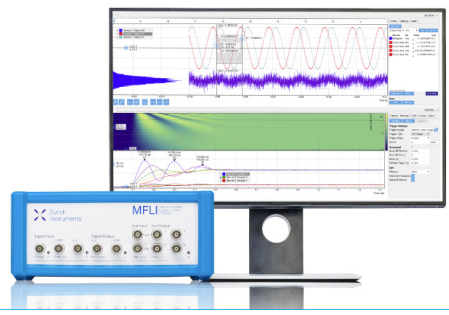
The Journal of Chemical Physics

Calculation of the water-octanol partition coefficient of cholesterol for SPC, TIP3P, and TIP4P water

Cite as: J. Chem. Phys. **149**, 224501 (2018); https://doi.org/10.1063/1.5054056
Submitted: 29 August 2018 . Accepted: 13 November 2018 . Published Online: 11 December 2018

Jorge R. Espinosa, Charlie R. Wand, Carlos Vega, Eduardo Sanz, and Daan Frenkel



AIP Publishing

Calculation of the water-octanol partition coefficient of cholesterol for SPC, TIP3P, and TIP4P water

Jorge R. Espinosa,[1] Charlie R. Wand,[2,3] Carlos Vega,[1] Eduardo Sanz,[1] and Daan Frenkel[2]
[1]*Departamento de Quimica Fisica, Facultad de Ciencias Quimicas, Universidad Complutense de Madrid, 28040 Madrid, Spain*
[2]*Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom*
[3]*School of Chemical Engineering and Analytical Science, University of Manchester, Manchester M13 9PL, United Kingdom*

(Received 29 August 2018; accepted 13 November 2018; published online 11 December 2018)

We present a numerical study of the relative solubility of cholesterol in octanol and water. Our calculations allow us to compare the accuracy of the computed values of the excess chemical potential of cholesterol for several widely used water models (SPC, TIP3P, and TIP4P). We compute the excess solvation free energies by means of a cavity-based method [L. Li *et al.*, J. Chem. Phys. **146**(21), 214110 (2017)] which allows for the calculation of the excess chemical potential of a large molecule in a dense solvent phase. For the calculation of the relative solubility ("partition coefficient," $\log_{10}P_{o/w}$) of cholesterol between octanol and water, we use the OPLS/AA force field in combination with the SPC, TIP3P, and TIP4P water models. For all water models studied, our results reproduce the experimental observation that cholesterol is less soluble in water than in octanol. While the experimental value for the partition coefficient is $\log_{10}P_{o/w}$= 3.7, SPC, TIP3P, and TIP4P give us a value of $\log_{10}P_{o/w}$ = 4.5, 4.6, and 2.9, respectively. Therefore, although the results for the studied water models in combination with the OPLS/AA force field are acceptable, further work to improve the accuracy of current force fields is needed. *Published by AIP Publishing.* https://doi.org/10.1063/1.5054056

I. INTRODUCTION

The solubility of a substance ("solute") in a given solvent denotes the maximum amount of such substance that can be dissolved. Solubility basically depends on the physicochemical properties of the solute and solvent, as well as on the temperature, pressure, and the chemical composition of the solution. Under certain conditions, the equilibrium solubility can be exceeded to reach a metastable supersaturated solution from which the solute can eventually precipitate. Accurate predictions of solubilities are of considerable practical importance. For instance, in the development of new pharmaceutical compounds,[1–4] it is crucial to know both its absolute solubility and relative solubility in hydrophilic-hydrophobic environments representing the blood-cell barrier. For similar reasons, the solubility of compounds is also of interest to the food industry.[5,6] In the petrochemical industry, it is extremely important to be able to predict the solubilities of various organic and inorganic compounds in aqueous solutions, as exceeding this limit might lead to blockages in pipelines.[7–9]

Given the importance of accurate solubility calculations during the two last decades, many efforts have been devoted in developing computational methods for estimating solubilities including direct coexistence simulations[10–12] and the osmotic ensemble[13–19] or making use of thermodynamic relations to calculate the solubility via chemical potential calculations.[20–26] Each of these methods has associated strengths and weaknesses. For example, direct simulation of the two phases of interest is a conceptually simple "brute force" technique, and it requires long simulations of large system sizes even for readily soluble solutes. It is effectively useless to estimate the solubility of sparingly soluble solutes. The osmotic ensemble has previously been successfully employed to calculate the solubility in both a solution and porous solids;[13–19] however, the method is less convenient for sparingly soluble solutes. Thermodynamic relationships rely on the equality of the chemical potential of the two phases in equilibrium and require calculation of the excess chemical potential of the solute in separate simulations of both phases. However, common chemical potential calculation methods such as Widom trial insertion[27] and direct growth[28–31] are sensitive to the choice of the coupling scheme used and may suffer particle-particle overlaps that can cause either low (or zero) insertion probabilities or an integrable singularity, respectively.

Therefore, most solubility calculations have only been applied to compute the solubility of small molecules such as NaCl in water.[10–12,14,20,21,32–37] A proof of how challenging it is to measure solubilities in simulations is that the first reported calculation for an ionic system was the one by Ferrario *et al.* in 2002 for KF.[22] Later in 2007, Sanz and Vega determined also the solubility of NaCl in water,[21] and since then, several groups calculated the solubility limit of NaCl in water for that model, the JC/SPC/E,[38,39] only reaching a consensus between the different methodologies[10,12,14,20,32] almost ten years later.

