Molecular Conformation of Linear Alkane Molecules: from gas phase to bulk water through the interface.

Ezequiel L. Murina,[†,‡] Roberto Fernández-Prini,[†,‡] and Claudio Pastorino[*,¶]

†*Departamento de Fisicoquímica de Fluidos, CAC-CNEA, Buenos Aires, Argentina*

‡*INQUIMAE, FCEN, UBA/CONICET, Buenos Aires, Argentina*

¶*Departamento de Física de la Materia Condensada, CAC-CNEA/CONICET, Buenos Aires, Argentina*

E-mail: pastor@cnea.gov.ar

Phone: +54 11 6772 7105

Abstract

We studied the behavior of long chain alkanes (LCA) as they were transferred from gas to bulk water, through the hydrophobic-aqueous interface. These systems were studied using umbrella sampling molecular dynamics simulation and we have calculated properties like free energy profiles, molecular orientation and radius of gyration of the LCA molecules. The results show changes in conformation of the solutes along the path; in bulk water their conformations up to dodecane are mainly extended. However, larger alkanes like eicosane present a more stable collapsed conformation as they approach bulk water. We have also found that the larger alkanes extend appreciably when partially immersed in the interface as shown by their average radius of gyration and have characterized the more probable configurations. The results obtained are of interest for the study of biomatter processes requiring the transfer of hydrophobic matter, especially chain-like molecules like LCA's, from gas to bulk aqueous systems through the interface.

INTRODUCTION.

The importance of hydrophobic matter in the presence of water is still a field of very active scientific research in Phys. Chem. as well as in Bio. Phys. Chem. This is due to the large number of studied systems , which are of interest for technology, as well as those related to biomatter processes, where it is frequent to encounter many hydrophobic molecules which are in contact with water, or in general with aqueous systems. Also the importance of hydrophobic moieties for stabilizing protein structures and their structurally stable conformation, thus permitting enzymatic processes to occur, is accepted; this is especially valid for membrane proteins[1]. Hydrophobic molecules penetrating through the water-hydrophobic interface is an important process to stabilize the protein conformation[2]; hence the protein's stable structure results from hydrophobic interactions[3]. Also they

are important in other cogent fields, like contamination remediation; this was illustrated by Hua and Wang[4] who referred to methods for remedial of oil spills (in earth or water) considering that there is a need for more information about interfacial (hydrophobic-hydrophilic phases) effects. This relates to the linear chain n-alkanes (LCAs) entering that interface and how this evolves as they are transferred to water where they may be processed by bacteria. This is in the same context as we studied here. One of the model hydrocarbons they used was n-hexadecane.

In a previous publication we have described the behavior of a methane-water system[5]. Our reported results agreed with those of other studies which described the behavior of short chain alkanes in contact with water. Other works have also studied this system by computer simulations[6–8] or density functional theory[9]. In this paper we studied butane (C-4), octane (C-8), dodecane (C-12), hexadecane (C-16) and eicosane (C-20) as solutes, to study the conformations of LCAs when entering into the alkane-water interface for systems of typical size computer simulations and using extensive simulation times. The low solubility in water of these molecules makes necessary the adoption of techniques in order to enhance the sampling of rare configurations. We have carried out molecular dynamics simulations with the umbrella sampling approach MDUS[10,11], fixing the center of mass of the molecules at different points in the water-vapor interface path from gas to bulk water. LCAs of similar length in bulk water have been studied by many authors. Ferguson et al.[12] have reported results of a very interesting study of LCAs– which shows that bulk water has a negligible effect– in the conformation adopted by LCAs, compared with that in the gas phase, at least up to C-20. This agrees with the careful experimental study carried out by Tolls et al.[13], up to C-15, which is consistent with the extended chain micro-configuration as the most stable one. Although the advances in this matter, there is a lack of publications which describe in detail the LCAs' conformations at the water-vapour interface. The present work covers, for the first time at the best of our

knowledge, the study of various properties of LCAs at different relevant positions of the water-vapor interface. They reveal conformational and orientational characteristics of the LCA molecules from the gaseous phase to bulk water..

Simulation Technique

We performed molecular dynamics simulations (MD) using the Gromacs 4.6.7[14] package and then analyzed the data with home-made codes that implement the MDAnalysis object-oriented library[15]. The simulated system consists of a rectangular box having a square base of 16 nm^2 and a height of 22 nm and we applyed periodic boundary conditions in the three cartesian directions. In the central region of the box, a volume of 4 nm $\times$ 4 nm $\times$ 10 nm was filled with 5 367 SPC/E[16] water molecules forming a slab, whose center of mass was fixed in the middle of the simulation box. In the bulk water region, the experimental water density value at 298 K was obtained. Atomic distance in water molecules was restrained by means of the Settle algorithm[17]. The other component of the system was a single LCA molecule as solute. The studied LCAs were butane (C4), octane (C8), dodecane (C12), hexadecane (C16) and eicosane (C20). All of them were described with the TraPPE-UA force field[18]. This force field was chosen due to its simplicity, computational efficiency, high accuracy and its use in similar systems[19,20]. Basically, it is an united-atom representation for alkyl segments and simple Lennard-Jones and Coulombic potentials. The initial configuration for each simulation consisted of a thermalized water slab and a single LCA placed in an all-trans conformation parallel to Gibbs water-vapor dividing surface (GS). Then, a MD equilibration was carried out during 1 ns. We find no influence of the initial configuration of the alkane in the calculations, in particular at interface. By computing the time auto-correlation function of both the radius of gyration and the orientational order parameter, we observe correlation times of roughly 100 ps and

