PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

Journal Name

ARTICLE TYPE

Cite this: DOI: 10.1039/xxxxxxxxxx

On the positional and orientational order of water and methanol around indole: a study on the microscopic origin of solubility

Andres Henao,[a,b] Andrew J. Johnston,[c] Elvira Guàrdia,[b] Sylvia E. McLain[†c] and Luis Carlos Pardo[*a]

Received Date
Accepted Date

DOI: 10.1039/xxxxxxxxxx

www.rsc.org/journalname

Although they are both highly polar liquids, there are a number of compounds, such as many pharmaceuticals, which show vastly different solubilities in methanol compared with water. From theories of the hydrophobic effect, it might be predicted that this enhanced solubility is due to association between drugs and the less polar -CH_3 groups on methanol. In this work, detailed analysis on the atomic structural interactions between water, methanol and the small molecule indole - which is a precursor to many drugs and is sparingly soluble in water yet highly soluble in methanol - reveal that indole preferentially interacts with both water and methanol via electrostatic interactions rather than with direction interactions to the -CH_3 groups. The presence of methanol hydrogen bonds with π electrons of the benzene ring of indole can explain the increase in solubility of indole in methanol relative to water. In addition, the excess entropy calculations performed here suggest that this solvation is enthalpically rather than entropically driven.

1 Introduction

There is a significant lack of knowledge concerning how solvation of biomolecules affects their subsequent structure and thereby function. Atomic scale interactions between water and other liquids with small molecules are complex and difficult to probe yielding the solution phase the least studied milieu, despite its importance in biology, chemistry and pharmacology. Biological molecules are quite often amphiphilic in nature, where both hydrophobic and hydrophilic motifs are present in close proximity to one another - such as for proteins and lipids - however there is still relatively little known about the interplay between hydrophobic/hydrophilic interactions that naturally occur in physiological solutions and biomolecules. This is especially evident for pharmacologically active molecules as there is no clear consensus regarding the structural features which facilitate either their solubility or to penetrate important amphiphilic environments such as that presented by the blood brain barrier (BBB).[2]

Indole is a small, aromatic compound which has similar molecular motifs to a range of drugs which are able to successfully cross the BBB.[3] Interestingly, similar to many drugs,[4,5] indole is sparingly soluble in water (in a relative proportion 1:3400 molecules of water per indole), but is highly soluble in methanol (relative proportion 1:3). Moreover, the addition of very small amounts of methanol to an indole/water mixture greatly enhances the solubility of indole in the mixture. As an example, the addition of only 30 methanol molecules per indole, increases water solubility 3000 times. Methanol must thus be somehow acting actively on the first hydration shell around indole, causing its solubility in water to increase.

The aim of this work is twofold. Firstly, this work is focused on providing an accurate description of the molecular ordering of water and methanol around indole in order to understand how

[a] Grup de Caracterització de Materials, Departament de Física, ETSEIB, Universitat Politècnica de Catalunya, Diagonal 647, E-08028 Barcelona, Catalonia, Spain
[b] Grup de Simulació per Ordinador en Matèria Condensada, Departament de Física, B4-B5 Campus Nord, Universitat Politècnica de Catalunya, E-08034 Barcelona, Catalonia, Spain
[c] Department of Biochemistry, University of Oxford, South Parks Road, OX1 3QU, UK
[†] email: sylvia.mclain@bioch.ox.ac.uk
[*] email: luis.carlos.pardo@upc.edu

† Electronic Supplementary Information (ESI) available: Comparison between neutron data and molecular dynamics. See DOI: 10.1039/b000000x/

these solvents act to keep indole in solution. This is important in order to ascertain what role the ordering of water and methanol plays in the huge differences of solubility in these two solvents and can guide future investigations of drug solubility: a major issue to design bioavailable drugs.

Secondly, in order to understand these methanol/water solutions and their solvation properties, a series of methods have been recently developed to study the local ordering present in disordered phases and have been implemented in the ANGULA[6] software package (which is freely available for download). These techniques can be applied to analyze snapshots produced by any computational method either theoretically driven as it is the case of Molecular Dynamics (MD) or experimentally driven as it is in the case of EPSR (Empirical Potential Structure Refinement)[7] or RMC (Reverse Monte Carlo).[8]

In the present case, the local ordering of water and methanol molecules around indole has been investigated using MD, both in an infinite solution of the binary mixtures indole+water and indole+methanol, and also in the experimentally available ternary mixture indole:water:methanol at a relative proportion of 1:29:30, respectively. This ratio of solvents has been chosen because it allows its comparison with previous neutron diffraction experiments where neither the solute (indole) nor the solvents have negligible contributions to the total diffraction pattern.[9] In addition, MD simulations have been performed on indole at 'infinite dilution', in both methanol and water separately. Measurement of indole as a single molecule in these solutions allows the local ordering of solvents around indole to be investigated without the complicating effects of indole-indole interactions. Further, a comparison between these binary systems and the highly soluble ternary mixtures of indole+methanol+water, allows for specific changes to the solvation of indole within a mixture to be directly assessed.

2 Methods

2.1 Molecular dynamics

MD simulations were performed for indole in three different solvent systems. Two simulations were performed on indole at an 'infinite' dilution in methanol and in water, IM and IW, respectively, where in each case there was one indole molecule and 1000 solvent molecules within the simulation box. In addition, simulations were performed on a more concentrated solution of indole in a methanol/water mixture to compare with previous neutron diffraction investigations.[9] In this case, the simulation box contained 100 indole molecules, 2900 methanol molecules and 3000 water molecules (IWM). All the simulations were performed using an NPT ensemble (T=300 K, P=1 bar) with periodic boundary conditions using the Gromacs 4.5.4 program.[10] The SPC model was used for the water molecules[11] and indole and methanol molecules were modeled using the OPLS-all-atom (OPLSAA) force field[12] but a set of different charges was used for indole (shown in Table 1). This new set of charges for indole was based on previous *ab initio* calculations for indole[13] and have successfully been used previously for EPSR investigations of indole in methanol water solutions.[9] The labeling scheme of the atoms is shown in Figure 1 and the parameters for all of the atoms in the simulations are shown in Table 1.

For each simulation, the time step was set to 2 fs and the Lennard-Jones interactions were treated using a switch cut-off from 8 Å to 10 Å. The real part of the electrostatic interactions were cut-off at 1.2 Å and the long range part treated using the Ewald summation technique. In addition, the Berendsen[14] thermostat and barostat with a coupling time of 0.1ps and 1.0ps, respectively, were used. The lincs algorithm[15] was used to keep all the bonds rigid. The density of the initial box was set to be low, in order to avoid overlapping of the atoms when inserting them randomly. A simulation was then run at 600 K until equilibrium was reached. Each system was subsequently cooled to 300 K until equilibration was reached. The final configurations of all three simulations were used to produce 1000 snapshots for the analysis herein. To avoid correlation between configurations, all the trajectories were saved every 1ps. A comparison between the previously published neutron data and the MD trajectories (see ESI†), which have been converted into putative diffraction patterns similar to previous investigations,[16] the ternary indole/water/methanol system shows good agreement between the MD simulations in the present work and the previously published neutron diffraction data.[9] The structure of pure methanol[17] and pure water[18] have been studied before.

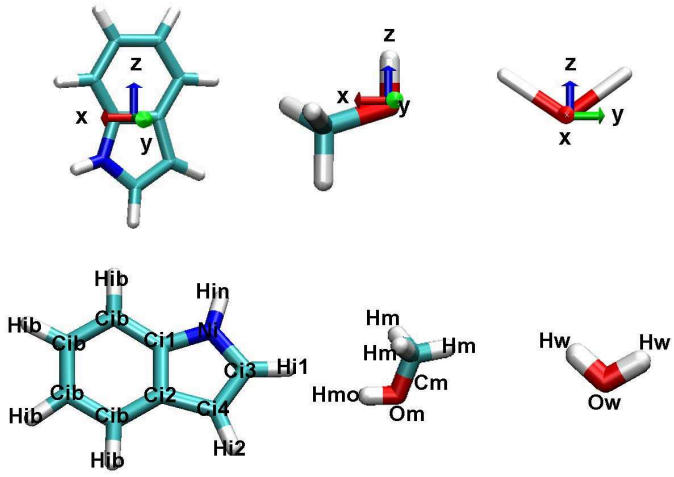

Fig. 1 (color online). Name of the atoms used in the force field of MD simulation (see table 1). Also shown are the axis definitions for all three molecules, used for characterization of the molecular ordering within the solutions.

