of Ref. [10]).

slope throughout the examined range with a steep increase at P greater than 8000 atm. Such behavior is not of easy interpretation in the absence of a clear physical meaning of these parameters. We merely note that this alternative parametrization leads to values which are in agreement with those obtained by fitting, in particular at 8000 atm. Thus, contact values of cavity-O rdfs at 8000 atm from the radial derivative of μ^* are in agreement with simulation results and fitted values (Fig. 6 (a)). In Fig. 6 (b), results from the same equation are shown over the range of P from 1 to 10000 atm.

4. Conclusions

Simulation results of water distribution around a cavity at 8000 atm presented in this work clearly show a very different behavior with respect to 1 atm when increasing the cavity radius. In particular, the concurrent progressive dewetting at the contact distance is peculiar of low pressure conditions, while at very high pressure there is persisting structure.

The rdf's contact value, G(r), is confirmed to be a valid quantity in parametrizing and testing heuristic expressions for the excess chemical potential. In contrast to the approximate SPT expression, simple expressions formulated within the thermodynamic of surfaces are able to catch the main features of cavities in water over a wide range of pressure. However, in this framework there are possible improvements by testing pressure derivatives of the excess chemical potential on excess volumes and excess compressibility. To this aim, the method adopted here to find pressure dependence of parameters appears quite convenient due to the use of pressure dependence of pure water quantities and only limited and easily available information on cavities in water. At least for the expression tested in this

work, which adopts the common division of the excess chemical potential in terms of the accessible surface and the exclusion volume, it was found that parameters determined with this method are in agreement with values obtained from normal fitting procedures.

5. Appendix A

We notice that Eq. (9) has been applied to small cavities such that the maximum number of oxygen centers observed in pure water in a sphere of radius r is 2. In this case, the average number of oxygen pairs observed in the same region is the probability of finding exactly 2 centers[3]. Given that this is a rare event, as first approximation one can assume a Poisson distribution [30]

$$P_2 = \frac{< n >^2}{2!} exp[- < n >]. \tag{11}$$

Thus, Eq. (9) corresponds to a modified Poisson distribution, namely

$$P_2 = [\frac{< n >^2}{2!} exp[- < n >]]^{B(r,<n>)} \tag{12}$$

where B is a function of the cavity radius and on the solvent density through the average number of oxygen centers in the spherical volume. It was found that, regardless of the pressure, the B value is generally greater than 1 when $< n >$ is less than 1. For radii at the extreme of the interval for which the maximum value of n is 2, $< n >$ is around of 1 and B approaches 1 (Fig. 8 (a)).

Therefore, a smaller average number of pairs is observed with respect to the Poisson distribution for smaller radii. To further examine deviations from the Poisson