Moving away from particle-based methods, machine learning[40] and analytical models[41–43] have also been developed; however, these methods lack physical or chemical understanding. To overcome the limitations mentioned above, in 2017, Li *et al.*[23] proposed a new method for computing

solubilities of general and arbitrarily complex solutes with low solubility, as, for instance, organic molecules in water. There the authors computed the solubility of two organic molecules in water, napthalene, and cyclohexane. This new methodology is based on calculating the solubility via computing the absolute chemical potential of the solid and the solution phases in solubility equilibrium. The method is composed of an extended version of the Einstein crystal[44,45] (Einstein molecule can be also used[46,47]) for evaluating the chemical potential of the solid in a molecular crystal, plus a novel cavity-based method, which allows for the calculation of the excess chemical potential of general solutes in the solution. In that sense, the fundamental idea is similar to other previous methodologies.[20,21,48] The significant advantage in the cavity-based method in comparison to other methodologies is that it does not suffer from inaccuracies as growing directly a solute in the solution nor is it only valid for small solutes[48] as it overcomes the integrable singularity due to particle-particle overlaps. Alongside this, it is simple to implement it in standard molecular dynamics packages.[49]

This methodology can also be applied for computing partition coefficients. An early calculation of partition coefficients for alkanes and alcohol solutes in octanol/water mixtures was reported in 2000.[50] In the present work, we use the cavity-based method,[23] a more straightforward technique than the configurational-bias Monte Carlo simulations in the Gibbs ensemble used in 2000,[50] to compute the partition coefficient of cholesterol between water and octanol. The partition coefficient ($P_{I/II}$) is the ratio of concentrations of a compound in a mixture of two phases (I and II) at equilibrium and can be related for sparingly soluble solutes to the difference in solubility of the compound in the two phases. Commonly one of the solvents is water, while the other is a hydrophobic one such as 1-octanol; hence, the partition coefficient measures how hydrophilic or hydrophobic is a particular substance. $\log_{10}P_{o/w}$ is given by

$$\log_{10} P_{I/II} = \log_{10} \frac{[S]_I}{[S]_{II}}, \tag{1}$$

where $[S]_x$ stands for the concentration of the solute (in molarity, the number of moles of solute per liter of solution) in x, where x is the solvent. Here the octanol phase is denoted by I whilst water is denoted by II. Although arbitrary concentration units can be used to define the partition coefficient, the use of molarity is particularly convenient as then the logarithm of the partition coefficient can be determined from the excess chemical potentials, as will be shown below. Experimentally the partition coefficient is measured for two immiscible solvents using a shake-flask containing both solvents from which the relative concentrations can be determined by spectroscopic, chromatographic, or volumetric methods. However, this method can be both expensive and time-consuming and may suffer from accuracy issues for sparingly soluble solutes.[51,52]

Here we relate $\log_{10}P_{o/w}$ to the excess chemical potential, μ_{excess}, of the solute in water and octanol. Note that for estimating the relative solubility between both phases, it is not necessary to know the chemical potential of the solid given that in both cases, it will be the same. From the Gibbs free energy of solvation in the two different phases at a given T and p, it is possible to calculate $\log_{10}P_{o/w}$ according to the following expression:[24,53]

$$\log_{10} P_{o/w} = \frac{\mu^w_{excess} - \mu^o_{excess}}{2.303RT}, \tag{2}$$

where μ^w_{excess} and μ^o_{excess} are the excess chemical potentials of cholesterol in water and octanol respectively, $2.303 \approx \ln 10$, R is the ideal gas constant, and T is the absolute temperature.

We employ the OPLS/AA force field[54] to represent both the cholesterol and octanol molecules in combination with several widely used water models such as the SPC,[55] TIP3P,[56] and TIP4P[56] models. We have selected the OPLS/AA force field because, over the past few years, it has become one of the most widely used force fields for simulating sterols, lipids, and membranes.[57,58] As OPLS/AA has been parameterized for being used in combination with SPC, TIP3P, and TIP4P water models, we have chosen these models for simulating water despite them being less accurate than other models in reproducing the behavior of pure water.[59]

We compare the prediction for the different models with the experimental value $\log_{10}P_{o/w} = 3.7$, calculated via Eq. (1) from the solubility of cholesterol in octanol and in water reported in Refs. 60 and 61, respectively. Predicting solubilities is a non-trivial quality test for a force field since $\log_{10}P_{o/w}$ depends on the partial solubility of a solute in two different phases and also because this property is not typically one of the main targets when developing force fields (including the water models investigated here).

The rest of this paper is organised as follows. In Sec. II, we present the model and simulation details including the cavity-based methodology. In Sec. III, we discuss our findings before summarizing in Sec. IV.

II. MODEL AND SIMULATION DETAILS

A. The cavity-based method

We employ the cavity-based method to calculate the excess chemical potential μ_{excess} of general and arbitrarily complex solutes. In principle, although it is possible to compute μ_{excess} directly by growing a solute directly in the solution without the use of a cavity, such an approach may suffer from inaccuracies[29] due to particle-particle overlap. The use of a pre-grown cavity eliminates such difficulties by the introduction of a softly repulsive potential that creates a cavity in the solution inside which the solute can be inserted. Once the solute is fully grown, then the cavity can be removed.

Thus the total $\Delta G_{solvation}$ is given by

$$\Delta G_{solvation} = \Delta G_{grow} + \Delta G_{insert} + \Delta G_{shrink}, \tag{3}$$

where ΔG_{grow}, ΔG_{insert}, and ΔG_{shrink} are the change in free energy associated with the cavity growth, insertion of the solute, and cavity shrinking, respectively (more details/sketch are described in Refs. 23 and 24). Note that $\Delta G_{solvation}$ is equal to μ_{excess} as we are inserting only one molecule of the solute ($N = 1$). In all cases, we calculate the change in free energy using thermodynamic integration.[44] For computing ΔG_{grow}, we first insert the cavity at the centre of our cubic simulation box and second we gradually increase the interaction of the softly repulsive potential for making the cavity.