Atom	$\varepsilon\,(kJ\,mol^{-1})$	σ (Å)	Atomic Mass	q_e
C_{i1}	0.35980	3.3997	12.011	0.0805
C_{i2}	0.35980	3.3997	12.011	-0.1055
C_{i3}	0.35980	3.3997	12.011	0.0805
C_{i4}	0.35980	3.3997	12.011	-0.1055
C_{ib}	0.35980	3.3997	12.011	-0.2245
H_{i1}	0.06276	2.5106	1.008	0.1814
H_{i2}	0.06276	2.5996	1.008	0.2091
H_{ib}	0.06276	2.5996	1.008	0.1587
H_{in}	0.65690	1.0691	1.008	0.3701
N_i	0.71130	3.2500	14.007	-0.4474
C_m	0.27614	3.5000	12.011	0.1450
H_m	0.12552	2.5000	1.008	0.0400
H_{mo}	0.00000	0.0000	1.008	0.4180
O_m	0.71128	3.1200	16.000	-0.6830
H_w	0.00000	0.0000	1.008	0.41
O_w	0.65000	3.1660	16.000	-0.82

Table 1 The parameters used in the MD simulation of indole, methanol and water molecules. The first subscript of each atom label indicates the molecule to which the atom belongs (i.e. subscripts i, m and w indicate the atoms belong in either an indole, a methanol or a water molecule, respectively). The second subscript indicates the position of that atom within the molecule. The labeled atoms can be seen in Figure 1

2.2 ANGULA analysis

2.2.1 Molecule selection criteria

In order to perform any calculation intended to describe local molecular ordering in a disordered system, a selection criteria of the studied molecules must be defined. Two goals are sought in this work: to describe at a high level of detail some selected sites of the solute molecule indole, and to provide a general description of molecular ordering around this molecule. Each description requires a different way to select the molecules chosen to perform the calculations, that are detailed in the following subsections.

2.2.1.1 Closest neighbors In order to study *only* the molecules closest to a specific site of a central molecule, the neighboring molecules have been ordered by the distance between two selected atoms of each molecule: the central molecule and the neighboring molecule. Subsequently, the closest neighbors to a particular atom on the solute are chosen for further analysis (see Fig. 2(a)). This selection process can also be performed, if necessary with second, third and following neighboring molecules.[16,19–24] It should be noted that with this closest neighbor selection criterion there is no condition imposed upon the distance which the neighboring molecules can occupy but only that these molecules must be the closest to a specific site (or the second closest and so on). This necessarily yields a distribution of distances, and therefore, indirectly a distance range, but the resultant distance distribution is a result rather than a restriction by the closest neighbor analysis.

2.2.1.2 Surface-atom criteria The next selection criteria is designed to provide a global description of molecular solvation rather than molecular ordering around a specific site. For this surface-atom criteria, the distance from the surface of a solute molecule to a given atom of a solvent molecule is defined as the minimum distance between all the atoms in the central molecule and the selected atom on the neighboring molecule (see Fig 2(b)). This selection allows the solvent molecules forming a sheet of selected solvent atoms surrounding a central molecule to be selected. This will be termed the surface-atom criteria.

For the present systems, two surface-atom criteria have been assigned for both water and methanol: from indole to the -OH hydrogens and -OH oxygen atoms of either methanol or water. This allows the spatial distribution of oxygen and hydrogen atoms around indole to be plotted. Using this criteria, the location of both solvent molecules as well as the atom (oxygen or hydrogen) through which they are interacting with indole can be determined. The surface-atom criteria can also be used to provide a calculation of the distribution of distances from oxygen and hydrogen atoms of both water and methanol to indole. Moreover, if the statistics are normalized per molecule, integration of this distribution allows for the number of oxygen or hydrogen atoms that surround the whole molecule to be calculated.

2.2.1.3 Surface-surface criteria The surface-surface selection criteria is useful to ascertain the number of solvent molecules associating with the solute irrespective of a specific atomic site. The surface-surface criteria has been defined as the distance between two molecules as the minimum one between *any* of the atoms on either molecule (Fig. 2(c)). This selection criterion has an additional advantage: if the atoms selected are recorded following this minimum distance rule, it can be established through which atoms the contact between solvent and solute molecules are made and a contact matrix can be generated. As in the previous case, the probability distribution normalized per molecule of the obtained distance has a clear quantitative meaning: its integration allows the calculation of the number of solvent molecules around a solute at a given distance.

The calculated distributions for both the surface-atom and the surface-surface analysis have a drawback, specifically, the calculated values will increase as a function of distance due to the fact that the volume of the sheets around the solute naturally increase with distance from the central solute molecule. While this effect is relatively easy to correct in a pair distribution function given its spherical symmetry, this is not the case for surface-atom or surface-surface analysis. This is due to the fact that successive shells partially 'mimic' the shape of the solvent molecule - as this is the central 'point' from which these distributions are generated.

Theoretically, both the surface-surface and surface-atom distributions should be normalized by the volume of the successive shells around the solvent molecule, as it is done for radial distribution functions when normalizing to $V_{shell} = 4\pi r^2$. However this is a non-trivial task for molecules lacking spherical symmetry, as is the case with indole. In order to correct for these volume effects, in the current analysis, a quadratic function has been fitted to at the longer distance ranges. The obtained function has then been used to normalize the surface-surface and surface-atoms distributions of distances. The obtained function cannot be called radial, since the spherical symmetry is lost due to the molecular shape, for this reason we will call this function a distance distribution function (DDF) which can be denoted by $g^s_{A,B}(d)$, where A and B might be a specific site or the surface of the solute.

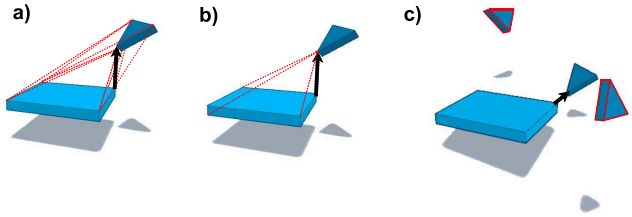

Fig. 2 (color online). Selection criteria for molecules used in this work. a) Closest contact criterion, b) minimum surface-atom distance criterion and c) minimum surface-surface distance criterion.

2.2.2 Spatial density maps (SDM)

Since indole is a planar, fairly rigid molecule, it is possible to study the shape of the solvation sheets surrounding it in solution. In order to do this, the first step is to select the solvent molecules to be investigated, where in the present work the surface-atom criteria uses the oxygen and hydrogen atoms for the solvent - methanol and/or water - around the surface of the solute indole. This allows the relative orientation of solvent molecules to be assessed, for instance to determine if the -OH groups on the solvents are pointing towards indole through either their hydrogen or oxygen atoms. In order to obtain the spatial density maps (SDM) of the methanol and/or water solvent around indole the coordinates of the selected atoms (O or H) have been calculated by using the axis sets from figure 1. In order to generate the SDMs for visualization, a 3D grid around a solute molecule is defined, and the density at each voxel is calculated using the VMD package.[25] SDMs have been used previously to determine the six degrees of freedom which describe the relative position and orientation of two molecules in the space to aid in the interpretation of liquid structures.[9,16,23,24,26–28]

2.2.3 Correlation functions

Correlation functions can be used in order to fully determine the relative position and orientation of two molecules relative to one another.[23,29] To calculate these functions, the first step is to place an orthonormal axis set on the molecules to be studied, where here the molecules have been assigned axes as defined in Figure 1. The relative position of a neighboring molecule at a distance d from the solvent to the solute can be subsequently described using the spherical polar coordinates, and obtaining the correlation function $g\left(d|\theta_{pos},\phi_{pos}\right)$. This calculation allows the regions where it is most probable to find the center of the axis set of a neighboring molecule to be defined.

