Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.
This journal is © the Owner Societies 2022

PCCP

ARTICLE TYPE

Cite this: DOI: 00.0000/xxxxxxxxxx

An atomistic explanation of the ethanol-water azeotrope

Vincenzo Carravetta,[a,*] Anderson Herbert de Abreu Gomes,[b,c,‡] Ricardo dos Reis Teixeira Marinho,[d,f,‡] Gunnar Öhrwall,[g,‡] Hans Ågren,[h,‡] Olle Björneholm,[h,‡] and Arnaldo Naves de Brito[b,¶]

Received Date
Accepted Date

DOI: 00.0000/xxxxxxxxxx

Ethanol and water form an azeotropic mixture at an ethanol molecular percentage of ∼ 91% (∼ 96% by volume) , which prohibits ethanol from being further purified via distillation. Aqueous solutions at different concentrations in ethanol have been studied both experimentally and theoretically. We performed cylindrical micro-jet photoelectron spectroscopy, excited by synchrotron radiation, 70 eV above C1s ionization threshold, providing optimal atomic-scale surface-probing. Large model systems have been employed to simulate, by molecular dynamics, slabs of the aqueous solutions and obtain an atomistic description of both bulk and surface regions. We show how the azeotropic behaviour results from an unexpected concentration-dependence of the surface composition. While ethanol strongly dominates the surface and water is almost completely depleted from the surface for most mixing ratios, the different intermolecular bonding patterns of the two components cause water to penetrate to the surface region at high ethanol concentrations. The addition of surface water increases its relative vapour pressure, giving rise to the azeotropic behaviour.

1 Hydrogen bond analysis

To obtain further insights into the atomistic mechanisms determining the surface composition of water ethanol mixtures, we analyse the differences of the hydrogen bonding (HB) networks between surface and bulk as a function of ethanol molecular percentage. A quantitative description of the HB network is strongly dependent on the, somewhat arbitrary, parameters used for the HB characterisation in terms of bonding energy and/or intermolecular geometry. For our analysis, we have adopted the geometric definition of the HB, using, as cutoff values for the distance R between the O atoms and the H-O-O angle, the default ones of the GROMACS code, namely R = 3.5 Å and angle = 30 degrees. The values of the average number of HBs we obtain are then dependent on the parameters employed, but they can be considered valid for identifying differences between bulk and surface regions and trends as the concentration of the solution varies.

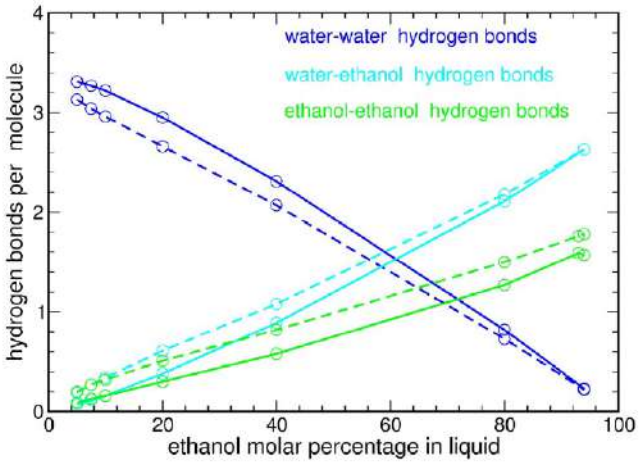

Fig. SI-1 Surface (dashed lines) and bulk (full lines) HB numbers for water-water, water-ethanol (counted per water molecule) and ethanol-ethanol pairs as a function of ethanol molar percentage X.

Figure SI-1 shows the average HB number between water-water, water-ethanol and ethanol ethanol pairs, in the different surface and bulk regions of the slab as a function of ethanol molecular percentage X; the HB numbers for water-ethanol are per water molecule. For both bulk and surface, the average number of ethanol-ethanol HBs is seen to rapidly increase as the ethanol molar percentage X increases, as it will gradually become