5 ps, respectively . We worked in the canonical ensemble and fixed the temperature by means of a Langevin thermostat. This thermostat allows for a correct velocity distribution of a single molecule and adds a dissipative force and a stochastic force to the conservative forces among particles; the thermostat acts on each atom of the system. The dissipative force on particle i is given by $\mathbf{F}_i{}^{\mathrm{D}} = -\gamma \mathbf{v}_i$, where γ is the friction coefficient and $\mathbf{v}_i$ the particle velocity. The random force, $\mathbf{F}_i^{\mathrm{R}}$, has zero mean value and its variance satisfies the following equation[21]:

$$\langle F_{i\mu}^{\mathrm{R}}(t) F_{j\nu}^{\mathrm{R}}(t') \rangle = 2\gamma T k_B \delta_{ij} \delta_{\mu\nu} \delta(t - t') \,, \tag{1}$$

where the indices i and j label particles, μ and ν are the cartesian components, and T is the temperature at which the system is simulated. We note that this thermostat was already used in related work[22]. In practice, this means integrating Langevin equations for each degree of freedom instead of the usual Newton equations, characteristic of the standard molecular dynamics scheme. In this way, the temperature is fixed to a desired value and the single thermostat parameter to be set is the friction constant. The inverse of friction constant was set to 5 ps, as implemented in GROMACS. This value is large compared to the characteristic time scales present in the system, so the stochastic-dynamics scheme behaves as MD with stochastic[21] temperature-coupling (NVT-ensemble with temperature fixed at 298 K).

The leap-frog algorithm was used to integrate the equations of motion setting a time step of 1 fs. Electrostatic interactions were calculated implementing the Particle Mesh Ewald method[23,24] with a mesh spaced 0.12 nm and a cutoff of 1.4 nm in real space. The protocol for Umbrella Sampling (US) technique was applied as follows[11]. The perpendicular direction to the square base of the simulation box (and to the water-vapor interface) was taken as reaction coordinate (see Fig. 1). The center of mass of a single LCA molecule was restrained *via* a harmonic biasing potential at different z values. These are called

umbrella windows (UW) with a step separation of 0.1 nm. We used a simulation sampling time of 40ns for each UW, storing frames every 0.1 ps. The harmonic biasing potential has a spring constant value of 1 600 kJmol^{-1}nm^{-2}. The sampled region extended 1.5 nm in z width at each side of the GS . The WHAM method was applied to merge the data of the UW into an unbiased free energy profile[25]. In addition to these simulations, a few independent simulations were carried out in bulk phases, setting a longer sampling time of 320 ns for the bulk liquid phase and 1 000 ns for the gas phase, to further study the configurations of the LCAs alkanes in bulk phases.

In the umbrella sampling technique the translational movement of the center of mass along the z direction is biased to fix this reaction coordinate in a chosen position (possibly an unlikely one). However, the rest of the degrees of freedom of the LCA are not affected by the umbrella potential. Therefore, the conformations of the molecules are not biased by the constraint of the center of mass and the molecular conformations in a given position of the interface can be correctly sampled. This lets us measure conformation and orientation properties of the LCAs in each window that we used to calculate the Free Energy Profiles (FEP). It must be taken into account however, that the longer molecules need much more time to cover the available configuration space and we needed to extend appreciably the simulation time, usually up to 40 ns. In the supplemental material (SM), Fig. S1, we show that with our appoach we overcome the kinetic traps for LCAs which remain for a considerable simulation time as reported by Ferguson et al.[12].

Results and discussion

The low solubility of LCAs leads to poor sampling in regions of high water density. To overcome this problem, we have implemented the US protocol and have been able to study single LCAs' behavior as they go through the interface from gas bulk phase to liquid bulk

phase. First, thermodynamics preferences of the LCAs have been evaluated by means of FEP profiles. Then, we have explored in each UW their conformation and orientation with respect to the GS.