Once the regions where there is a high probability of finding neighboring solute molecules has been established, the next step is to select them and proceed with the calculation of the relative orientation of these neighboring molecules with respect to the central molecule. The most probable relative orientation of a neighboring molecule at the chosen location can then be obtained by calculating the correlation function $g\left(d,\theta_{pos},\phi_{pos}|\theta_{ori},\phi_{ori},\psi_{ori}\right)$, where θ_{ori},ϕ_{ori} and ψ_{ori} are the Euler angles (using the ZYZ convention). This procedure has been described in more detail elsewhere.[23,30] Although the correlation function of the Euler angles at a given a position should be written as $g\left(d,\theta_{pos},\phi_{pos}|\theta_{ori},\phi_{ori},\psi_{ori}\right)$, for simplicity, the notation for the positional angles for which the orientations have been calculated can be omitted. This allows the correlation function to be written as $g\left(d|\theta_{ori},\phi_{ori},\psi_{ori}\right)$, where it is assumed that this calculation is performed in the selected high probability areas of $g\left(d|\theta_{pos},\phi_{pos}\right)$.

The main goal of the present work is to investigate the solvation around indole, which poses a problem in the aforementioned calculation. Similar to the calculation of the DDFs, the spherical symmetry is broken by the trivial fact that indole is not spherical. Therefore, the correlation function describing the position, should be read as $g\left(d|\theta_{pos},\phi_{pos}\right)$, where d is not a distance to a point, or specific atom, but to the surface of the solute. As a result, this correlation function will be slightly biased since what it is calculated is a convolution of the shape of the solvation shell and the shape of the solute molecule. This means that the exact probabilistic meaning of $g\left(d|\theta_{pos},\phi_{pos}\right)$ is lost. Nevertheless, the function is still highly useful in determining the "high" and "low" probability locations of neighboring molecules, and thus the most probable relative molecular ordering of solvents around a central solute molecule. Although, this is not the goal of the present work, the problem can indeed be solved by determining $g\left(d|\theta_{pos},\phi_{pos}\right)$ for the successive sheets for which the correlation function is calculated. The only drawback of not doing that is that an integration of $g\left(d|\theta_{pos},\phi_{pos}\right)$ will not be proportional to the number of molecules in a given location. This is, however, not a problem since the quantitative solvation can be determined by the DDF functions.

2.2.4 Entropy calculations

The differences of the relative position or orientation of molecules from that of a highly disordered state or a hypothetical 'ideal gas'

state at the same temperature and density conditions are encoded in the excess entropy,[29] and can be calculated as:

$$S^{excess}(d) = S^{liquid}(d) - S^{i.g.}(d) \quad (1)$$

where $S^{excess}(d)$ is the excess entropy, $S^{liquid}(d)$ is the entropy of the liquid and $S^{i.g.}(d)$ is the entropy of an hypothetical ideal gas at the same density and temperature as the liquid which is used as a reference state. In this equation d is a characteristic distance from the solute to the solvent, and in the present case will be calculated as specified in the previous section. $S^{excess}(d)$ can be used to quantify the magnitude of ordering in a liquid relative to the 'ideal gas' state at a given length scale d. Since liquids are more ordered than an ideal gas this value is always negative. Indeed the more negative the excess entropy values are, the more ordered the liquid is. On the contrary, small negative values indicate relatively little ordering.

In order to calculate the excess entropy it is necessary to calculate the distance-dependent correlation functions that encode at each distance from the solute two quantities: the amount of ordering at each distance and the correlation between variables. This has been described in detail elsewhere.[30] Although bulk S^{excess} has been often calculated, calculations which detail the distance-dependent magnitudes in order to quantify special structural motifs at a given length scale are not often performed. Here, the distance-dependent excess entropy $S^{excess}(d)$ will be used in order to determine the differences in solvation for both the infinite dilution simulations and the ternary mixtures of indole in methanol/water solutions.

Excess entropy is determined through the calculation of the excess entropies associated to all combinations of the angles defining the position and orientation of two particle through the equation:

$$S^{excess}_{\{\Omega\}}(d) = -\frac{1}{\{\Omega\}} \int g\left(d|\{\Omega\}\right) \ln\left[g\left(d|\{\Omega\}\right)\right] d\{\Omega\} \quad (2)$$

Where $S^{excess}_{\{\Omega\}}(d)$ is the excess entropy related to any angle or combination of the five angles $\{\Omega\} = \left\{\theta_{pos}, \phi_{pos}, \theta_{ori}, \phi_{ori}, \psi_{ori}\right\}$ which describes the position and orientation of two molecules relative to one another and d is the distance derived from the criterion to calculate the correlation function $g\left(d|\{\Omega\}\right)$ - as described in Sec.2.2.3 - and $\{\Omega\}$ is the integral of the angles participating in the calculation. The distance-dependent excess entropy $S^{excess}_{\{\Omega\}}(d)$ associated at some angle combination Ω is calculated as a function of a length variable such as a distance between atoms or a surface-atom criteria, or any other suitable distance criteria[30].

It should be noted that in the case of molecules for which a spherical symmetry can be assumed, such as liquid water, the integration of these excess entropies weighted by the appropriate radial distribution function allows the calculation of the thermodynamic excess entropy.[29] However, if the molecule can not be approximated to have spherical symmetry - such as is the case for indole - the formulae provided by Lazaridis et al. cannot be used in a straight forward manner. This is due to the aforementioned fact that the correlations functions - such as $g\left(r|\theta_{pos}, \phi_{pos}\right)$ - are the result of the convolution of the shape of the molecule and the solvation shells. However, here the excess entropy for solvents around *the same* molecule will be compared and, as a result, the contribution of the molecular shape can be considered to be the same for all of the molecules between simulations.

In the present work, the distance-dependent excess entropie associated with the angles describing the relative position of a neighboring molecule with respect to a central one have been calculated by :

$$S^{excess}_{\Omega_{pos}}(d) = -\frac{1}{\Omega_{pos}} \int g\left(d|\theta_{pos}, \phi_{pos}\right) \ln\left[g\left(d|\theta_{pos}, \phi_{pos}\right)\right] d\Omega_{pos} \quad (3)$$

where $\Omega_{pos} = 4\pi$ is the integral of the angles describing the position in spherical coordinates: θ_{pos} and ϕ_{pos}. The excess entropy related to the relative orientation of two associated molecules is calculated as:

$$S^{excess}_{\Omega_{ori}}(d) = -\frac{1}{\Omega_{ori}} \int g\left(d|\theta_{ori}, \phi_{ori}, \psi_{ori}\right) \ln\left[g\left(d|\theta_{ori}, \phi_{ori}, \psi_{ori}\right)\right] d\Omega_{ori} \quad (4)$$

where $\Omega_{ori} = 8\pi^2$ is the integral of the Euler angles describing the relative orientation of two molecules with respect to one another: θ_{ori} and ϕ_{ori} and ψ_{ori}. Finally, the total excess entropy due to all angular contributions can be calculated by:

$$S^{excess}_{total}(d) = -\frac{1}{\Omega} \int g\left(d|\Omega\right) \ln\left[g\left(d|\Omega\right)\right] d\Omega \quad (5)$$

where $\Omega = 32\pi^3$ is the integral of all angles, both describing molecular position and orientation. This formula gives rise to an integral in a 5D space, and as it is described in[30–32], has many computational drawbacks. For this reason, similar to previous work, the total excess entropy as an expansion to third order angular contributions has been calculated here, following Huggins notation *viz*[32]

$$S^{excess}_{total}(d) = \sum_{3}^{5} C S(d|\alpha,\beta,\gamma) - 2\sum_{2}^{5} C S(d|\alpha,\beta) + 3\sum_{1}^{5} C S(d|\alpha) \quad (6)$$

where α,β and γ are any of the five angles $\{\Omega\}$ describing the molecular position and orientation. In Eq. 6, the entropies $S(d|\alpha)$ are those calculated for one of the five angular variables, $S(d|\alpha,\beta)$ are the joint entropies for two variables, and $S(d|\alpha,\beta,\gamma)$ is that for three variables. The summations $\sum_n^5 C$ are performed by taking all possible combinations of n=1,2 or 3 angles from the five describing molecular arrangements.