[a]CNR-IPCF, Institute of Chemical and Physical Processes, via G.Moruzzi 1, I-56124 Pisa, Italy.
[b]Dept. of Applied Physics, Institute of Physics "Gleb Wataghin", Campinas University, CEP: 13083859 CampinasSP, Brazil.
[c]Brazilian Synchrotron Light Laboratory (LNLS), Brazilian Center for Research on Energy and Materials (CNPEM), PO Box 6192, 13083-970, Campinas, SP, Brazil
[d]Institute of Physics, Federal University of Bahia, 40.170-115, Salvador, BA, Brazil
[e]Institute of Physics, Brasilia University (UnB), Box 4455, Brasília 70910-970, Brazil
[f]MAX IV Laboratory, Lund University, Box 118, SE-22100 Lund, Sweden
[g]Division of X-ray Photon Science, Department of Physics and Astronomy, Uppsala University, Box 516, 75120 Uppsala, Sweden
‡these authors contributed equally to this work
∗vincenzo.carravetta@pi.ipcf.cnr.it
¶arnaldo@ifi.unicamp.br

easier for ethanol molecules to find each other. The number of HBs is generally higher at the surface than in the bulk, reflecting the surface enrichment of ethanol, see Figures 7 and SI-1, at practically every concentration. Turning to water, we see that for both bulk and surface, the average number of water-water HBs rapidly decreases as the ethanol molar percentage X increases, as it will gradually become more difficult for water molecules to encounter each other to form water clusters. The water HB number is lower in the surface region than in the bulk, especially for lower X. This can be correlated to both the generally reduced coordination at the surface, and the more pronounced breaking of the water HB network in the ethanol-enriched surface region. Both in the bulk and at the surface, the loss of water-water HBs is thus partially compensated by the formation of water-ethanol HBs with increasing X. Note that the number of water-ethanol HBs displayed in Figure SI-1 is counted per water molecule, illustrating the increased importance of water-ethanol HB for water solvation at high X. Also note that what we mean by "surface" is a 6 Å corrugated layer which always include molecules not precisely at the liquid vacuum interface. Above X=60%, there are actually more water-ethanol than water-water HBs.

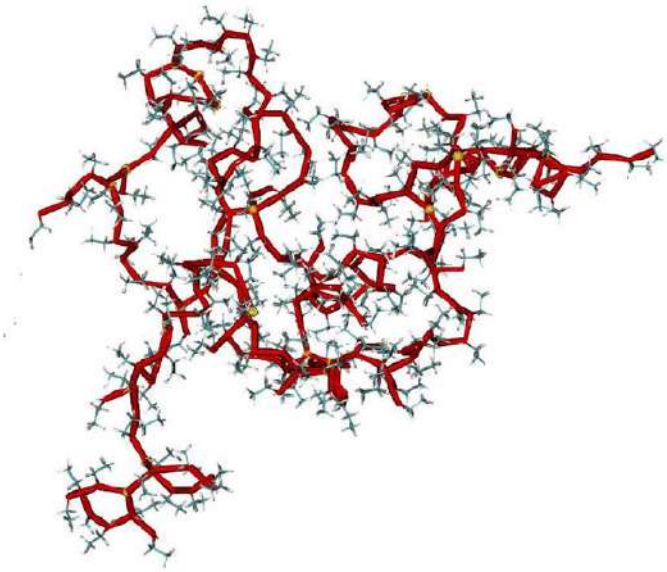

Fig. SI-2 Example of an ethanol-water chain. The connection line in red highlights the sequence of hydrogen bonds O-H—O which makes up the chain; the yellow circles point out the O atoms that belong to water molecules.

This means that, in the average, a water molecule instead of taking part in the water HB network, becomes an integrated part of the ethanol clusters (chains), and this happens more easily in the surface than in the bulk, as shown by the light blue lines in Figure SI-1. This is a key point. In fact, being close to the surface is a necessary condition for a molecule to evaporate, because the inelastic mean free path of molecules in a liquid is in the order of an intermolecular distance. The composition of the vapour phase will thus reflect the surface composition and this connects to the development of the partial vapour pressure with X, see Figure 1. On the macroscopic scale, the azeotropic behaviour of water/ethanol mixtures at high X is due to a reduction of the ethanol partial vapour pressure below that expected from Raoult's law.