FEP: adsorption layer and strong repulsion to enter the interface

We studied the free energy profile FEP(z) as a function of the position of the molecule, relative to the vapor-liquid interface. To this end, we fixed the molecule at different z positions, covering from bulk vapor phase up to bulk liquid phase, analyzing intermediate positions in the interface (1 nm from both sides of the GS). In each UW a local FEP$_i(z)$ was measured using the Equation:

$$\mathrm{FEP}_i(z) = -k_B T \ln P^*(z) - w_i'(z) + C_i$$

where P^* is the biased probability of finding the center of mass of LCA at the i-th UW, w_i' is the bias umbrella potential energy and C_i is a local relevant constant with an implicit dependency on w_i'[10]. The global FEP profile was generated after determining w_i' and suitable values of C_i for all the UW. This was done by applying the WHAM method[25].

The FEPs are shown in Fig. 2. They present a minimum values close to $z = -0.2$ nm, which indicates a thermodynamic preference of the LCAs to adsorb on the interface. The adsorption at the interface is related to the attractive van der Waals water-solute interaction the LCAs feel, as they go from gas bulk to the interface, which makes LCAs adopt a tangential orientation with respect to the GS exposing all the carbons to the interface. For this reason, the minima of the FEP curves decrease with the chain size and are shifted according to the size of the molecule. The minimum is reached in the proximity of the GS, where the attractive interaction is balanced by a repulsive contribution to the FEP. This contribution is related to the work necessary to break partially the surrounded

water structure. Repulsion overcomes the van der Waals interaction as the LCA goes deeper into water and the FEP curve increases[26].

We also analysed the FEP profiles scaled by the number of carbons of the corresponding LCA (see Fig. S2 of the SP). We observe that the free energy per carbon atom at bulk solution decreases with the alkane length, saturating for C12. We attribute this to the fact that the $-CH_3$ groups contribute more prominently to the hydrophobic effect in water than the $-CH_2$ groups. The reported value of the contribution to the standard Gibbs hydration free energy for $-CH_3$ groups is $3.63\,\mathrm{kJ\,mol^{-1}}$, while for $-CH_2$ group is $0.72\,\mathrm{kJ\,mol^{-1}}$[27]. As regards the behavior in the adsorption region, C4 presents lower free energy. This may be due to the fact that longer LCAs can explore a wider range of conformations in which some carbons experiment a weaker interaction with water.

The mean radii of gyration in gas bulk phase are depicted in the inset of Fig. 2, to give a reference of the molecular size. The density profile (dotted line in Fig. 2) reachs a constant bulk value for the higher z-values. We explored z-coordinate values well beyond the molecular size inside the water layer to insure that the molecules are completely immersed in bulk water. This is also supported by the constant value observed for the FEP for all the cases at high z.

For the FEPs of the LCAs the difference in energy from the bulk phase increases monotonously with the alkane's size as expected, since it is more difficult dissolving them and their solubility will decrease with the size of the alkane. Nevertheless the difference of free energy between bulk gas and bulk water between one LCA and another having four $-CH_2$ more is 7.75 ± 1.5 kJ/mol. This increase in the free energy value is due to the increasing difficulty of alkanes to enter bulk water for larger chain lengths and therefore a higher number of hydrophobic $-CH_2$ units. .

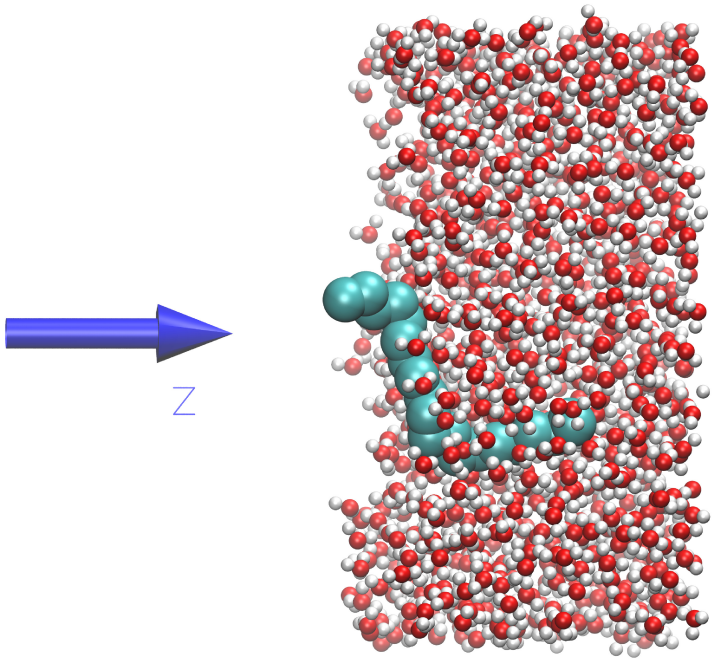

Figure 1: Snapshot of one of the studied systems showing a fragment of the water slab close to the interface. The water layer is located in the center of the simulations box and the LCA molecule is fixed at different points in the interface, from bulk vapor phase towards water phase. The z-coordinate is taken to be normal to the interface. Only half of the water slab is shown. Periodic boundary conditions are set in x and y directions. A molecule of C20 is shown in this example. Vector $\boldsymbol{r}$ is in the direction of the line that links the methyl groups (end to end distance)