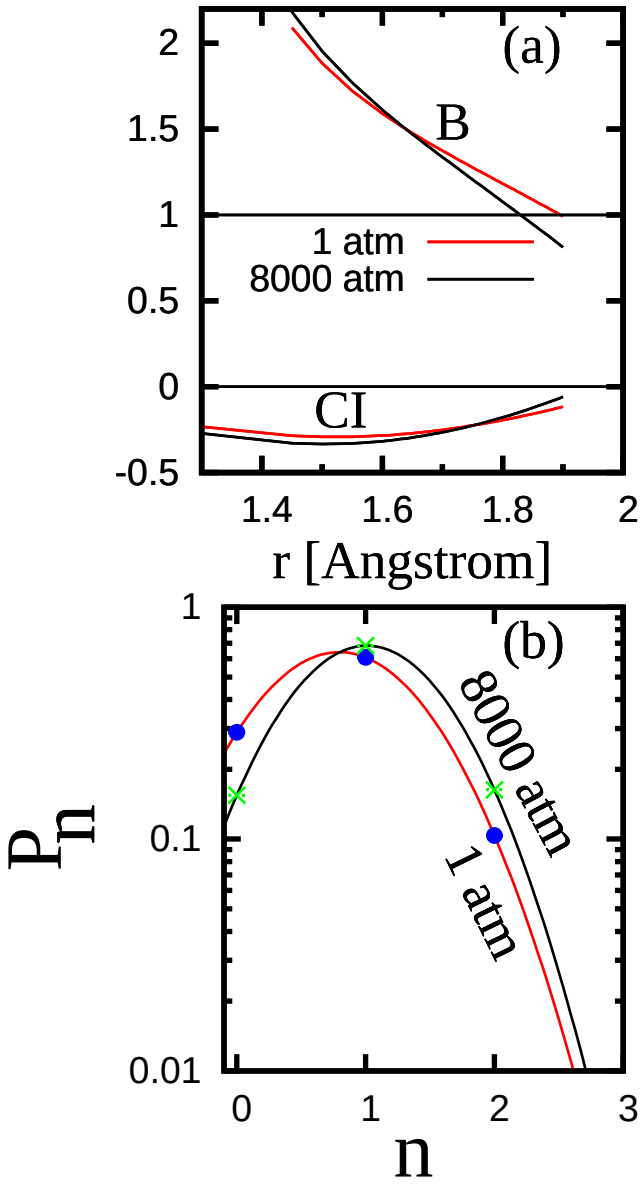


Figure 8: (a)The radial dependence of the exponent B in the modified Poisson distribution (Eq. (12)) and the correlation index CI (Eq. (14)) at 1 atm and 8000 atm. B computed from Eq. 15 in terms of $< n >$ using Eq. 7 for ρ. Crosseover with horizontal lines indicate cavity radii at which the modified distribution coincides with a Poisson distribution. (b) P_n, the probability to find exactly n water-oxygen in a spherical region. Results refer to a radius of 1.8 $\mathring{A}$ at 1 atm and 8000 atm. Points indicates simulation results, while the curves comes from parabolic fitting of the probability described with the modified Poisson distribution with B from eq. 15.

distribution, also the ratio between variance and $< n >$ has been taken into account. This ratio is 1 for the Poisson distribution of a certain, unlimited number of events. When n can assume a maximum value of 2, only three possible events can occur, and if also P_1 follows a Poisson distribution, this ratio would be:

$$\frac{\sigma_n^2}{< n >} = \frac{< n >}{1 + < n >} + 1 - < n > . \tag{13}$$

By assuming that $< n >$ is given by Eq. (10) and computing variance from $< n >$ and the average number of pairs, the following index was computed:

$$CI = \frac{\sigma_n^2}{< n >} - \left(\frac{\sigma_n^2}{< n >} \right)_{Poisson} . \tag{14}$$

Here the same symbol used for the clustering index introduced for comparison with the Poisson distribution in studies of drop clustering is adopted. A plot of CI versus the cavity radius is shown in Fig. 8 (a) for pressures of 1 atm and 8000 atm. It can be noted that CI is zero approaching the radius at the extreme of the interval for which the maximum value of n is 2. This radius is quite close to that at which the exponent B in Eq. (12) is 1.

In order to compute the pressure derivative of the excess chemical potential, dependence of B on $< n >$ was studied at fixed r by looking for a model that at the same time works well for different cavity radii. In this respect it was also important that optimized parameters showed continuous dependence on r. For cavities in which the maximum number of oxygen centers is 2, it was found that such dependence is reasonably described by the following expression

$$B\Big[r, < n > \Big] = b_0(r) + b_1(r) < n > + b_2(r) < n >^2 + \\ + \frac{b_3(r)}{< n >} \tag{15}$$

where parameters b_0, b_1, b_2 and b_3 are functions of r. Also a second-degree polynomial fit was taken into account as it is also able to describe the observed dependence. Nevertheless, Eq. (15) was prefered on the basis of a systematically lower value of χ^2_ν and generally better behavior of residual observed by comparing the two models for the same choice of weights.

Hence, from a study of the number of pairs against cavity radius at a fixed pressure, parameters entering Eq. (15) can have the general form

$$b_k(r) = \alpha_0 + \sum_{l=1}^{l_1} \alpha_l r^l + \sum_{l=1}^{l_2} \beta_l r^{-l} \tag{16}$$

with α_l and β_l independent of density. A strict test of radial dependence of these functions regards the prediction of G(r) values by computing the derivative of $\mu*$ with respect to r. In fact, this implies also a good description of first derivatives of these functions. A possible radial dependence of b_0, b_1, b_2 and b_3 is given by the following expressions:

$$b_0(r) = \alpha_0 + \alpha_1\, r + \alpha_2\, r^2 + \frac{\beta_1}{r} + \frac{\beta_2}{r^2} \tag{17}$$

$$b_1(r) = \frac{\beta_1}{r} + \frac{\beta_2}{r^2} + \frac{\beta_3}{r^3} + \frac{\beta_4}{r^4} + \frac{\beta_5}{r^5} \tag{18}$$

$$b_2(r) = \frac{\beta_4}{r^4} + \frac{\beta_5}{r^5} + \frac{\beta_6}{r^6} + \frac{\beta_7}{r^7} + \frac{\beta_8}{r^8} \tag{19}$$

$$b_3(r) = \alpha_1\, r + \alpha_2\, r^2 + \alpha_3\, r^3 + \alpha_4\, r^4 + \alpha_5\, r^5. \quad (20)$$

According to these equations and Eq. (10) , radial dependence of B is the same as that of b_0. Appropriate values of α_l and β_l were found for each coefficient b_k entering Eq. (15) by least squares minimization [27].

Finally, for a specific cavity radius at constant pressure, the probability P_n was obtained from Eqs. (9) and Eq. (10) using ρ from Eq. (7) and fitted with the form derived from an information theory based on the first two moments and using a flat default model, namely

$$P_n = exp[\lambda_0 + \lambda_1 n + \lambda_2 n^2]. \quad (21)$$

The curves obtained in this manner for a cavity of 1.8 Å are shown in Fig. 8 (b) at 1 and 8000 atm. These curves practically overlap those obtained fitting the simulation results. At 1 atm, our results compare very well with the information theory results obtained by Hummer et al. [12] using another model potential for water.