In theory, the μ_{excess} calculated is independent of the cavity potential used; however, recently we have found that this is

not always the case due to irreversible nucleation events.[24] However, given the same force fields are employed by Li *et al.*,[23] we chose to use the same cavity potential described by the following equation:

$$U_{cavity}(r_{\alpha-\beta}, \lambda) = A \exp(-r_{\alpha-\beta}/B + \lambda), \tag{4}$$

where λ is the thermodynamic coupling parameter and accounts for the size of the cavity formed, $r_{\alpha-\beta}$ is the distance between the centre of the cavity and a fixed point of a solvent molecule, which for convenience can be chosen as a specific atom of the solvent molecule, and A and B are the fixed parameters with the values of A = 300 kcal/mol and B = 1 Å, respectively. The typical range of λ used in this work has been from −10 to 6. When $\lambda = -10$, the interaction between the cavity potential is negligible and we take as the equivalent of having no cavity, and for $\lambda = 6$, the volume of the cavity is noticeably larger than the volume of a single cholesterol molecule, our solute of interest. The thermodynamic integration expression for computing ΔG_{grow} is

$$\begin{aligned}\Delta G_{grow} &= \int \langle \partial U_{cavity}(r_{\alpha-\beta}, \lambda)/\partial\lambda \rangle_\lambda d\lambda \\ &= \int \langle U_{cavity}(r_{\alpha-\beta}, \lambda) \rangle_\lambda d\lambda,\end{aligned} \tag{5}$$

where $U_{cavity}(r_{\alpha-\beta}, \lambda)$ refers to the interaction energy between the cavity and the surrounding solvent molecules as a function of λ. Once the cavity has been created, the solute is then inserted into it. The insertion of the solute has to be done by fixing the centre of the solute into the centre of the cavity in order to avoid displacements towards the border of the cavity during the activation of the interactions between the solute and the solvent. For computing the free energy work of inserting the solute into the cavity, ΔG_{insert}, we first switch on the van der Waals interactions and then we turn on the Coulombic interactions.

For computing ΔG^{vdW}_{insert}, we use

$$\Delta G^{vdW}_{insert} = \int_0^1 \langle \partial U_{vdW}(\lambda'_{vdW})/\partial\lambda'_{vdW} \rangle_{\lambda'_{vdW}} d\lambda'_{vdW}, \tag{6}$$

where $U_{vdW}(\lambda'_{vdW}) = U_{vdW}\lambda'_{vdW}$ is the van der Waals interaction energy between the solute and the solvent along the path λ'_{vdW} that connects no interaction between them (λ'_{vdW} = 0) to full interaction (λ'_{vdW} = 1). To evaluate the electrostatic contribution ΔG^{Coul}_{insert} between the solute and the solvent, the same way described in Eq. (6) is used but for the Coulombic interactions.

As long as the cavity size is large enough for the host solute, vdW and Coulombic interactions can be switched on in a single step (λ'_{vdW} or λ'_{coul} from 0 → 1 directly). We have verified that activating separately both interactions in a single step or in several steps (0 → 0.1, 0.1 → 0.2, . . ., and 0.9 → 1), the results, when the cavity size is large enough, are the same within the uncertainty.

Finally, the last step is to shrink the cavity until the solute molecule is completely solvated. This free energy work can be computed analogously to that in Eq. (5).

B. Model

The force field used in this work for modeling cholesterol and octanol was the OPLS/AA.[54] In this force field, the atoms interact via the Lennard-Jones potential plus a Coulombic interaction using partial charges. 1,4 intramolecular interactions are also considered. For the case of water, we used three different water models, and all of them are rigid and non-polarizable, SPC,[55] TIP3P[56], and TIP4P.[56] Both SPC and TIP3P are 3-site water models, and TIP4P is a 4-site one having the negative charge of the oxygen in the bisector of the angle formed by the oxygen and the two hydrogens. The cross interactions between cholesterol and octanol and cholesterol and water are described by the geometric combination rule $\sigma_{ij} = \sqrt{\sigma_i\sigma_j}$ and $\epsilon_{ij} = \sqrt{\epsilon_i\epsilon_j}$, where ϵ_{ij} is the depth of the potential well for the Lennard-Jones interactions between atoms of types i and j and σ_{ij} is the same but for the atomic diameter. The simulations were carried out using LAMMPS[49] in the isobaric-isothermal ensemble (NpT). The pressure was fixed using the Nose-Hoover[62] barostat with a relaxation time of 0.5 ps, and temperature was fixed by using the Nose-Hoover[63] thermostat with a relaxation time of 0.5 ps. For integrating the equations of motion, we used the velocity-Verlet integrator[64] with a time step of 0.3 fs. We used particle-particle-particle-mesh summations[65,66] to deal with the electrostatic interactions. The cut-off radius for both dispersive interactions and the real part of the electrostatic interactions were 14 Å. Long-range Lennard-Jones tail corrections were included. The SHAKE algorithm[67] was used to constrain the O-H bond length in water and the angle of the molecule. All our simulations were run at T = 298 K and p = 1 bar, with N = 512 molecules in the case of octanol (being 0.012 mol/l the concentration of cholesterol when one molecule is inserted) and N = 4096 molecules for water (being 0.013 mol/l the concentration of cholesterol when one molecule of cholesterol is inserted). The experimental solubilities of cholesterol in water and in octanol are $4.7{\cdot}10^{-6}$ mol/l[61] and 0.022 mol/l,[60] respectively. We shall assume that the excess chemical potential of cholesterol does not change much with concentration as it is a sparingly soluble solute.

III. RESULTS

A. Computing the partition coefficient of cholesterol between octanol and SPC water

In this section, we illustrate step-by-step how $\log_{10}P_{o/w}$ has been computed for the case of SPC water.[55] For the other water models studied in this work, the same protocol was followed.

First, we have computed the free energy work needed (ΔG_{grow}) for creating a cavity as a function of its size in both solvents water and octanol. In Fig. 1(a), we show the value for the interaction between the cavity potential and the solvent molecules $U_{cavity}(\lambda)$ (octanol is represented by red squares, and SPC water is represented by blue circles) as a function of λ, the parameter which controls the size of the cavity. The value of the integral for both curves evaluated via Eq. (5) is also included in Fig. 1(b), a dashed red line for octanol and a dashed blue one for SPC water. As can be extracted from