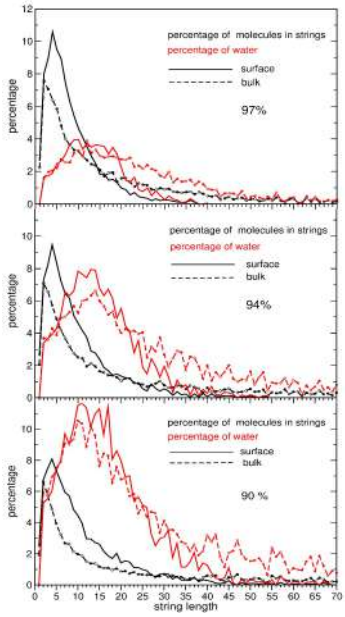

Fig. SI-3 The percentage probability (black lines) of finding chains of a given length, understood as the number of molecules forming the chain, as well as the percentage of water (red lines) present in a chain vs the length of the chain, for the bulk (dotted lines) and surface (solid lines) regions at various ethanol molar concentrations.

On the microscopic scale, we see that the reduction of the ethanol partial vapour pressure is caused by ethanol being partially replaced by water in the surface layer for X>85At the surface of mixtures with high ethanol concentration, the water molecules are practically isolated and "solvated" in ethanol chains and due to the reduced intermolecular interaction, in comparison with liquid water, can more easily evaporate.

Further analysis of the MD simulations reveals more details about the microscopic structure of the solution. At sufficiently high ethanol molar percentage X, ethanol molecules tend to form long chains, also branched and with rings, as exemplified in Figure SI-2.

The length of the chains found in the simulation sample depends on the definition adopted for the hydrogen bond, but, as will be seen, there is a clear difference between chains in the bulk and those on the surface. For a further characterization of the chains, we consider not only the length but also the composition, in terms of water molecules present in the chain. At the highest concentrations of ethanol, the water molecules have little chance of forming bonds between them, as confirmed by blue lines in Figure SI-1, and generally, they are instead inserted in the ethanol chains forming two hydrogen bonds, but also, occasionally, three or four (see Figure SI-2) resulting in branches.

We analyzed the length and composition of these chains, and

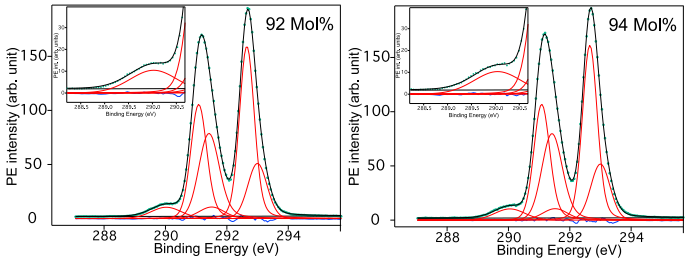

Fig. SI-4 Experimental spectra at 92 and 94 mol% together with the fitted lines. See text for details. We divided the analyser detector counts by 10^6.

how do they change with ethanol concentration; pointing out a difference between surface and bulk. Figure SI-3 shows the percentage probability of finding chains of a given length, understood as the number of molecules forming the chain, as well as the percentage of water present in a chain vs the length of the chain. The analysis was done for the bulk region and the surface region. From this, we can make some observations. It is clear that the shorter chains (<15 molecules) are present in both bulk and surface, with a clear preference for the surface, while the longer chains (> 40 molecules) are practically only present in the bulk. We also see that the relatively short chains (<15 molecules) have higher water content. This difference is even more pronounced for chains belonging to the surface region instead of the bulk region and is particularly evident for the 94% solution. Both sets of curves, therefore, seem to indicate that the insertion of water molecules in the ethanol chains is preferably carried out in relatively shorter chains and that these are more numerous in the surface region. This serves as an atomistic mechanism for the water molecules, having an increasing preference for the surface with the highest content in ethanol and then having more chances of evaporating, i.e. an apparent greater volatility.

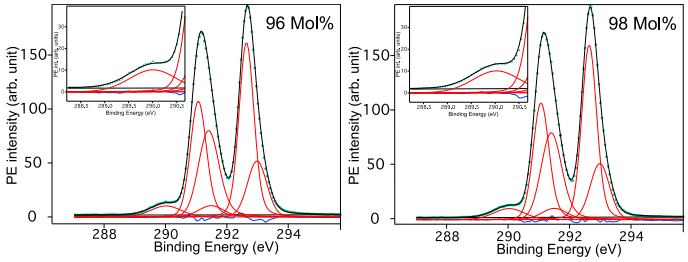

Fig. SI-5 Experimental spectra at 96 and 98 mol% together with the fitted lines. See text for details. We divided the analyser detector counts by 10^6.