Supplementary Material

Parameters entering Eqs. 7-9.

[1] A. Ben-Naim, Statistical Thermodynamics for Chemists and Biochemists, Plenum, New York, 1992.

[2] D. Ben-Amotz, F. O. Raineri, G. Stell, Solvation thermodynamics: Theory and apllications., J. Phys. Chem. B 109 (2005) 6866.

[3] H. Reiss, H. Frish, J. Lebowitz, Statistical mechanics of rigid spheres, J. Chem. Phys. 31 (1959) 369.

[4] F. H. Stillinger, Structure in aqueous solutions of nonpolar solutes from the standpoint of scaled-particle theory, J. Solution Chem. 141 (1973) 197.

[5] F. Floris, M. Selmi, A. Tani, J. Tomasi, Free energy and entropy for inserting cavities in water: Comparison of monte carlo simulation and scaled particle theory results, J. Chem. Phys. 107 (1997) 6353.

[6] H. S. Ashbaugh, L. Pratt, Scaled-particle theory and the length scales of hydrophobicity, Rev. Mod. Phys. 78 (2006) 159.

[7] J. Tomasi, B. Mennucci, R. Cammi, Quantum mechanical continuum solvation models, Chem. Rev. 105 (8) (2005) 2999–3093.

[8] C. Amovilli, C. Filippi, F. M. Floris, Quantum monte carlo formulation of volume polarization in dielectric continuum theory, J. Chem. Phys. 129 (2008) 244106.

[9] C. Amovilli, F. M. Floris, Study of dispersion forces with quantum monte carlo: Toward a continuum model for solvation, J. Phys. Chem. A. 5327 (2015) 244106.

[10] F. M. Floris, Modeling the cavitation free energy, J. Phys. Chem. B 109 (2005) 24061.

[11] T. L. Hill, Statistical Mechanics, Dover Publications, New York, 1956.

[12] G. Hummer, S. Garde, A. E. Garcia, A. Pohorille, L. R. Pratt, An information theory model of hydrophobic interactions, Proc. Natl. Acad. Sci. USA 93 (1996) 8951–8955.

[13] A. Pierotti, A scaled particle theory of aqueous and nonaqueous solutions, Chem. Rev. 76 (1976) 717.

[14] J. Postma, H. Berendson, J. Haak, Thermodynamics of cavity formation in water, Faraday Symp. Chem. Soc. 17 (1982) 55.

[15] J. Rowlinson, B. Widom, Molecular Theory of Capillarity, Clarendon Press, Oxford, 1951.

[16] L. Pratt, A. Pohorille, Theory of hydrophobicity: Transient cavities in molecular liquids, A. Proc. EBSA Workshop on Water-Biomolecular Interactions 43 (1992) 261.

[17] H. S. Ashbaugh, T. M. Truskett, Putting the squeeze on cavities in liquids: Quantifying pressure effects on solvation using simulations and scaled-particle theory, J. Chem. Phys. 134 (2011) 014507.

[18] W. Jorgensen, J. Chandrasekhar, J. Madura, R. Impey, M. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926.

[19] W. Jorgensen, BOSS, version 3.5, Yale University Press, New Haven, CT, 1994.

[20] In order to do this we modified the program BOSS [see Ref. 18].

[21] F. M. Floris, Nonideal effects on the excess volume from small to large cavities in tip4p water., J. Phys. Chem. B 108 (2004) 16244.

[22] A. Kalinichev, Y. Gorbaty, A. Okhulkov, Structure and hydrogen bonding of liquid water at high hydrostatic pressures: Monte carlo npt-ensemble simulations up to 10 kbar, J. Mol. Liq. 82 (1999) 57.

[23] M. S. Moghaddam, H. S. Chan, Pressure and temperature dependence of hydrophobic hydration: Volumetric, compressibility, and thermodynamic signatures., J. Chem. Phys. 126 (2007) 114507.

[24] M. W. Mahoney, W. L. Jorgensen, A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpoalarizable potential functions, J. Chem. Phys. 112 (2000) 8910.

[25] H. Sato, M. Uematsu, K. Watanabe, A. Saul, W. Wagner, A fundamental equation for water covering the range from the melting line to 1273 k at pressures up to 25000 mpa, J. Phys. Chem. Ref. Data 17 (1988) 1439.

[26] J. R. MacDonald, Some simple isothermal equations of state, Rev. Mod. Phys. 38 (1966) 669.

[27] See Supplemental Material Document No. for parameters entering radial functions b_0(r), b_1(r), b_2(r), b_3(r) .

[28] H. Reiss, E. Helfand, H. Frish, J. Lebowitz, Aspects of the statistical thermodynamics of real fluids, J. Chem. Phys. 32 (1960) 119–124.

[29] G. Graziano, Scaled particle theory of the length scale dependence of cavity thermodynamics in different liquids, J. Phys. Chem. B 110 (2006) 11421–11426.

[30] P. Vallery, A. J. Patel, D. Chandler, An improved coarse-grained model of solvation and the hydrophobic effect, J. Chem. Phys. 134 (2011) 074109.