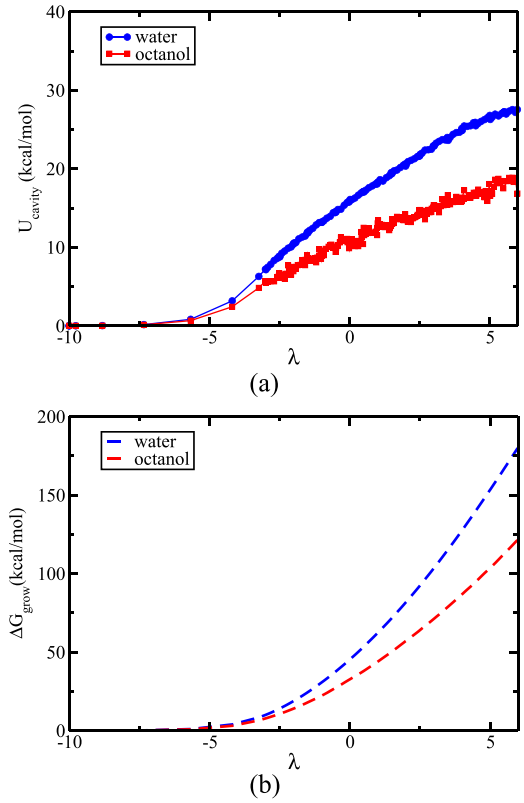


FIG. 1. (a) Potential energy interaction (U_{cavity}) between the cavity and solvent molecules as a function of λ; red squares account for octanol and blue circles account for SPC water. (b) Free energy change associated with creating the cavity (ΔG_{grow}) in each solvent; the red dashed curve stands for octanol, and the blue one stands for water.

this plot 1(b), ΔG_{grow} for creating a cavity of the same size in octanol costs less energy than that in water. This result is consistent with the fact that the liquid-vapour interfacial free energy in octanol is much lower than that in water.[68,69]

Once the cavity formed in the fluid is large enough to accommodate one molecule of cholesterol, the second step is to introduce the molecule inside it and to switch on the potential interaction between the cholesterol molecule and the solvent. To check that our results are unlikely to be affected by finite-size effects, we evaluated a density profile from the center of the cavity to the half of the box side. We observe that the liquid density beyond the cavity is the same as that of the bulk for the liquid under the same conditions.

First we have to attach the closest atom to the center of mass in cholesterol to the center of the cavity, thus avoiding the cholesterol to move towards the borders of the cavity. For switching on the interactions, as commented before, it is advisable first to activate the van der Waals interaction, and once these are at full strength, then do the same for the electrostatic contribution. When the size of the cavity is big enough, the activation of such interactions can be done in one single step or in several steps. We have obtained similar results within the uncertainty when switching on the dispersive/electrostatic interactions in one or in several steps (e.g., 5 steps) for the largest λ values, from 4.5 to 6.

In Fig. 2, we show the values for ΔG_{insert} for the van der Waals, the electrostatic contribution, and the sum of both interactions between cholesterol and the surrounding solvent

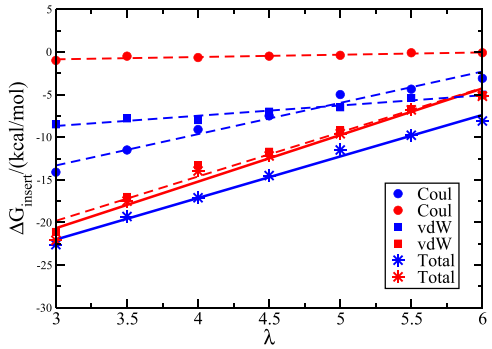


FIG. 2. ΔG_{insert}^{vdW}, ΔG_{insert}^{Coul}, and ΔG_{insert} (total) as a function of λ defined in Eq. (4) (and not of λ' used to determine ΔG_{insert}) for the insertion of cholesterol in SPC water (blue symbols) and in octanol (red ones).

molecules, as a function of λ. Blue symbols correspond to the interaction between cholesterol and SPC water, and red ones correspond to that between cholesterol and octanol. Circles represent the energy for the electrostatic contribution, squares represent that for the van der Waals one, and stars represent the sum of both contributions. As can be seen from Fig. 2, the electrostatic interaction between cholesterol and octanol is much weaker than the one between cholesterol and water. That is not surprising since the partial charges in the octanol molecule are much lower than those of water. On the other hand, the dispersive interactions between cholesterol and octanol are larger than the ones with water. Although globally the sum of both interactions in both solvents follows a similar trend as a function of λ, ΔG_{insert} in water is 2-3 kcal/mol more favorable than that in octanol.

Finally the third step required for computing $\Delta\mu_{excess}$ is to evaluate the free energy change associated with shrinking the cavity ΔG_{shrink} to recover the final state with the molecule of cholesterol fully solvated in the solution. The results obtained for ΔG_{shrink} are shown in Fig. 3 where blue circles stand for the energy between the cavity potential and water U_{cavity}, red squares are the same but for octanol, and dashed lines ΔG_{shrink} are for each solvent. For octanol, the free energy required to shrink the cavity is lower than that in water, as it was for growing it.

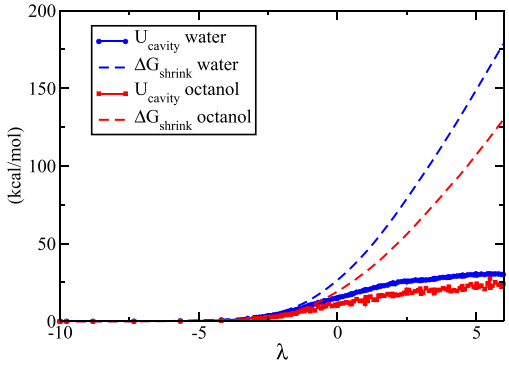


FIG. 3. Potential energy of interaction between the cavity and solvent molecules U_{cavity} as a function of λ while shrinking the cavity. Red squares represent the octanol data, and blue circles are those for SPC water. The free energy change associated with shrinking the cavity once the interactions between cholesterol and each solvent are fully activated is represented by dashed curves using the same code color. Both magnitudes U_{cavity} and ΔG_{shrink} are in *kcal/mol* as indicated in the y axis.

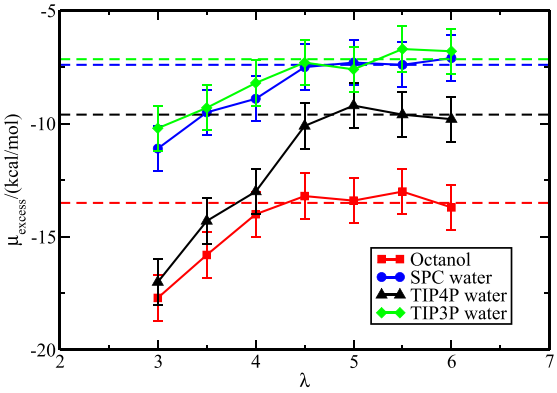


FIG. 4. Excess free energy of solvation of cholesterol as a function of λ in octanol and in water for the three different water models studied. Meaning of the symbols is indicated in the legend.