2 Experimental details

2.1 Data analysis specifics

The spectra we show in Figure SI-4,SI-5,SI-6 plus the spectra two C1s Spectra shown in the main text taken at 85,90 mol %, Figure 5, are all nine experimental spectra we measured exclusive to this study. The data point below 15 mol % in Figure 2 are not critical to the present analysis and the original data turned out to have poor quality. We therefore used two data points at 15 and 9 mol % from present set of data to normalize previous data taken also with 360 eV at concentration 14.26, 8.59 down to 0.18 mol %. These low concentration data were taken from ref. [1].

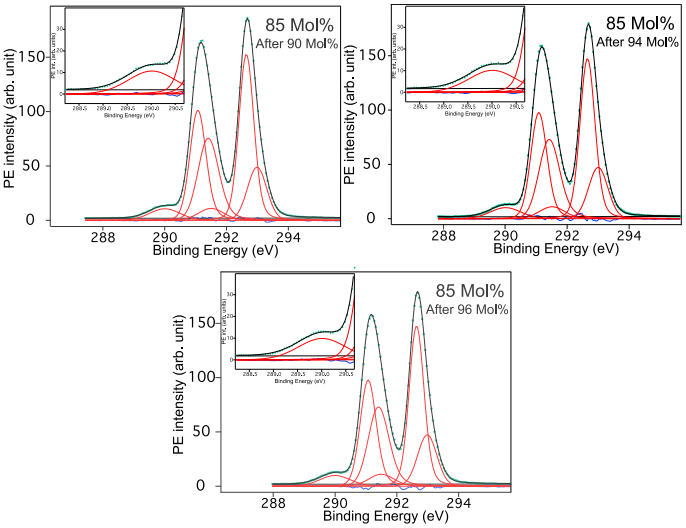

Fig. SI-6 Detail of the spectra and fitting procedure employed to obtain the area marked as black circles in Figure SI-8 at time "t" about 110 min ("85 mol% after 90 mol%") upper left. At time "t" about 240 min ("85 mol% after 94 mol%") upper right. At time "t" about 290 min ("85 mol% after 96 mol%") lower center. All spectra used in the normalisation are taken from the same sample with 85 mol% concentration of ethanol. We divided the analyser detector counts by 10^6.

We show in Figure 5 and Figures SI-4,SI-5, the fitted spectra between 85 and 98 mol % concentrations. In all spectra we also included an insert containing an amplified region of the C1s CH_3 liquid peak. As can be seen, the fitted line matches the experimental data closely. The areas of these peaks were used to compose Figure 2 because they show the smallest overlap with the gas phase peaks comprising of the four most intense curves in the main window. These are the most critical spectra exhibiting the relative decrease in ethanol C1s area above the azeotrope.

In Figure SI-7, we show the C1s XPS signal from the evaporated gas surrounding the liquid micro-jet. The black dotted line represents the ethanol vapor molar percentage according to Raoult's law. The curve resembles the ethanol vapor molar percentage shown in Figure 1, even though we did not build our setup with this type of measurement in mind. The consistency of the results with the known vapor pressure data shows that the X-ray cross-section with the micro-jet was constant and or properly corrected to high accuracy in this study. This result has an important consequence. It suggests the existence of a negligible spurious intensity "exchange" with the CH_3 liquid signals and the partially overlapping gas-phase peaks shown in Figure 5 and Figs. SI-4,SI-5.

Now we turn our attention on how we performed the final steps of the normalisation. The area from the ethanol intensity points 90, 92, 94, 96, 98 mol% were normalised by the number of sweeps and ring current. In order to reach a higher accuracy in the determination of the intensity variation of these points, we also recorded spectra at 85 mol% in between the high concentration points. This procedure is described in Figure SI-8.