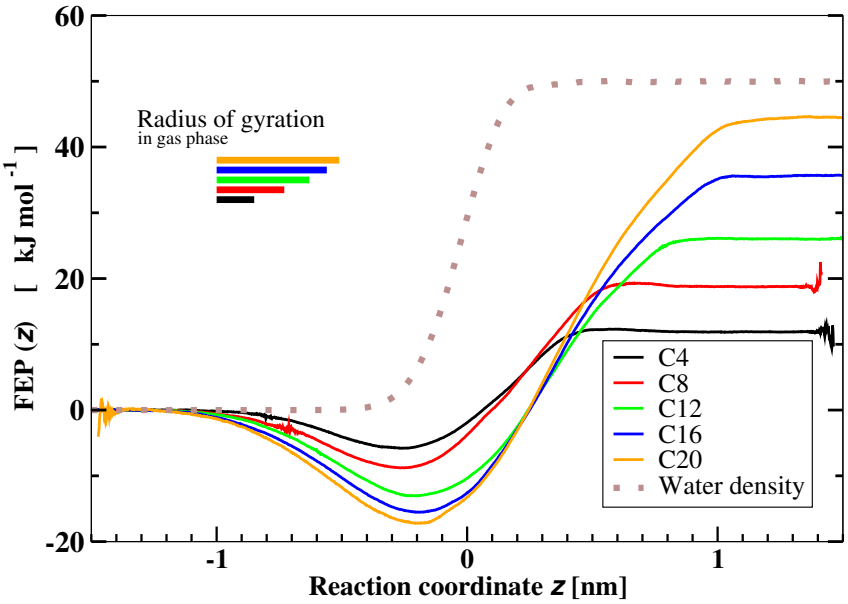

Figure 2: Free energy profile FEP(z) through the gas-liquid interface for all the studied molecules. The water density, in units of $\mathrm{kg\,m^{-3}}$ times 50 divided by water bulk density, is also shown to provide a reference in the reaction coordinate z, perpendicular to the interface (dotted line). The origin of coordinates in z is set at the GS. The FEPs are taken to be zero at the bulk gas phase. The inset shows the length scale given by the radius of gyration of the molecules in gas bulk phase. An analysis of the criterion of convergencean errors is included in the SM.

Orientation and conformations of LCA molecules

In the US protocol we implemented, the center of mass of a single LCA molecule was fixed at different distances from the GS. We exploited this to analyze molecular orientations and conformations, as a function of the center of mass position of the molecule from gas phase to bulk water through the interface.

Alkane Orientation

The orientation of the alkanes with respect to the GS was characterized by the order parameter $\lambda(z)$ defined by:

$$\lambda(z) = \frac{3 \cdot < \boldsymbol{r} \cdot \boldsymbol{z} >^2 - 1}{2}$$

where $< \boldsymbol{r} \cdot \boldsymbol{z} >$ is the time-averaged scalar product over the whole MDUS simulation, $\boldsymbol{r}$ is the vector in the direction of the end to end distance of the LCA, and $\boldsymbol{z}$ is the vector depicted in Fig. 1.

Additionally, we also computed an order parameter $\lambda^{'}$, using a principal axis of inertia $\boldsymbol{r'}$, associated with the minor eigenvalue of the inertia moment. The curves for λ and $\lambda^{'}$ are coincident all along the z-axis, except in zones 3 and 4, where the curves differ negligibly, *i.e.* within the statistical error. This justifies the use of the line that links the methyl groups as a suitable and simple definition of the orientation of the solute. A time average of λ was calculated for each UW and then the full profile was obtained as a function of z. This is shown in Fig. 3 for all the molecules studied in this work.

The order parameter profile is qualitatively similar to that obtained for the study of the interface between water and hexane by Chapman and coworkers[9]. We define five regions, labeled in Fig. 3, to analyze different behaviors: three interfacial regions (2,3,4), in which interfacial effects are present, and two bulk-regions (1,5), without interfacial