Having evaluated ΔG_{grow}, ΔG_{insert}, and ΔG_{shrink}, by means of Eq. (3), we can evaluate μ_{excess} for solvating cholesterol either in water or in octanol. The results for μ_{excess} as a function of λ are shown in Fig. 4, for all the studied water models and octanol. Blue circles represent the values for SPC water, and red squares represent those for octanol (green diamonds for TIP3P and black up triangles for TIP4P models). When the size of the cavity is too small to accommodate the molecule of cholesterol, the values of μ_{excess} depend on λ, but from $\lambda = 4.5$ to $\lambda = 6$, the value of the free energy of solvation remains approximately constant within the uncertainty which implies that the cavity size is large enough for hosting the solute. As can be seen from Fig. 4, μ_{excess} for all water models is higher than that for octanol, which clearly indicates that cholesterol will be more soluble in octanol than in water, no matter which model one chooses. It is interesting to note that whilst ΔG_{insert} was lower for water than for octanol, the overall μ_{excess} of cholesterol in octanol was lower than that in water. This is because the cycle of forming and removing the cavity requires less reversible work in the case of octanol than in water, which shows the importance of considering all the free energy contributions when calculating μ_{excess} and not only ΔG_{insert}.

With the obtained values of μ_{excess} for SPC water and octanol, the partition coefficient can be calculated, giving a value of $\log P_{o/w} = 4.5$ in acceptable agreement with the experimental value[60,61] of $\log_{10} P_{o/w} = 3.7$.

B. Comparison of $\log_{10} P_{o/w}$ for SPC, TIP3P, and TIP4P water models

Following the procedure explained in Sec. III A, we have estimated $\log_{10} P_{o/w}$ for the TIP3P[56] and the TIP4P[56] water models (see Fig. 4 and Tables I and II).

The estimated partition coefficients for TIP3P and TIP4P were found to be 4.6 and 2.9, respectively. Notice that the chemical potential of cholesterol in TIP3P and SPC water is about 2 kcal/mol higher than that in TIP4P, and this explains the differences in the partition coefficient. Therefore, both 3-site models SPC and TIP3P overestimate the partition coefficient in comparison to the experimental one,[60,61] while TIP4P underestimates it by an amount of similar magnitude.

TABLE I. Excess free energy of solvation of cholesterol in octanol (μ^o_{excess}), in water μ^w_{excess} in kcal/mol, and its partition coefficient ($\log_{10} P_{o/w}$) in both solvents, for the OPLS/AA force field combined with SPC, TIP3P, and TIP4P. Experimental results calculated from Refs. 60 and 61 are also included.

System	μ^o_{excess}	μ^w_{excess}	$\log_{10} P_{o/w}$
OPLS/AA-SPC	−13.5	−7.4	4.5
OPLS/AA-TIP3P	−13.5	−7.2	4.6
OPLS/AA-TIP4P	−13.5	−9.6	2.9
Experimental	...	...	3.7

Although, TIP4P is a more accurate water model than SPC and TIP3P in describing the behavior of pure water,[59] in this case, for predicting the partition coefficient in combination with the OPLS/AA force field, its estimation of $\log_{10} P_{o/w}$ is similar in accuracy than those of the others.

From Table II, we can evaluate the ratio between ΔG_{grow} and the liquid-vapour interfacial free energy (γ_{lv}) for these models at the same conditions, reported in Ref. 70, which are 54.7 mJ/m^2 for SPC, 52.3 mJ/m^2 for TIP3P, and 59.0 mJ/m^2 for TIP4P. In all cases, we obtain a constant value of 2.7-2.8 indicating that the free energy work computed for creating the cavity inside the liquid is consistent with the previous results of (γ_{lv})[70] since (γ_{lv}) = $\Delta G_{grow}/A_{cavity}$, where A_{cavity} means the exposed area of the cavity to the fluid. Also observing term by term the different contributions of Eq. (3) in Table II (where $\lambda = 5$ was used because the cavity size is large enough for hosting a cholesterol molecule and μ^w_{excess} is converged; see Fig. 4), one can notice that ΔG_{insert} for TIP4P is around 4 kcal/mol lower than that for SPC and TIP3P. This difference is partially compensated when taking into account the free energy work of growing and shrinking the cavity ($\Delta G_{grow} + \Delta G_{shrink}$), which for TIP4P is about 2 kcal/mol higher than for the other models. Therefore, taking into account the sum of the three contributions, we obtain that for the TIP4P model, μ^w_{excess} is around 2 kcal/mol lower than for the other ones, which makes its partition coefficient also lower.

A possible explanation why TIP4P gives a lower $\log_{10} P_{o/w}$ than the other models is related to its partial charges. For TIP4P, the negative charge is −1.04 e, and for TIP3P and SPC, it is −0.834 and −0.82 e, respectively. The higher charges in TIP4P favor the electrostatic cross interaction between cholesterol and water, and this might lead to an overestimate of its solubility in water.

These results stress that models which are successful in describing the behavior of pure substances need further tuning when applying them to mixtures.

In this context, it might be interesting to analyze the effect of using different types of combination rules (for instance, the

TABLE II. Different contributions of Eq. (3) in kcal/mol for the three employed water models at $\lambda = 5$. The sum of them gives μ^w_{excess}.