The recorded spectra for normalisation purpose are the black

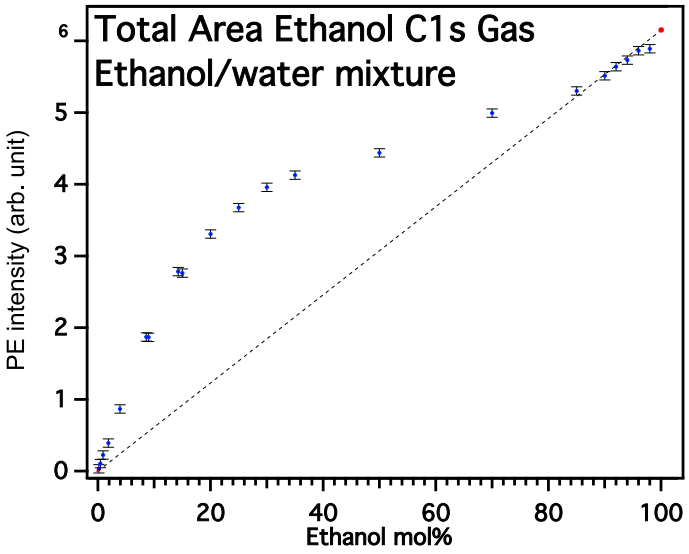

Fig. SI-7 Total C1s signal from the gas phase surrounding the liquid micro-jet. The dotted line reflects Raoult's law for ethanol/water mixture vapor. We divided the analyser detector counts by 10^6.

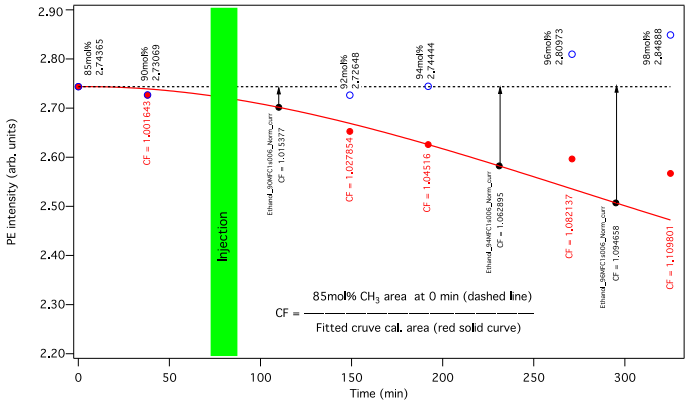

Fig. SI-8 Experimental details of the normalization procedure. We divided the analyser detector counts by 10^5.

points in the figure. Additionally, the spectra using concentration 85 mol % at time 0 was also used in the fitting procedure shown as solid red curve. Variations in the intensity of the 85 mol % points is a clear indication we needed to correct the intensity further. In order to obtain the best possible results, we plotted all the points against the time in minutes the experimental point was recorded. In this way, we could better correct for changes that would occur as a function of time. These changes could be due, for example, to a small misalignment between the X-ray beam and the micro-jet. Usually, this misalignment is more likely to occurs just after an electron injection in the synchrotron ring (the injection is also displayed as a green bar in Figure SI-8.

The normalised points by ring current and number of sweeps are marked as red dots in the figure. The correction factors according to the solid red curve are indicated as CF. The final corrected points are shown as open blue circles. Nearest to each point we wrote the final normalised intensity as well as the concentration representing the point in mol% .

The fitting procedure performed in order to obtain the areas representing the black points in SI-8 is presented in Figure SI-6.

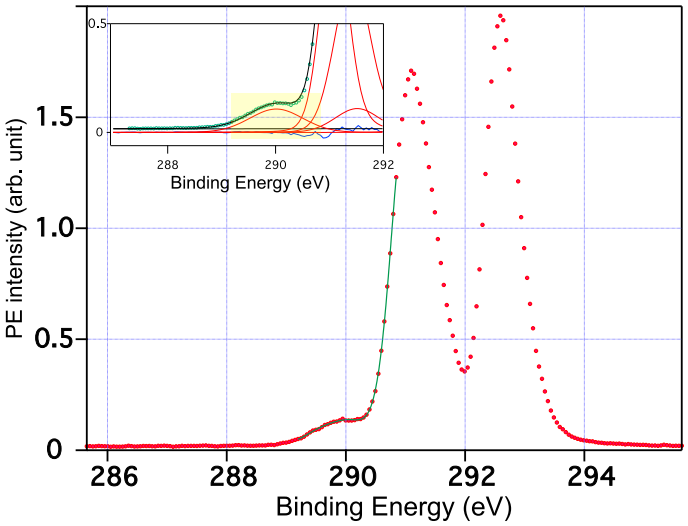

Fig. SI-9 Description of the error calculation procedure. We divided the analyser detector counts by 10^8.