3 Results

As stated above, three simulations have been performed - two on indole at infinite dilution in both methanol and water solvents (IM and IW) - and a third consisting of a ternary mixture of indole, water and methanol (IWM), where simulations on the binary mixtures allow any indole-indole interactions which might affect the indole-solvent interactions to be eliminated. Figures 3 & 4 show the DDFs and SDMs for the indole solvent interactions for the IM and IW systems, respectively. In both of these figures, the solid lines represent the binary mixtures DDFs have been further compared with the ternary IWM DDFs for the salient solvent in each case (dashed lines). The SDM shown in the insets of both figures are from the binary mixtures, however, the SDMs for the ternary mixtures yielding the similar results (not shown).

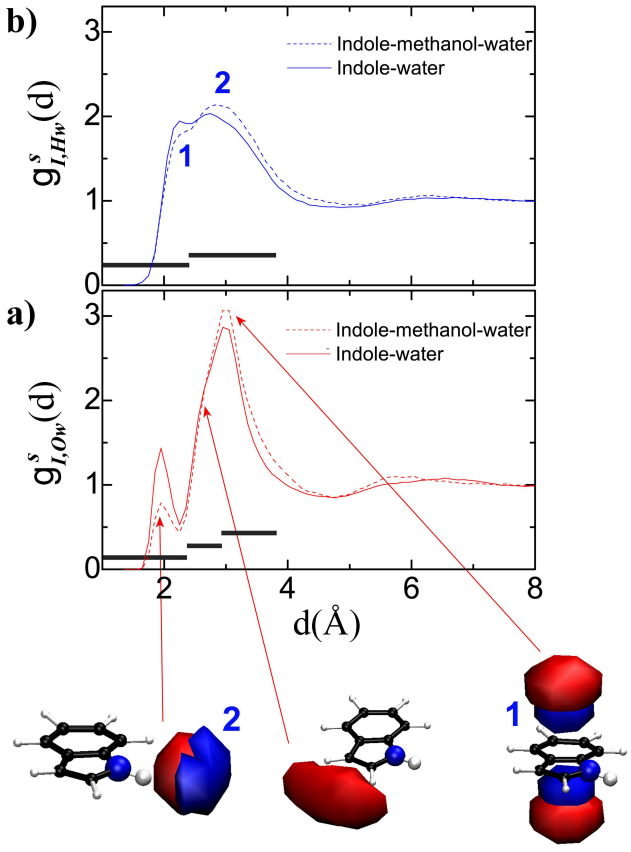

Fig. 3 (color online). Surface-atom DDF from indole to a) the oxygen (Ow) $g^s_{I,O_W}(d)$ and b) hydrogen (Hw) $g^s_{I,H_W}(d)$ atoms of water in the binary (solid lines) and ternary (dashed lines) systems. For each selected region the SDMs obtained for Ow around indole are shown (red clouds). The SDMs for some selected Hw atoms are also shown (blue clouds), they are labeled with numbers in panel b) to their corresponding location in the DDF. The black bars indicate the distance ranges used for the analysis and serve as an eye guide.

3.1 Hydration in the indole-water binary mixture

$g^s_{I,O_W}(d)$ (the DDF from the surface of indole to the water oxygen (O_W)) in Fig. 3(a) shows two main peaks that define the Ow-hydration shells around indole. The second peak has an additional shoulder before the main peak which might be indicative of an additional molecular arrangement beyond the two most prominent hydration shells. An analysis of the SDMs obtained from molecules contained within this second peak at both shorter and longer distances (which have been defined as above or below 2.89 Å, as depicted in Fig. 3a)) shows that there is a difference in molecular ordering of the water molecules around indole in these two regions. The lower panel of Fig. 3 shows the water oxygen SDMs associated with three different regions: 0 Å$< d_{I,O_W} <$ 2.28 Å, 2.28 Å$< d_{I,O_W} <$ 2.89 Å and 2.89 Å$< d_{I,O_W} <$ 3.8 Å. From these SDMs, the water molecules in the surrounding shells are located in front of the amine group in the nearest shell, in front of the hydrogen atoms of the pyrrole ring at intermediate indole-Ow distances, and in the faces of the benzene ring for the Ow atoms within the 2.89-3.8 Å distance range.

In addition to Ow-indole interactions, the $g^s_{I,H_W}(d)$ and SDMs for the water hydrogens around indole are shown in Fig. 3(b) in order to ascertain the approximate orientation of water molecules around the solute. For the Hw-indole DDF, there are only two peaks in $g^s_{I,H_W}(d)$ from which the two following regions can be defined: 0 Å$< d_{I,O_W} <$ 2.28 Å and 2.28 Å$< d_{I,O_W} <$ 3.8 Å. Hydrogen atoms within the shorter distance range (labeled as 1 in Fig. 3(b)) are associated with water molecules located on the faces of the indole benzene ring forming a hydrogen-bonding type interaction with indole, consistent with what has been previously observed for indole in methanol/water solutions.[9] Water hydrogens within the second region in $g^s_{I,H_W}(d)$ (labeled as 2 in Fig. 3(b)) are associated with water molecules that are hydrogen bonding, through the water oxygen, to the amine hydrogen on indole. The water molecules in this region are thus oriented such that, as expected, H_W is pointing away from indole towards the bulk water solvent.

While $g^s_{I,O_W}(d)$ is sufficient in determining the locations of Hw or Ow contacts around indole, this quantity is not necessarily indicative of the minimum distance of water molecules to indole. This is due to the fact that direct contacts between water and indole can be done by virtue of either water hydrogen or oxygen atoms, so that a simple distance range assessment of nearest neighbor hydration around indole is not straight forward to determine from Fig. 3. This is clear from the $g^s_{I,O_W}(d)$ where water molecules located in front of the amine group and in front of the faces of the benzene ring are pointing to indole through the Ow atom and the HW atom, respectively. Although at both locations water molecules are in direct contact with indole, the water oxygens are located at different distances in the DDF. However, in order to ease the subsequent discussion of the hydration of indole - 'first', 'second' and 'third' shells will be classified by virtue

of their ordering in $g^s_{I,O_W}(d)$ DDF (Fig. 3 (a)).

3.2 Indole-methanol interactions in the binary mixture

The $g^s_{I,O_M}(d)$ DDFs for the IM simulations are shown in Fig. 4(a)(solid line), where four distance regions can be distinguished: 0 Å< d_{M,O_W} < 2.28Å, 2.28 Å< d_{M,O_W} < 2.8 Å, 2.8 Å< d_{M,O_W} < 3.68 Å and 3.68 Å< d_{M,O_W} < 5.0 Å. Similar to the hydration interactions with indole, the closest contacts between indole and methanol oxygen (Om) are by virtue of hydrogen bonding to the amine group hydrogen, which can be seen in the indole-Om SDMs for this distance range (Lower panel, Fig. 4(a)). The Om solvation shells at further distances around indole are also similar to the water coordination, with the oxygen atoms being located at the edges of pyrrole ring in the second region and in the faces of the benzene ring in the third shell. The methanol orientation has also been qualitatively determined through reference to both Fig. 4(b) - the DDF associated to the hydrogen of its hydroxyl group ($g^s_{I,H_M}(d)$) - and the associated SDMs (lower panel, Fig. 4). In a coordination that is similar to the hydration, the orientation of the methanol -OH group is such that it forms a hydrogen bond with the amine group on indole and coordinates through its -OH hydrogen to the benzene ring, also in agreement with previous work on IWM solutions.[9] This result, i.e. that benzene induces structural ordering on methanol, is consistent with a recent work in which the formation of methanol chains is enhanced by the addition of benzene.[33]

There is however a difference between the hydration and the methanol coordination, specifically, there is a further shell of density (3.68 Å< d_{M,O_W} < 5.0 Å) around indole which does not appear in the water DDFs (Fig. 3). In order to probe the locations the methanol molecules in three dimensions around the indole molecule Fig. 4 also shows the SDM for methanol oxygens at all of the four aforementioned distance ranges. From these SDMs it can be clearly seen that methanol is occupying the space left between the oxygens located face of the benzene rings as well as the hitherto unsolvated spaces located are the edges of the indole molecule.