Model	ΔG_{grow}	ΔG_{insert}	ΔG_{shrink}
SPC	153.4	−11.5	−149.0
TIP3P	146.5	−12.0	−142.0
TIP4P	157.7	−15.8	−151.1

Lorentz-Berthelot rule[71]) on the partition coefficient. Clearly, more work is needed to improve the accuracy of the current force fields to obtain reliable predictions of the properties of solutions, such as solvation free energies. Here we have focused on the partition coefficient which eliminates the need to consider the solid phase. Of course, the absolute solubility is also of importance. Here, we did not include absolute solubility calculations because the accuracy of such calculations is likely to suffer from the fact that the OPLS/AA (and most other commonly used force fields) has not been parameterized for the solid phase. Moreover, in the case of cholesterol, there exist multiple crystalline polymorphs depending on the hydration of the crystal. From a more practical perspective, the relevance of absolute solubility calculations of cholesterol seems limited as to our knowledge crystalline cholesterol does not form in biological systems. To be more precise, the highly relevant "cholesterol deposits" in atherosclerosis are made of a mixture of cholesterol and lipids. Predicting the conditions under which such deposits would form would require a study of the complete cholesterol/lipid/water phase diagram, which is beyond the scope of the present work.

IV. CONCLUSIONS

In this work, we have applied the recently developed cavity-based method to evaluate the partition coefficient of cholesterol between octanol and water. Using this method, we have evaluated the partition coefficient of cholesterol in octanol and water using the OPLS/AA force field (for modeling cholesterol and octanol) in combination with three different models of water, SPC, TIP3P, and TIP4P. For the three studied water models, we reproduce the fact that cholesterol is more soluble in octanol than in water. When using SPC and TIP3P, the partition coefficients predicted are higher, $\log_{10}P_{o/w} = 4.5$ and 4.6, respectively, than the experimental one $\log_{10}P_{o/w} = 3.7$ whilst when OPLS/AA is used with TIP4P, it gives a lower $\log_{10}P_{o/w} = 2.9$. These results might turn out a bit unexpected since in overall TIP4P is a somewhat better model in describing pure water and one should expect a better performance of a force field when combining it with a better model of water. As shown in this paper, this is not the case since the predictions based on the three water models studied deviated about equally from the experimental $P_{o/w}$ values. We hypothesize that the underestimation in TIP4P is because the higher partial charges than in either SPC or TIP3P are enhancing the electrostatic stabilization of cholesterol in water. By contrast, the lower partial charges in SPC and TIP3P lead to an overestimate of $\log_{10}P_{o/w}$.

In summary, the cavity-based method has allowed us to carry out a rigorous validation test for the applicability of force fields to the prediction of partition coefficients. Our results show that although the performance of the models is fair, our simulations suggest that current force fields yield predictions for $\frac{[S]_o}{[S]_w}$ that may be off by an order of magnitude. In the case of cholesterol, the observed discrepancy is important because octanol is often used as a model fluid that is representative for the medium inside a lipid membrane. Hence, if simulations predict the wrong partition coefficient between octanol and water, one should expect that the same models will not perform well when predicting how strongly cholesterol is absorbed in bio-membranes.

Clearly, further work is needed to improve the predictive capability of force fields for solubility predictions. Such ability to predict solubilities is particularly important in the case of sparingly soluble solutes, in which case, experiments may be very challenging.

ACKNOWLEDGMENTS

This work was funded by Grant No. FIS2016/78117-P of the MEC. J. R. Espinosa acknowledges financial support from the FPI Grant No. BES-2014-067625. C.R.W. and D.F. gratefully acknowledge the generous funding in the early stages of this work from BP Plc through the International Centre for Advanced Materials (ICAM). All the simulations carried out within this work were conducted using the HPC resources from the Department of Chemistry, Cambridge.

[1] G. L. Amidon, H. Lennernas, V. P. Shah, and J. R. Crison, "A theoretical basis for a biopharmaceutic drug classification: The correlation of *in vitro* drug product dissolution and *in vivo* bioavailability," Pharm. Res. **12**, 413 (1995).
[2] W. L. Jorgensen and E. M. Duffy, "Prediction of drug solubility from structure," Adv. Drug Delivery Rev. **54**, 355–366 (2002).
[3] W. L. Jorgensen and E. M. Duffy, "Prediction of drug solubility from Monte Carlo simulations," Bioorg. Med. Chem. Lett. **10**, 1155 (2000).
[4] M. H. Abraham and J. Le, "The correlation and prediction of the solubility of compounds in water using an amended solvation energy relationship," J. Pharm. Sci. **88**(9), 868–880 (1999).
[5] M. Esmaili, S. M. Ghaffari, Z. Moosavi-Movahedi, M. S. Atri, A. Sharifizadeh, M. Farhadi, R. Yousefi, J.-M. Chobert, T. Haertlé, and A. A. Moosavi-Movahedi, "Beta casein-micelle as a nano vehicle for solubility enhancement of curcumin; food industry application," LWT–Food Sci. Technol. **44**(10), 2166–2172 (2011).
[6] J. Gajdoš, K. Galić, Ž. Kurtanjek, and N. Ciković, "Gas permeability and DSC characteristics of polymers used in food packaging," Polym. Test. **20**(1), 49–57 (2000).
[7] K. U. G. Raju and G. Atkinson, "Thermodynamics of 'scale' mineral solubitities. III. Calcium sulfate in aqueous NaCl," J. Chem. Eng. Data **35**, 361 (1990).
[8] T. Chen, A. Neville, and M. Yuan, "Calcium carbonate scale formation-assessing the initial stages of precipitation and deposition," J. Pet. Sci. Eng. **46**, 185–194 (2005).
[9] O. C. Mullins, A. E. Pomerantz, J. Y. Zuo, and C. Dong, "Downhole fluid analysis and asphaltene science for petroleum reservoir evaluation," Annu. Rev. Chem. Biomol. Eng. **5**, 325–345 (2014).
[10] J. R. Espinosa, J. M. Young, H. Jiang, D. Gupta, C. Vega, E. Sanz, P. G. Debenedetti, and A. Z. Panagiotopoulos, "On the calculation of solubilities via direct coexistence simulations: Investigation of NaCl aqueous solutions and Lennard-Jones binary mixtures," J. Chem. Phys. **145**(15), 154111 (2016).
[11] H. M. Manzanilla-Granados, H. Saint-Martín, R. Fuentes-Azcatl, and J. Alejandre, "Direct coexistence methods to determine the solubility of salts in water from numerical simulations. Test case NaCl," J. Phys. Chem. B **119**(26), 8389–8396 (2015).
[12] A. L. Benavides, J. L. Aragones, and C. Vega, "Consensus on the solubility of NaCl in water from computer simulations using the chemical potential route," J. Chem. Phys. **144**, 124504 (2016).
[13] M. Lisal, W. R. Smith, and J. Kolafa, "Molecular simulations of aqueous electrolyte solubility. I. The expanded ensemble osmotic molecular dynamics method for the solution phase," J. Phys. Chem. B **109**, 12956 (2005).
[14] F. Moucka, M. Lisal, J. Skvor, J. Jirsak, I. Nezbeda, and W. R. Smith, "Molecular simulation of aqueous electrolyte solubility. II. Osmotic ensemble Monte Carlo methodology for free energy and solubility calculations and application to NaCl," J. Phys. Chem. B **115**(24), 7849–7861 (2011).