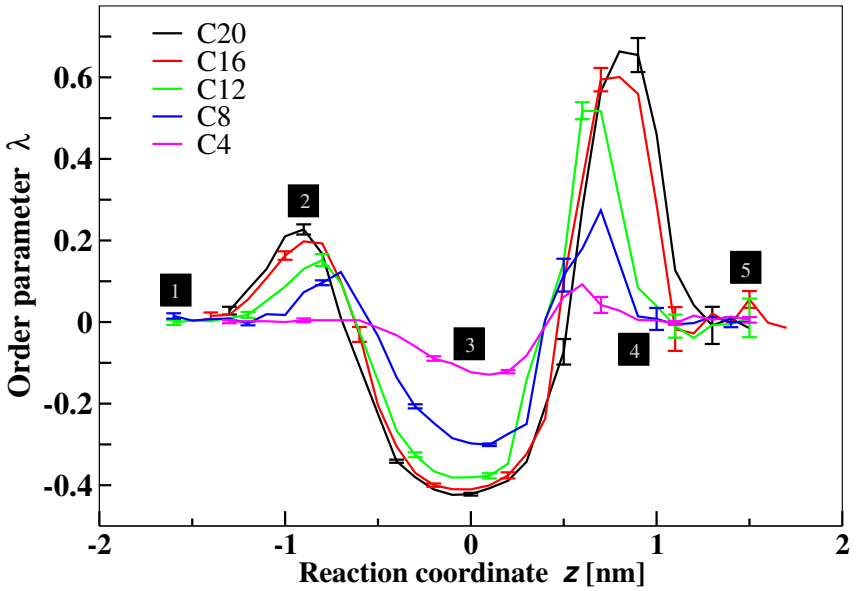

Figure 3: Orientational order parameter λ for the five studied molecules across the interface. The numbers indicate relevant regions of the interface with a distinctive behavior. The error bars in each curve are shown for some points.

effects.

Region 1 corresponds to bulk vapor phase and region 5 to bulk liquid phase. In both regions, LCAs have a zero-value order parameter due to a random orientation of the molecules in isotropic bulk phases, as expected.

In region 2, λ reachs a local maximum, that increases with the size of the LCAs. This maximum is attributed to the onset of the van der Waals attraction, at roughly 1 nm from GS, between H_2O molecules and the carbons of one chain end of the organic solute that lies close to water surface. The alkane-water interactions force the solute to rise partially towards a more vertical orientation to the interface, and the carbons of the opposite chain end of the LCA, that lie further away from water surface, do not feel the water attraction. For longer alkane sizes, the maximun value of λ increases due to the fact that more carbons can lie close to the water surface and solutes adopt a more pronounced normal orientation.

In region 3, since water density progressively increases, in order to accomodate the

organic chain as the LCA approachs water toward region 4, it will be increasingly necessary to overcome the repulsive water-solute interaction, arising from the tendency of water to conserve hydrogen bonding structure. LCAs lie down, tending to adopt a parallel position to the GS, and λ becomes negative. The λ profile in this region is similar to the one found for *n*-hexane/water interface[8].

Region 4 is characterized by a steep increase in λ, leading the alkane molecules to orient perpendicular to the GS. This is indeed an interesting region where orientational interfacial effects are important, before the solute is completely disolved in bulk water. The random orientation characteristic of bulk regions for the shortest LCA is attained at a distance of roughly 0.8 nm from GS towards bulk water. This distance is in agreement with the one estimated for the complete disolution of methane[5].

Molecular radius of gyration

Taking into account the polymeric nature of LCAs, the radius of gyration becomes a suitable quantity to characterize the longitudinal extension of the solute chain. It is defined by:

$$R_g^2(z) = \frac{1}{M} \sum m_i \langle (\mathbf{r_i} - R_{cm}(z))^2 \rangle \,,$$

where $R_{cm}(z)$ is the position vector of the center of mass at a distance z from GS, m_i and r_i are the mass and the position vector of the i-th residue, and M the molecular weight of the LCA, respectively. A time average of $R_g(z)$ was calculated for each UW and then the full profile was obtained as a function of z.

The curves shown in Fig. 4 confirmed the absence of any water effect over the longitudinal extension of the chains shorter than C12 in interface regions and bulk water. From C12 on, there are two effects of water on the solute. The first one appears in the interface region 4, around $z = 0.75\,\mathrm{nm}$ in the plot of Fig 4, and consists in a chain stretching of the

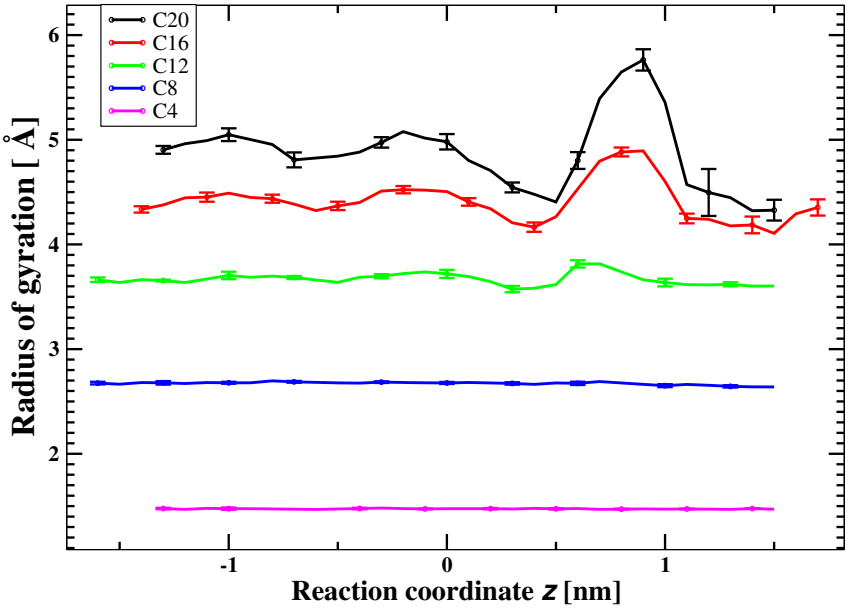

Figure 4: Radius of gyration profiles across the interface for the five studied molecules. The origin for the reaction coordinate z is located at the GS. Error bars are shown only for some points on each curve.