2.2 Experimental uncertainties

Uncertainty evaluation is certainly critical to establish our results. With that in mind we create a procedure that could be followed step wise providing a clear meaning to what the uncertainty bar stands for. The first step was to find the best mean value of the experimental points. One way to achieve this is to subtract the real data from a fit where all links are removed and even extra Voigt are added so that the fitted line represents, as close as possible, the mean. This is a good strategy when we are interested in extracting the residue from the whole energy rage. In our case, we are specifically interested in the energy region around 290 eV delimited by the yellow square in the insert in the upper left corner in figure Figure SI-9. This is the region where the C1s CH_3 liquid peak is situated. It turned out we could obtain a good mean for this region by fitting a nine degree polynomial to the experimental points. The result of this fit is the solid green line also shown in the Figure SI-9 insert. For the above evaluation, we choose the spectra at 96 mol% where the effect we are studying is expected to be most pronounced. We subtracted this spectrum from the fitted polynomial obtaining our residue shown in Figure SI-10 left side.

As a next step, we used the function called "Wave Statistics" which is part of IGOR PRO 6.29 program to obtain the standard deviation (SD) of this residue. In the left side of Figure SI-10 we show a graphical evaluation of the residue SD. We produce an histogram from the residue in the right side. As a second step we fitted a Gaussian to the histogram. We show in the figure the SD (sigma). Following the electronic supplementary information (ESI) of ref.[2] a procedure was employed to obtain the best evaluation of the background noise in the energy regions where there is no signal, only noise. The criteria to establish the regions that can be considered free from signal is somewhat arbitrary. We used the first 10 points in the low binding energy side and the 5 points

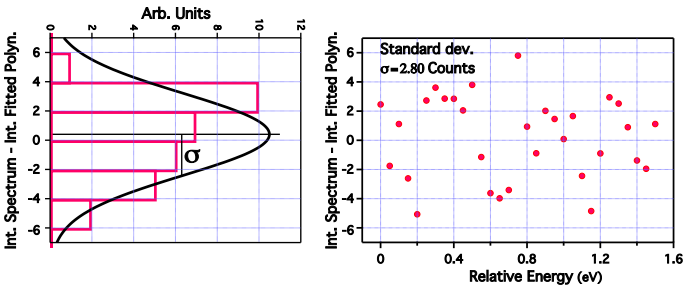

Fig. SI-10 In the right side We show the resulting points equal to the difference between the fitted polynomial and the 96 mol% spectra. The standard deviation is also shown for these points. We divided the analyser detector counts by 10^5.

in the high binding energy side to fit a linear background. Finally, the SD was used to produce an artificial data set by adding the SD to the original data to all the points with intensity larger than two sigma compared to the background. See an illustration of this procedure in ref.[2], ESI section "Calculation of Error Bars" figure S 3.

To this new artificial spectrum we include all the converged fitting curves. We run a new fit keeping the energy of the peaks fixed allowing the gas phase peaks to change only intensity and the liquid peaks to alter width and intensity. The new area of the fitted liquids peaks was used as the upper uncertainty bar while the lower error bar was symmetric.

For the final error calculation, we also included the uncertainty from the normalization spectra taken at 85mol%. This procedure generated the calculated error bar for the normalized areas.

Notes and references

1 R. R. Marinho, M.-M. Walz, V. Ekholm, G. Ohrwall, O. Bjorneholm and A. N. de Brito, *J. Phys. Chem. B*, 2017, **121**, 7916–7923.

2 G. Gopakumar, I. Unger, C.-M. Saak, G. Öhrwall, A. Naves de Brito, T. C. Rizuti da Rocha, C. Nicolas, C. Caleman and O. Björneholm, *Environ. Sci.: Atmos.*, 2022, **2**, 441–448.