[15]I. Nezbeda, F. Moučka, and W. R. Smith, "Recent progress in molecular simulation of aqueous electrolytes: Force fields, chemical potentials and solubility," Mol. Phys. **114**(11), 1665–1690 (2016).
[16]F. Moučka, J. Kolafa, M. Lísal, and W. R. Smith, "Chemical potentials of alkaline earth metal halide aqueous electrolytes and solubility of their hydrates by molecular simulation: Application to $CaCl_2$, antarcticite, and sinjarite," J. Chem. Phys. **148**(22), 222832 (2018).
[17]A. Ghoufi and G. Maurin, "Hybrid Monte Carlo simulations combined with a phase mixture model to predict the structural transitions of a porous metal-organic framework material upon adsorption of guest molecules," J. Phys. Chem. C **114**(14), 6496–6502 (2010).
[18]L. J. Dunne and G. Manos, "Statistical mechanics of binary mixture adsorption in metal-organic frameworks in the osmotic ensemble," Philos. Trans. R. Soc., A **376**, 20170151 (2017).
[19]F.-X. Coudert, "The osmotic framework adsorbed solution theory: Predicting mixture coadsorption in flexible nanoporous materials," Phys. Chem. Chem. Phys. **12**(36), 10904–10913 (2010).
[20]Z. Mester and A. Z. Panagiotopoulos, "Mean ionic activity coefficients in aqueous NaCl solutions from molecular dynamics simulations," J. Chem. Phys. **142**(4), 044507 (2015).
[21]E. Sanz and C. Vega, "Solubility of KF and NaCl in water by molecular simulation," J. Chem. Phys. **126**, 014507 (2007).
[22]M. Ferrario, G. Ciccotti, E. Spohr, T. Cartailler, and P. Turq, "Solubility of KF in water by molecular dynamics using the Kirkwood integration method," J. Chem. Phys. **117**, 4947 (2002).
[23]L. Li, T. Totton, and D. Frenkel, "Computational methodology for solubility prediction: Application to the sparingly soluble solutes," J. Chem. Phys. **146**(21), 214110 (2017).
[24]C. Wand, T. Totton, and D. Frenkel, "Addressing hysteresis and slow equilibration issues in cavity-based calculation of chemical potentials," J. Chem. Phys. **149**(1), 014105 (2018).
[25]D. L. Mobley, C. I. Bayly, M. D. Coope, M. R. Shir, and K. A. Di, "Small molecule hydration free energies in explicit solvent: An extensive test of fixed-charge atomistic simulations," J. Chem. Theory Comput. **5**, 350–358 (2009).
[26]G. Koenig, M. T. Reetz, and W. Thiel, "1-butanol as a solvent for efficient extraction of polar compounds from aqueous medium: Theoretical and practical aspects," J. Phys. Chem. B **122**, 6975–6988 (2018).
[27]B. Widom, "Some topics in the theory of fluids," J. Chem. Phys. **39**, 2808 (1963).
[28]T. Steinbrecher, D. L. Mobley, and D. A. Case, "Nonlinear scaling schemes for Lennard-Jones interactions in free energy calculations," J. Chem. Phys. **127**(21), 214108 (2007).
[29]T. Simonson, "Free energy of particle insertion," Mol. Phys. **80**, 441 (1993).
[30]T. C. Beutler, A. E. Mark, R. C. van Schaik, P. R. Gerber, and W. F. Van Gunsteren, "Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations," Chem. Phys. Lett. **222**(6), 529–539 (1994).
[31]M. Zacharias, T. Straatsma, and J. McCammon, "Separation-shifted scaling, a new scaling method for Lennard-Jones interactions in thermodynamic integration," J. Chem. Phys. **100**(12), 9025–9031 (1994).
[32]J. Kolafa, "Solubility of NaCl in water and its melting point by molecular dynamics in the slab geometry and a new BK3-compatible force field," J. Chem. Phys. **145**(20), 204509 (2016).
[33]H. Jiang, Z. Mester, O. A. Moultos, I. G. Economou, and A. Z. Panagiotopoulos, "Thermodynamic and transport properties of H_2O + NaCl from polarizable force fields," J. Chem. Theory Comput. **11**(8), 3802–3810 (2015).
[34]Z. Mester and A. Z. Panagiotopoulos, "Temperature-dependent solubilities and mean ionic activity coefficients of alkali halides in water from molecular dynamics simulations," J. Chem. Phys. **143**(4), 044505 (2015).
[35]F. Moucka, I. Nezbeda, and W. R. Smith, "Chemical potentials, activity coefficients, and solubility in aqueous NaCl solutions: Prediction by polarizable force fields," J. Chem. Theory Comput. **11**(4), 1756–1764 (2015).
[36]J. L. Aragones, E. Sanz, and C. Vega, "Solubility of NaCl in water by molecular simulation revisited," J. Chem. Phys. **136**(24), 244508 (2012).
[37]A. S. Paluch, S. Jayaraman, J. K. Shah, and E. J. Maginn, "A method for computing the solubility limit of solids: Application to sodium chloride in water and alcohols," J. Chem. Phys. **133**(12), 124504 (2010).
[38]I. S. Joung and T. E. Cheatham, "Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations," J. Phys. Chem. B **112**(30), 9020–9041 (2008).
[39]H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, "The missing term in effective pair potentials," J. Phys. Chem. **91**, 6269 (1987).
[40]N. Habibi, S. Z. M. Hashim, A. Norouzi, and M. R. Samian, "A review of machine learning methods to predict the solubility of overexpressed recombinant proteins in Escherichia coli," BMC Bioinf. **15**(1), 134 (2014).