LCAs. Taking as reference the value of $R_g(z)$, measured in the gas phase far from GS, the stretching is around 4%, 13% and 18% for C12, C16 and C20, respectively. From the comparison of Fig 3 and Fig 4, the positions of the maxima of $R_g(z)$ for each LCA are in agreement with the those observed (*vide infra*) for λ in region 4. In the case of C20, there is also a decrease of 11% of the value of $R_g(z)$ in bulk water due a partial collapse of the alkane which will be describe in detail in the following section. For these alkanes, the second effect of the water consists in a small oscillation in the value of $R_g(z)$, which extends along the gas phase up to $z = 0.5\,\text{nm}$, towards solution and whose amplitude increases with molecular size.

Probability distribution of radius of gyration

As the value of R_g decreases, the multiplicity of molecular conformations with the same value of R_g increases. A more suitable quantity to obtain information about the confor-

mations adopted by the alkanes is the distribution of values of R_g taken on a complete production MD run. In order to have proper statistics to enable accurate measurements of this distribution, the total simulation time was set to 40 ns for each UW. The data can be presented either as free energy landcapes[12,28] or as probability distributions[19,22] parametrized by R_g. We used the last type of presentation. Alternatively, the free energy landscapes are shown in Fig S4 of SM.

The probability distributions for the interfacial regions 2, 3 and 4 are shown in Fig. 5, while that of bulk vapor and water regions are shown in Fig. 6. The well that separates the two peaks in C4 arises from geometric arguments reported by Ferguson et al.[12] and corresponds to a free energy barrier to and from an all-trans conformation. That barrier is also present for C8 and C12, but their heights decrease and their locations are shifted to greater values of R_g, for increasing chain lengths. This is consequence of the increment in the number of accessible conformations with the same R_g for longer LCA molecules. It can also be observed a less deep well for C8 at $R_g = 2.9$ Å and a weak trace of it for C12 at $R_g = 4.5$ Å.

The hydrophobic interaction will drive the conformations adopted by an LCA. It has been reported that the plasticity of bulk water forms structures which, conserving hydrogen bonds, are able to accomodate long chain alkanes (16 carbon or even more) in extended conformations[19,29,30]. On the other hand, in a previous work we described the surface water as a region with large spontaneous density fluctuations[5]. If an alkane does not perturb the hydrogen bonded network of bulk water, it will not affect the structure of a lower water density region with large spontaneous fluctuations, as the surface water. This explains the similarity of the probability distributions in all regions for the three shortest LCAs studied: C4, C8 and C12. However, in the case of C12 a slight enhacement of longer values of R_g appears in interfacial regions 2, 3 and 4, in comparison with the bulk region 1. That enhancement progressively increases from region 2 to region 4 and

correlates with the orientational transition of the LCA from a tangential conformation to a normal one, as shown in Fig 3. This is also in agreement with the stretching observed in Fig. 4. For C16 and C20 it becomes so strong that it modifies the probability distribution from region 3 to 4. In the case of C20 the change is staggering: in region 2 and 3 there are mostly conformations with $R_g \leq 5$ Å while in region 4 almost all conformations present values in the interval $R_g \geq 5$ Å.

According to Ferguson et al.[12], the effect of the solvent on the LCAs consists in the appearance of a free energy barrier of heigth $\sim k_B T$ from C12 onwards and, additionally, in the thermodynamic destabilization of conformations with longer R_g from C16 onwards. In the energy landscapes in Fig. S4 of the SM, we confirmed the appearance of that free energy barrier for C16 and C20. In this sense, our results agree with the experimental data reported up to C15[13]. However, we found for C20 that the conformations with lower R_g values are the most stable in bulk water, as indicated by the depth of the well around $R_g = 3.8$ Å. Additionally, we observed that the mean value of R_g is smaller than the value at gas phase. That more closed molecular configuration may allows for a partial screening of carbons located in the internal region of the molecule which should be, in average, less exposed to water than the external carbons, decreasing the overall hydrophobic effect of the molecule. This small differences with Ferguson et al.[12] could arise from the REMD approach they implemented to overcome the structural hysterisis between extended and compact chain conformations. In the present work, as explained in Section , we performed MDUS simulations for each UW during 40 ns. This lets us to access to a sufficient sampling of the whole configuration phase space, avoiding any effect related to the structural hysterisis.