3.3 Comparison of hydration and methanol-indole interactions for binary mixtures

The comparison of the hydration in Fig. 3 with the methanol solvation in Fig. 4 indicates that there are similar features between the molecular ordering of water and methanol molecules around indole within the short to medium distance ranges (those less than 2.89Å for water and 2.8Å for methanol), both with respect to the $g^s_{I,B}(d)$ and the SDMs. While some solvent molecules hydrogen bond to the amine group through an oxygen contact (at low r values (less than 2.28Å), others are located in front of the hydrogens of the pyrrole ring (at more intermediate values), and

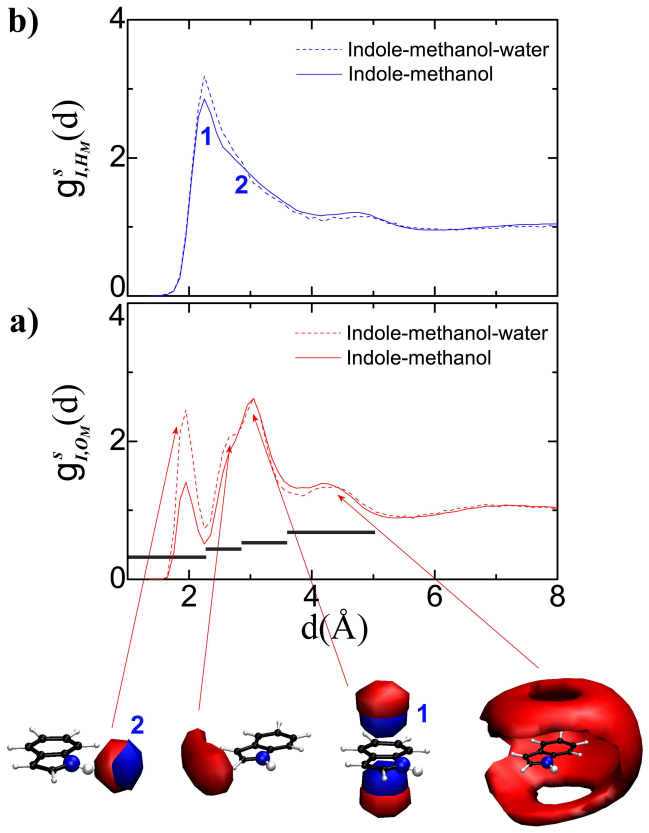

Fig. 4 (color online). Surface-atom DDF from indole to a) the oxygen (Ow) $g^s_{I,O_M}(d)$ and b) hydrogen (Hw) $g^s_{I,H_M}(d)$ atoms of methanol in the binary (solid lines) and ternary (dashed lines) systems. For each selected region the SDMs obtained for Ow around indole are shown (red clouds). The SDMs for some selected Hw atoms are also shown (blue clouds), they are labeled with numbers in panel b) to their corresponding location in the DDF. The black bars indicate the distance ranges used for the analysis and serve as an eye guide.

finally the solvating molecules are also located around the faces of the benzene ring, with the -OH hydrogens for both solvents pointing towards the center of the benzene ring. However, for methanol, at distances greater than 2.8Å, there is an additional solvation shell via the methanol oxygen atom as shown in Fig. 4. Oxygen atoms of methanol occupy the spaces left by molecules in first and second shells, forming two rings parallel to the plane defined by indole, where this feature is not present in water. Since the differences between the two solvents arise at distances $d > 2.8$ Å from indole, the distance ranges where the differences between the two solvents arise (eg 2.89 Å< d_{I,O_W} < 4.3 Å for Ow and 2.8 Å< d_{M,O_W} < 3.68 Å for Om) have been assessed in more detail below, with the aim of determining if there is any difference between the two solvents in their arrangements within these dis-

tance ranges that can impinge a different molecular ordering for the other solvation shells.

Water and methanol molecules at distances in the third region (as determined by $g^s_{I,O_W}(d)$ and $g^s_{I,O_M}(d)$ in Figs. 3 & 4, respectively) are largely localized above and below the benzene motif on indole. In order to select the molecules at these positions, a pseudo-atom has been defined at the center of the benzene ring. Next, the first neighbor molecules to that pseudoatom have been selected using the "closest neighbors" criteria described in Sec. 2.2.1. Choosing a position at the center of the benzene ring, allows avoids 'contamination' of the positional correlation functions coming from molecules located at the edges of indole. For both methanol and water molecules nearest to the faces of the benzene ring, the correlation functions related to their position $g\left(d|\theta_{pos},\phi_{pos}\right)$ (Fig. 5(a)) have been calculated. In Fig. 5 the two spots of density - which show a high probability of finding a molecule are located at $\theta_{pos}\approx 66^\circ$ and $\phi_{pos}=\pm 90^\circ$, above and below the benzene ring.

Once the most probable location of molecules has been determined, the orientations of the molecules in these locations has been further assessed by selecting one of the two spots representing a high probability location. Any of them can be chosen due to the plane symmetry of indole, here the one located at $\theta_{pos}\approx 66^\circ$ and $\phi_{pos}=90^\circ$ has been chosen. The 2D correlation functions of Euler angles for molecules at the chosen location which determine their relative orientation are then assessed ($g\left(\theta_{ori},\phi_{ori}\right)$ and $g\left(\theta_{ori},\psi_{ori}\right)$). A region of high density in the first correlation function means that the z axis associated with the solvent molecule is most likely pointing to a given direction given by the azimuthal and equatorial angles θ_{ori} and ϕ_{ori}. For both water and methanol, there is a single localized spot of density (Fig. 5), from which it can be concluded that there is a single orientation for the z axis, and thereby the solvent molecules, determined by $\theta_{ori}=90^\circ$ and $\phi_{ori}=-90^\circ$. The molecules with their z-axis located at this orientation are further selected and then analyzed to assess the rotation angle ψ_{ori} around the new z-axis.

In Fig. 5(c) the correlation function $g\left(\theta_{ori},\psi_{ori}\right)$ for both water and methanol are shown. Although there is no difference between water and methanol, with respect to $g\left(\theta_{pos},\phi_{pos}\right)$ or in $g\left(\theta_{ori},\phi_{ori}\right)$, Fig. 5(c) shows a difference between the two molecules for $g\left(\theta_{ori},\psi_{ori}\right)$. Specifically, here water shows two spots of density while methanol only shows a single spot. This means than for water there are two possible orientations of these water molecules around indole, while for methanol there is a single orientation. To further investigate this difference in relative orientations, the most probable orientations for both water and methanol molecules have been generated using the maximum probability regions defined by the spots in all previous correlation functions in Fig. 5.

Fig. 5(d) shows that both water and methanol in the studied region are oriented such that the -OH bond is pointing to the center of the benzene ring, consistent with what has been previously observed for indole in methanol/water solutions.[9] However, while water has two different most probable orientations with one hydrogen pointing perpendicular to an imaginary axis bonding the center of the two rings, methanol has only one orientation where its methyl group sits in front of the pyrrole ring. In other words: for methanol -OH oxygen and hydrogen atoms are aligned with the imaginary axis for benzene as defined above.

The orientation observed for methanol suggests that the additional oxygen shell at distances $3.68\ \text{Å} < d_{M,O_W} < 5.0\ \text{Å}$ observed for methanol in Fig 4 may be related to methanol molecules forming a hydrogen bond through their oxygen atom, with other methanol molecule which sit front of the faces of the benzene ring.