[41]H. K. Hansen, C. Riverol, and W. E. Acree, "Solubilities of anthracene, fluoranthene and pyrene in organic solvents: Comparison of calculated values using UNIFAC and modified UNIFAC (Dortmund) models with experimental data and values using the mobile order theory," Can. J. Chem. Eng. **78**(6), 1168–1174 (2000).
[42]A. Jouyban, A. Shayanfar, and W. E. Acree, "Solubility prediction of polycyclic aromatic hydrocarbons in non-aqueous solvent mixtures," Fluid Phase Equilib. **293**(1), 47–58 (2010).
[43]A. Shayanfar, S. H. Eghrary, F. Sardari, W. E. Acree, Jr., and A. Jouyban, "Solubility of anthracene and phenanthrene in ethanol +2, 2, 4-trimethylpentane mixtures at different temperatures," J. Chem. Eng. Data **56**(5), 2290–2294 (2011).
[44]D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic Press, London, 1996).
[45]D. Frenkel and A. J. C. Ladd, "New Monte Carlo method to compute the free energy of arbitrary solids. Application to the fcc and hcp phases of hard spheres," J. Chem. Phys. **81**, 3188 (1984).
[46]C. Vega, E. Sanz, J. L. F. Abascal, and E. G. Noya, "Determination of phase diagrams via computer simulation: Methodology and applications to water, electrolytes and proteins," J. Phys.: Condens. Matter **20**(15), 153101 (2008).
[47]J. L. Aragones, E. Sanz, C. Valeriani, and C. Vega, "Calculation of the melting point of alkali halides by means of computer simulations," J. Chem. Phys. **137**(10), 104507 (2012).
[48]C. H. Bennett, "Efficient estimation of free energy differences from Monte Carlo data," J. Comput. Phys. **22**, 245 (1976).
[49]S. Plimpton, J. Comput. Phys. **117**, 1 (1995).
[50]B. Chen and J. I. Siepmann, "Partitioning of alkane and alcohol solutes between water and (dry or wet) 1-octanol," J. Am. Chem. Soc. **122**(27), 6464–6467 (2000).
[51]E. Baka, J. E. A. Comer, and K. Takács-Novák, "Study of equilibrium solubility measurement by saturation shake-flask method using hydrochlorothiazide as model compound," J. Pharm. Biomed. Anal. **46**(2), 335–341 (2008).
[52]J. Alsenz and M. Kansy, "High throughput solubility measurement in drug discovery and development," Adv. Drug Delivery Rev. **59**(7), 546–567 (2007).
[53]N. M. Garrido, A. J. Queimada, M. Jorge, E. A. Macedo, and I. G. Economou, "1-octanol/water partition coefficients of n-alkanes from molecular simulations of absolute solvation free energies," J. Chem. Theory Comput. **5**(9), 2436–2446 (2009).
[54]W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, "Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids," J. Am. Chem. Soc. **118**, 11225 (1996).
[55]H. Berendsen, J. Postma, W. F. van Gunsteren, and J. Hermans, *Intermolecular Forces*, edited by B. Pullman (Reidel, Dorderecht, 1982), p. 331.
[56]W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," J. Chem. Phys. **79**, 926 (1983).
[57]W. Kuliga, M. Pasenkiewicz-Gierulab, and T. Rog, "Cis and trans unsaturated phosphatidylcholine bilayers: A molecular dynamics simulation study," Chem. Phys. Lipids **195**, 12–20 (2016).
[58]W. Kulig, H. Mikkolainen, A. Olyska, P. Jurkiewicz, L. Cwiklik, M. Hof, I. Vattulainen, P. Jungwirth, and T. Rog, "Bobbing of oxysterols: Molecular mechanism for translocation of tail-oxidized sterols through biological membranes," J. Phys. Chem. Lett. **9**(5), 1118–1123 (2018).
[59]C. Vega and J. L. F. Abascal, "Simulating water with rigid non-polarizable models: A general perspective," Phys. Chem. Chem. Phys. **13**, 19663–19688 (2011).
[60]S. Baluja and K. V. Chavda, "Prediction of solubility parameters of cholesterol in some organic solvents," Int. J. Pharm., Chem. Biol. Sci. **5**, 171–176 (2015), available at http://www.ijpcbs.com/files/volume5-1-2015/19.pdf.
[61]M. E. Haberland and J. A. Reynolds, "Self-association of cholesterol in aqueous solution," Proc. Natl. Acad. Sci. U. S. A. **70**(8), 2313–2316 (1973).
[62]W. G. Hoover, "Constant-pressure equations of motion," Phys. Rev. A **34**, 2499–2500 (1986).

[63]W. G. Hoover, "Canonical dynamics: Equilibrium phase-space distributions," Phys. Rev. A **31**, 1695 (1985).
[64]L. Verlet, "Computer experiments on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules," Phys. Rev. **159**, 98 (1967).
[65]P. P. Ewald, Ann. Phys. **369**, 253 (1921).
[66]R. W. Hockney and J. W. Eastwood, *Computer Simulation Using Particles* (McGraw-Hill, New York, 1981).
[67]J. P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen, "Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes," J. Comput. Phys. **23**, 327 (1977).
[68]A. W. Adamson and A. P. Gast, *Physical Chemistry of Surfaces* (Wiley-Interscience, New York, 1997).
[69]J. Hruby, V. Vins, R. Mares, J. Hykl, and J. Kalová, "Surface tension of supercooled water: No inflection point down to −25° C," J. Phys. Chem. Lett. **5**(3), 425–428 (2014).
[70]C. Vega and E. de Miguel, "Surface tension of the most popular models of water by using the test-area simulation method," J. Chem. Phys. **126**, 154707 (2007).
[71]H. A. Lorentz, "Ueber die anwendung des satzes vom virial in der kinetischen theorie der gase," Ann. Phys. **248**(1), 127–136 (1881).