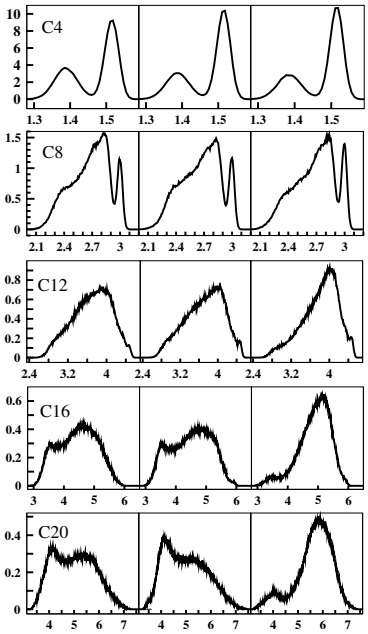

Figure 5: Probability distributions (percent) of the radius of gyration (Å) sampled by the five studied molecules in three characteristic interfacial regions (regions 2, 3 and 4 from left to right in Fig. 3).

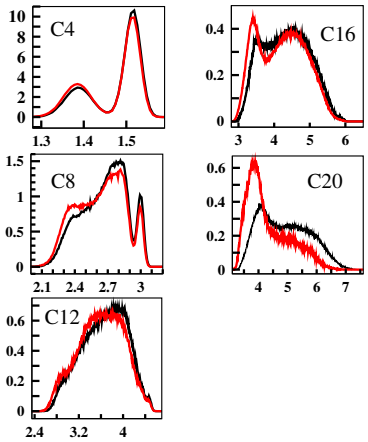

Figure 6: Comparison of the probability distributions (percent) of radius of gyration (Å) in bulk vapor (black curves) and bulk water (red curves) for the five studied molecules.

Visual inspection of the alkanes' conformations

One can obtain a better description of the probability distribution in interfacial regions by means of visual inspection of the MDUS molecular trajectories. In Fig. 7 representative

conformations adopted by C20 in each region are illustrated. A video included in the as SM provide an insight into representative LCAs' conformations at interface region 4. Interestingly, C20 has very frequently between four and seven carbons of one of its end-

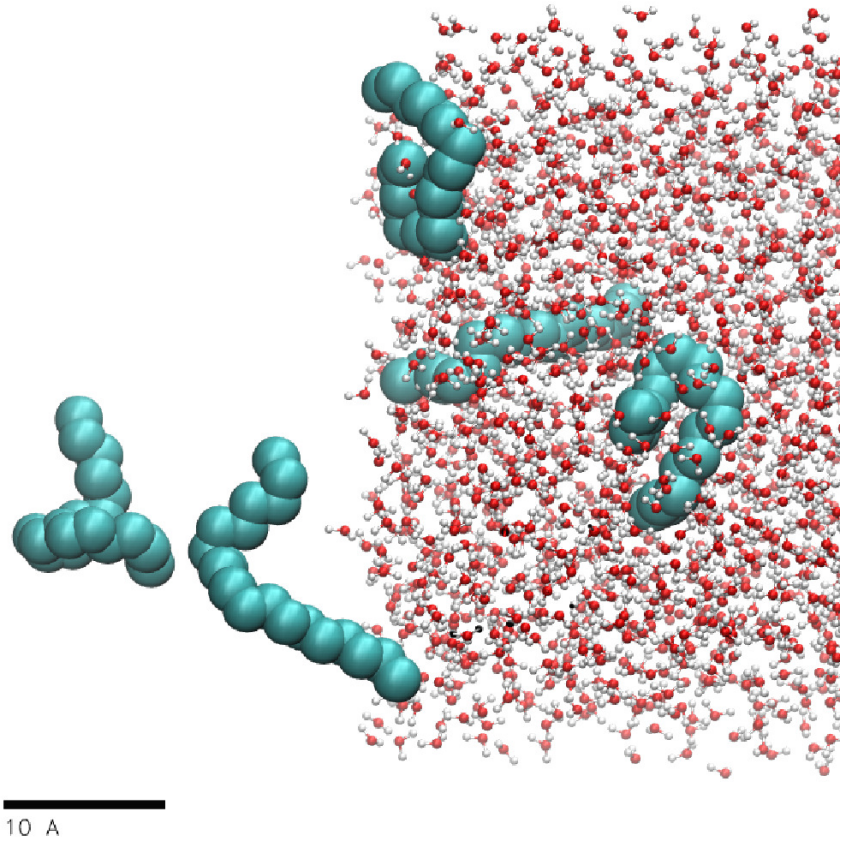

Figure 7: Superposition of representative snapshots for C20 corresponding to each region indicated in Fig. 3. LCAs in cyan color, water oxygens in red and hydrogens in white. Water is depicted only for the LCA in the bulk liquid phase.

chains adsorbed on the water surface. Those carbons adopt an extended conformation with a tangential orientation to the GS, while the rest of the chain, in contact with higher water density, adopts a quasi-trans conformation with a normal orientation to the GS. The conformation of the whole chain for C20 results, in most cases, in a L-shaped conformation having its shortest segment adsorbed (see Fig. 7). In the case of C16, due to its shorter chain length, only one or two carbons remain adsorbed, while roughly 6 carbons at the middle of the chain adopt a quasi-trans conformation (normal orientation to GS). For this

LCA the end chain, which is in contact with higher water density, adopts preferentially collapsed conformations. This effect must be caused by the water repulsion, before C16's end-chain reachs the region of constant free energy values, in the profile shown in Fig 2. That collapse is absent for C20, due to the fact that last carbons in its end-chain do get deeper into liquid phase, where free energy has a constant value, since water repulsion has been overcome. For C12, C8 and C4, the visualization does not show any appreciable changes in the extended conformations they adopt.