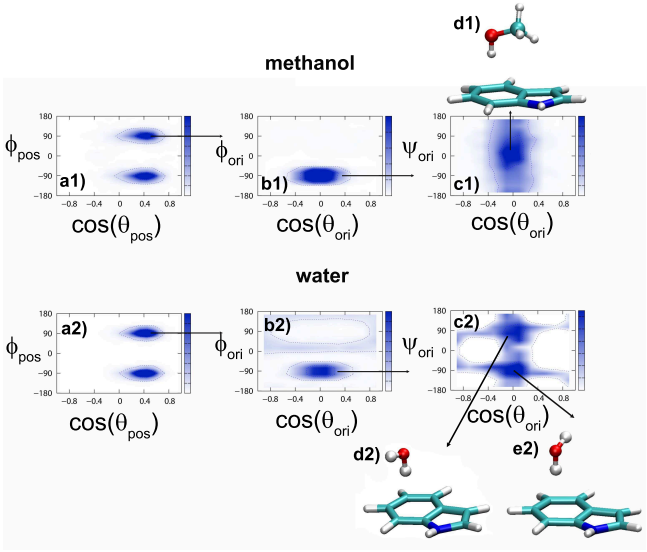

Fig. 5 (color online). First neighbour ordering of the solvent molecules around indole. Panels a_1 and a_2 show the two dimensional PDF $g(\theta_{pos},\phi_{pos})$ maps giving the most probable position of oxygen atom of solvent molecules around Indole. Panels b_1 and b_2 show the two dimensional PDF $g(\theta_{ori},\phi_{ori})$ giving the orientation of Z axis of the solvent with respect to the Z axis of Indole for the position selected in figures a_1 and a_2. Panels c_1 and c_2 show the two dimensional PDF $g(\theta_{ori},\psi_{ori})$ related to the rotation of the molecule around their Z axis in its new position. Inset d_1 show the most probable position and orientation of methanol around indole calculated for the maximum probability regions determined in figures a_1, b_1 and c_1. Finally, insets d_2 and e_2 show the most probable position and orientation of water around indole calculated for the maximum probability regions determined in figures a_2, b_2 and c_2

.

3.3.1 Excess entropy calculations in the binary mixtures

In order to quantify how relatively ordered the molecules around indole are, the solid lines in Fig. 6 show the excess entropy associated to the position ($S^{excess}_{\Omega_{pos}}(d)$) and the orientation ($S^{excess}_{\Omega_{ori}}(d)$)

of methanol (red) and water (blue) with respect to indole[30] for the binary mixtures. In order to ease the identification of the different distance regions given by $g^s_{I,O_W}(d)$ and $g^s_{I,O_M}(d)$ in Figs. 3& 4, the positional entropy has been calculated using the position of oxygen for both water and methanol.[29]

It is clear in Fig 6 that methanol is more ordered around indole compared with water, both with respect to its position and its orientation as both excess entropies are smaller than those for water. The main difference between the excess entropies of methanol and water is especially important within the distance region where the molecules are located in the faces of the benzene ring of indole. This might be expected from the results shown in Fig. 5 where it is clear that there are less degrees of freedom in the orientation of methanol around indole compared with water; namely water has two possible orientations and methanol only one. However, that this fact is also captured in the positional ordering of both molecules is rather unexpected. Thus, it seems that in the present case, the higher solubility of indole in methanol compared with water is likely to be dominated by enthalpy and not by entropy as methanol is clearly more ordered around indole compared with water.

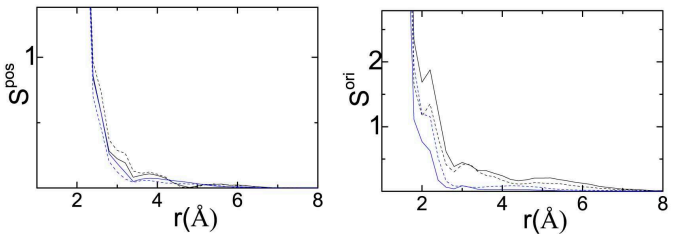

Fig. 6 (color online). Excess entropies related to the positional (a) and orientational (b) degrees of freedom for water (blue line) and methanol (red line) in their respective binary(solid line) and ternary (dashed line) mixtures.

3.4 The ternary mixture I-M-W

In order to understand the changes that occur in the ordering of methanol around indole upon the addition of water, the same calculations as those performed for the binary mixtures have also been performed for a ternary mixture of indole, water and methanol. Figure 7 shows the surface-atom and surface-surface DDFs for all atoms of water and methanol around the indole molecule both in the binary systems (solid lines) and in the ternary mixture (dashed lines). In the first instance, it is desirable to understand if there is any change in the number of water or methanol molecules surrounding indole of the ternary system compared with the binary systems. The goal of these calculations is to quantify if methanol is actively changing the hydration of indole, or if it is a passive agent that avoids water contacts. In order to quantify the number of molecules in the first solvation

	$N^{bin}_{1^{st}shell}$	$N^{ter}_{1^{st}shell}$	$N_{bin\to ter}$	$\rho^{W}_{bin\to ter}$
Water	$N^W_{IW}=30$	$N^W_{IMW}=7.38$	$N^W_{IW\to IMW}=0.25$	$\rho^W_{IW\to IMW}=0.30$
Methanol	$N^M_{IM}=18.2$	$N^M_{IMW}=12$	$N^M_{IM\to IMW}=0.66$	$\rho^M_{IM\to IMW}=0.68$

Table 2 The first and second columns contain the coordination numbers for the first solvation shell for methanol and water in the binary (bin) and ternary (ter) systems. The relative increase of the coordination numbers for water and methanol when going from the binary to the ternary system can be found in the third column and are calculated as $N_{bin\to ter}=N^{bin}_{1^{st}shell}/N^{bin}_{1^{st}shell}$ (see text for details). The last column contains the relative increase of density for water and methanol of the ternary system with respect to the ternary system $\rho_{bin\to ter}=\rho^{bin}_{1^{st}shell}/\rho^{bin}_{1^{st}shell}$ (see text for details)

shell, the nearest shell has been defined by the first minimum of the surface-surface DDFs $g^s_{I,W}(d)$ and $g^s_{I,M}(d)$. In this way, the first solvation shell can be defined taking into account any contact between solvent and solute. This overcomes the aforementioned problem when using $g^s_{I,O_W}(d)$ and $g^s_{I,O_M}(d)$: that functions do not order molecules, but atoms pertaining to molecules.

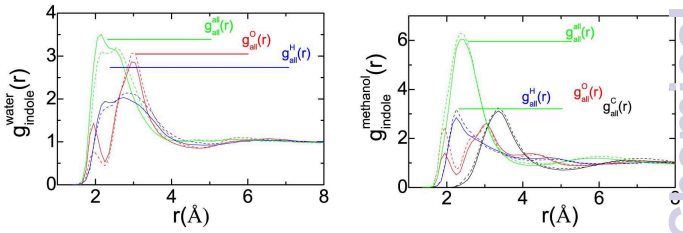

Fig. 7 (color online). Surface-surface and surface-atom distribution functions for water and methanol. Straight lines are calculated in the binary systems and dashed lines are the same functions calculations done for the ternary system.

In the binary systems, the number of water and methanol molecules found in this first solvation shell are $N^W_{IW}=30$ and $N^M_{IM}=18.2$, and for the ternary mixture the coordination numbers obtained are $N^W_{IMW}=7.38$ and $N^M_{IMW}=12$ for water and methanol, respectively. The relative increase of molecules within this closest solvation shell around indole for the binary systems compared to the ternary system for water is $N^W_{IW\to IMW}=N^W_{IMW}/N^W_{IW}=0.25$ and for methanol $N^M_{IM\to IMW}=N^M_{IMW}/N^M_{IM}=0.66$.

These values are, nonetheless, not only affected by interaction between molecules, but also by their density in the solution: in a ternary solution with a high density of water molecules there will be more water molecules available to form a solvation shell. To eliminate this effect, the increase of the density of the binary

system with respect to the ternary system has also been calculated. The results are $\rho^W_{IW\to IMW} = \rho^W_{IMW}/\rho^W_{IW} = 0.30$ for water and $\rho^M_{IM\to IMW} = \rho^M_{IMW}/\rho^M_{IM} = 0.68$ for methanol. These two calculations show the decrease of the occupancy of both water and methanol per unit of volume of the ternary system with respect to the binary systems. The table 2 with all these calculated values has been produced in order to ease the comparison for both methanol and water.

For methanol molecules around indole, both values for the relative increase of molecules in the first solvation shell $N^M_{IM\to IMW} = 0.66$ and the increase in the density for the binary methanol/indole systems compared with methanol in a solution of indole and water $\rho^W_{IM\to IMW} = 0.68$ are very close. This suggests that the mechanism of methanol-solvation around indole is largely the same with and without water. In other words, water seems to not affect the ordering of methanol molecules around indole.

On the other hand, the relative increase of water molecules in the first solvation shell around indole in the binary system is $N^W_{IW\to IMW} = 0.25$, and is smaller than what would be expected from a trivial effect of molecular occupancy ($\rho^W_{IW\to IMW} = 0.30$). Methanol is thus actively influencing the hydration of indole since the increase of the coordination number for the first hydration shell is smaller than that would be expected from a greater availability of water molecules due to an increase of its density. The smaller value of $N^W_{IW\to IMW} = 0.25$ with respect to $\rho^W_{IW\to IMW} = 0.30$ suggests that methanol expels water from the first solvation shell.

In order to ascertain which water molecules are removed from the first solvation shell in the ternary system, Fig. 7 shows the surface-atom DDFs for all atoms of both methanol and water in the binary (dashed lines) mixtures compared with those from the ternary (solid lines) mixtures. From this figure, there is an apparent increase of water molecules forming hydrogen bonds with the amine group upon the addition of methanol, concomitant with a decrease in amine-methanol hydrogen bonding interactions. This increased hydration of the amine group is likely due to the smaller volume of water molecules compared to methanol molecules. Indeed this increase in hydration around the amine group would lead to the conclusion that water should be added rather than expelled from the first hydration shell of indole, which it clearly is not. It should be noted, however, that although in the surface-atom distribution functions this hydration effect seems to be very important, it is amplified by the fact that $g^s_{I,B}(d)$ is normalized by the volume of the shells surrounding the indole molecule as explained in Sec. 2.2.1. In fact, this redistribution of water around the amine group has a very small effect on the water-indole coordination number when going from the binary (3.8 molecules) to the ternary system (3.59 molecules) in the first solvation shell. The key of water expulsion from the first solvation shell around indole is thus not due to molecules accepting a hydrogen bond from the amine group on indole.

Binary and ternary mixtures surface-oxygen DDFs are also different in the region 2.28 Å$< d_{M,O_W} <$ 2.89 Å, where specifically, there is a decrease on the number of water molecules at these distances in the ternary system with respect to the binary IW solution. From Figs. 3 & 5, water and methanol molecules within this distance range are largely located in the faces of the benzene ring of indole. Therefore, it can be concluded that the decrease in hydration in the IWM solutions within this region is due to the methanol molecules occupying the positions formerly occupied by water molecules in the binary system, above and below the benzene ring.

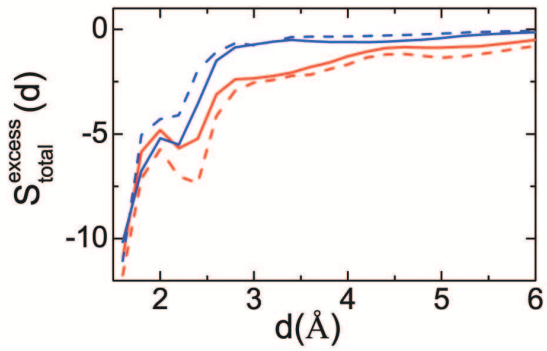

Fig. 8 (color online). Surface-surface and surface-atom distribution functions for water and methanol. Solid lines are calculated in the binary systems and dashed lines are the same functions calculations done for the ternary system.

Figure 6 shows the entropy calculations for the ternary system (dashed lines) compared with those for both IW and IM binary systems (solid lines). Interestingly, there is an increase in the positional excess entropy for methanol in the ternary system relative to the binary system, while there is a decrease in the positional excess entropy for water in the ternary system compared with the binary IW system. This means that the position of water is more ordered in the ternary mixture in comparison with that of the binary system, and that methanol is more disordered (less ordered) in the IWM mixture relative to the IM system. Further, the positional entropy for both methanol and water in the ternary system are quite similar to one another. The opposite phenomenon occurs with respect to the orientational entropy, that is water is more disordered around indole in the ternary compared with water in the binary system and methanol is more ordered.

Figure 8 shows the total excess entropy for both water and methanol around indole in binary and ternary mixtures calculated from equation 6. Comparison of this figure with the positional and orientational entropy calculations shown in Fig. 6, shows that the features of the curve in Fig. 8 are largely driven by the ori-

entational excess entropy. This gives a more a coherent picture of the process of solvation of indole in methanol/water mixtures compared with those from the IW and IM 'infinite dilution' simulations. That is the addition of methanol to indole in water has two effects. First of all methanol appears to be expelling water molecules from the first solvation shell around indole. Second when adding methanol to the binary mixture IW, water becomes slightly more disordered around indole given its increase in the excess entropy depicted in Fig. 8, with the side effect of methanol becoming more ordered around the solvent which is seen in Fig. 8 as a decrease in its excess entropy.

For completeness, the closest contacts between indole molecules has been calculated for the ternary system. Contrary to what occurs with water or methanol, it is not possible to select one molecular site, since this selection will bias the results towards the chosen site. For this reason, the analysis of the short range order between indole molecules has been done using the contact matrix shown in Fig. 9a. From this figure, the most probable contact between two closest indole molecules occurs through the benzene ring Cib atoms and the amine Hin hydrogen. This is comparable to the strong interaction of both methanol and water through a cation-π interaction with the benzene ring as is depicted by the indole molecules shown in Fig. 9b. The most probable number of molecules interacting through this contact is two, suggesting that there is no formation of long range ordered structures, or stacking, involving indole molecules. This result is consistent with previous findings for indole in methanol water solutions using neutron diffraction experiments.[9].

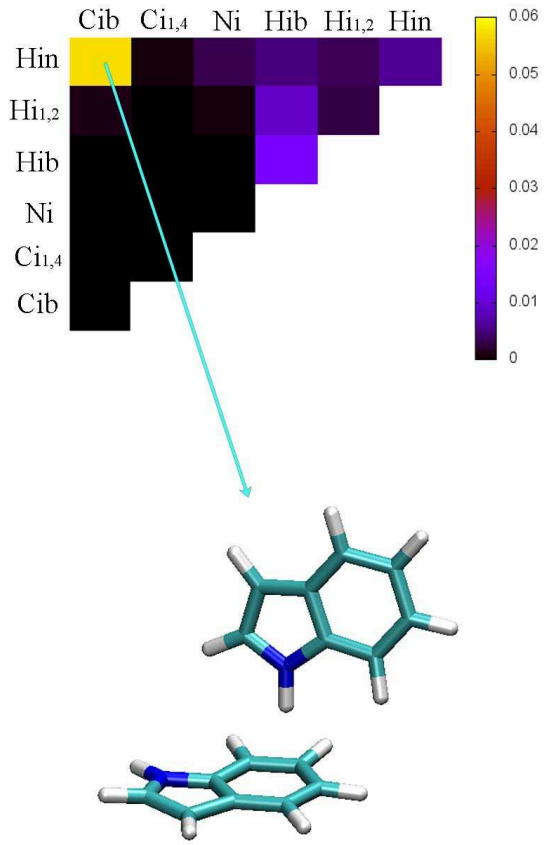

Fig. 9 (color online). Contact matrix using the atom names of figure 1. (b) Two Indole molecules interacting thrpugh the most probable ontact (Cb-Hn) determined by the contact matrix

4 Conclusions

The solvation of indole by either methanol or water seems to be predominately determined by electrostatics, that is overall the -OH motifs of both solvent molecules interact with indole such that both the amine group and benzene ring motifs are solvated by the hydroxyl groups in both solutions. This is consistent with previous neutron diffraction measurements on indole/methanol/water solutions,[9] as well as with infra-red measurements on indole derivatives in solution which have been shown to be sensitive to -OH bonding to the benzene ring motifs.[34,35] From the excess entropy calculations here, this appears to be enthalpically driven rather than entropically driven given that methanol is more ordered compared with water around indole in solution and indole has a much higher solubility in methanol compared with water.

Explanations of the 'hydrophobic effect' have suggested that it is the expulsion of water molecules from hydrophobic surfaces that provides the driving force for biomolecular association and solvation as this expulsion results in an increase of entropy.[36,37] Interestingly here, while the addition of more hydrophobic motifs in solution - namely the -CH_3 groups in methanol - does lead to an expulsion of water around the surface of indole - the replacement of water with methanol results in a *decrease* in entropy of solvation around indole rather than an increase. This is consistent with calorimetry measurements which showed that an increase in binding affinity was determined by an increase in enthalpy rather than entropy for aromatic molecules bound to protein receptor sites,[38] and that in protein binding sites unfavorable entropic changes can be overcome by favorable enthalpic changes.[39]

On the molecular scale, the key to understanding this enhanced solubility for indole in methanol compared with water appears to be related to the solvation of the benzene ring. The results here indicate that methanol forms more ordered interactions around indole compared with water in ternary indole/water/methanol solution, that replaces the benzene water contacts which are present in the absence of methanol. This preferred solvation by methanol resulting in a stronger -OH hydrogen bond with π electrons of benzene ring. This is most likely due to the fact that, with water, there is a competition between the orientation impinged by the molecular symmetry and the dipole moment that causes a less oriented hydrogen bond and results in a less intense contact

between water and the benzene ring on indole. This is not the case for methanol, for which both dipole and H-bond are, by in large, roughly in the same direction as one another. Therefore, the solubility of indole in methanol/water solutions seems to be dominated by the molecular ordering of the methanol -OH groups rather than any substantial change in water ordering which might have occurred upon the addition of methanol to an aqueous indole solution.

5 Acknowledgments

This work was supported by the Spanish MINECO (grants No. FIS2014-54734-P and FIS2012-39443-C02-01) and by the Government of Catalonia (grants No. 2009SGR-1003 and 2014SGR-00581)and the UK Engineering and Physical Sciences Research Council (EP/J002615/1). One of the authors (L. C. Pardo) would like to thank E. Píscore for inspiring part of the present work.

References

1 K. T. Savjani, A. K. Gajjar and J. K. Savjani, ISRN Pharmaceutics, 2012, **2012**, year.
2 S. D. Campbell, K. J. Regina and E. D. Kharasch, Journal of Biomolecular Screening, 2014, **19**, 437–444.
3 L. K. Chico, L. J. Van Eldik and D. M. Watterson, Nat. Rev. Drug. Discov., 2009, **8**, 1474–1776.
4 S. Baluja, R. Bhalodia, M. Bhatt, N. Vekariya and R. Gajera, International Letters of Chemistry, Physics and Astronomy, 2013, **17**, 36–46.
5 K. G. H. Desai, R. K. Anandrao and T. M. Aminabhavi, Journal of Chemical & Engineering Data, 2003, **48**, 942–945.
6 This program can be downloaded from, `http://gcm.upc.edu/en/members/luis-carlos/angula/ANGULA`, [Online; accessed 15-June-2016].
7 A. K. Soper, Mol. Sim., 2012, **38**, 1171–1185.
8 R. L. McGreevy, Journal of Physics: Condensed Matter, 2001, **13**, R877.
9 A. J. Johnston, Y. Zhang, S. Busch, L. C. Pardo, S. Imberti and S. E. McLain, The Journal of Physical Chemistry B, 2015, **119**, 5979–5987.
10 B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, J. Chem. Comput., 2008, **4**, 435–447.
11 J. Hermans, H. J. Berendsen, W. F. Van Gunsteren and J. P. Postma, Biopolymers, 1984, **23**, 1513–1518.
12 G. A. Kaminski, R. A. Friesner, J. Tirado-Rives and W. L. Jorgensen, The Journal of Physical Chemistry B, 2001, **105**, 6474–6487.
13 K. E. Norman and H. Nymeyer, Biophys. J., 2006, **91**, 2046–2054.
14 H. J. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. Haak, J. Chem. Phys., 1984, **81**, 3684–3690.
15 B. H. et al., J. Comput. Chem., 1997, **18**, 1463–1472.
16 S. Busch, C. D. Lorenz, J. W. Taylor, L. C. Pardo and S. E. McLain, J. Phys. Chem. B., 2014, **118**, 14267–14277.
17 T. Yamaguchi, K. Hidaka and A. Soper, Molecular physics, 1999, **96**, 1159–1168.
18 S. Dixit, A. Soper, J. Finney and J. Crain, EPL (Europhysics Letters), 2002, **59**, 377.
19 A. Silva-Santisteban, A. Henao, S. Pothoczki, F. J. Bermejo, J. L. Tamarit, E. Guardia, G. J. Cuello and L. Pardo, Journal of Physics: Conference Series, 2014, **549**, 012014.
20 N. Veglio, F. Bermejo, L. Pardo, J. L. Tamarit and G. Cuello, Physical Review E, 2005, **72**, 031502.
21 L. Pardo, N. Veglio, F. Bermejo, J. L. Tamarit and G. Cuello, Physical Review B, 2005, **72**, 014206.
22 M. Rovira-Esteva, N. A. Murugan, L. Pardo, S. Busch, J. L. Tamarit, G. Cuello and F. Bermejo, The Journal of chemical physics, 2012, **136**, 124514.
23 L. Pardo, A. Henao, S. Busch, E. Guàrdia and J. L. Tamarit, Physical Chemistry Chemical Physics, 2014, **16**, 24479–24483.
24 N. Steinke, R. J. Gillams, L. C. Pardo, C. D. Lorenz and S. E. McLain, Phys. Chem. Chem. Phys., 2016, **18**, 3862–3870.
25 W. Humphrey, A. Dalke and K. Schulten, Journal of Molecular Graphics, 1996, **14**, 33–38.
26 S. Busch, L. C. Pardo, W. B. O'Dell, C. D. Bruce, C. D. Lorenz and S. E. McLain, Phys. Chem. Chem. Phys., 2013, **15**, 21023–21033.
27 R. J. Gillams, J. V. Busto, S. Busch, F. M. Goñi, C. D. Lorenz and S. E. McLain, The journal of physical chemistry. B, 2015, **119**, 128–39.
28 A. J. Johnston, S. Busch, L. C. Pardo, S. K. Callear, P. C. Biggin and S. E. McLain, Phys. Chem. Chem. Phys., 2016, **18**, 991–999.
29 T. Lazaridis and M. Karplus, J. Chem. Phys., 1996, **105**, 4294.
30 L. C. Pardo, A. Henao and A. Vispa, Journal of Non-Crystalline Solids, 2015, **407**, 220–227.
31 H. Matsuda, Phys. Rev. E., 2000, **62**, 3096–3102.
32 D. J. Huggins, J. Chem. Phys., 2012, **136**, 064518.
33 J. Shephard, S. Callear, S. Imberti, J. Evans and C. Salzmann, Physical Chemistry Chemical Physics, 2016, **18**, 19227–19235.
34 W. Zhang, B. N. Markiewicz, R. S. Doerksen, A. B. Smith, III and F. Gai, Phys. Chem. Chem. Phys., 2016, **18**, 7027–7034.
35 B. N. Markiewicz, D. Mukherjee, T. Troxler and F. Gai, The Journal of Physical Chemistry B, 2016, **120**, 936–944.
36 D. Chandler, Nature, 2005, **437**, 640–647.
37 N. T. Southall, K. A. Dill and A. D. J. Haymet, The Journal of Physical Chemistry B, 2002, **106**, 521–533.

38 P. W. Snyder, J. Mecinović, D. T. Moustakas, S. W. Thomas, M. Harder, E. T. Mack, M. R. Lockett, A. HÃ©roux, W. Sherman and G. M. Whitesides, Proceedings of the National Academy of Sciences, 2011, **108**, 17889–17894.

39 D. Huggins, Biophysical Journal, 2015, **108**, 928 – 936.