We have also checked the visualizations of the MDUS trajectories of C16 and C20 in region 1 and 5 (bulk phases) and we observed that, for R_g values close to the global minimum of energy landscapes, LCAs adopt the U-shaped conformations as reported by Sun et al.[19] for octadecane.

Conclusions

This study provides insight in the dissolution process of LCA molecules through a water-vapor surface, from bulk gas to bulk water. We describe the conformations that LCAs adopt at a given distance from the GS, by fixing the center of mass of the molecule at different locations of the reaction coordinate and extensively computing the adopted orientations and conformations. We used the pecularities of orientations of the molecules fixed at different points of the water-vapor interface, to define representative zones of the interface. In these zones we carefully computed the mean radius of gyration and probability distributions of the radius of gyration, to characterize the interfacial configurations and compare with the behavior at liquid and vapor bulks for various molecular chain lengths.

The first important interaction between LCAs and water is found to be of van der Waals type, and appears at a distance of 1nm from the GS. One end-chain of the alkane

gets adsorbed on water surface giving to the solute a slightly normal orientation to the GS. That orientation increases with the size of the LCA. Closer to water surface, LCAs lie adsorbed with a parallel orientation to the GS. The longer the alkane, the more stable the adsorption is. The shorter LCAs, C4 and C8, do not experiment any conformational change in the interfacial regions. However, from C12 on we observe changes of conformation in the interfacial region of high water density. This is in agreement with an increment in the normal orientation to the GS. A stretching of the alkane chain is revealed in a peak of the profile of radius of gyration and an increment of the probability density for longer radius of gyration. We confirm by visual inspection the existence of a characteristic L-shaped conformations for C20, at the global maxima of the order parameter.

For long LCA's in bulk phases there are relatively few experimental datato compare with, due to the difficulty implied by their very low solubility. However, our simulation results coincide with the data of Tolls et al.[13] up to C15, which carried out an extremely careful and precise work. Up to C15, their solubility agrees with the values of the group contribution method and with data from Ferguson et al.[12], that favor an extended conformation as the most stable, with a slight tendency to have higher solubility than those determined through simulation. They are also somewhat closer to the contribution group results. Our results show that up to C12 in bulk water there is practically no effect of the solvent on the most stable conformations adopted by the LCA molecules. On the other hand, for C20 we observed a clear preponderance (more stable) of the collapsed conformation in bulk waters and, as a consequence, the mean value of the radius of gyration is smaller than the gas bulk phase value. This is in agreement with the reduction of the free energy per carbon atom with the alkane length, we have found at bulk solution.

Acknowledgment

The authors are grateful to Dr. M. Laura Japas for useful discussions about the contents of this work. E.L.M. received a fellowship from the National Scientific and Technical Research Council (CONICET). C.P. and R.F-P. are members of the Career Research of CONICET. Financial support was obtained from the ANPCyT (Programs PICT-2011-2378, PICT-E 134-2014), CONICET (112 2015-0100417) and CNEA (INN-CNEA 2011).

References

(1) Jayaraj, V.; Suhanya, R.; Vijayasarathy, M.; Anandagopu, P.; Rajasekaran, E. Role of large hydrophobic residues in proteins. *Bioinformation* **2009**, *3*, 409–412.

(2) von Heijne, G. Membrane protein structure prediction. *Journal of Molecular Biology* **1992**, *225*, 487 – 494.

(3) Vacha, R.; Slavicek, P.; Mucha, M.; Finlayson-Pitts, B. J.; Jungwirth, P. Adsorption of Atmospherically Relevant Gases at the Air/Water Interface:Free Energy Profiles of Aqueous Solvation of N2, O2, O3, OH, H2O, HO2, and H2O2. *The Journal of Physical Chemistry A* **2004**, *108*, 11573–11579.

(4) Hua, F.; Wang, H. Uptake and trans-membrane transport of petroleum hydrocarbons by microorganisms. *Biotechnol Biotechnol Equip* **2014**, *28*, 165–175.

(5) Murina, E. L.; Pastorino, C.; Fernández-Prini, R. Entrance dynamics of {CH4} molecules through a methane-water interface. *Chemical Physics Letters* **2015**, *637*, 13 – 17.

(6) Ghoufi, A.; Malfreyt, P. Numerical evidence of the formation of a thin microscopic

film of methane at the water surface: a free energy calculation. *Phys. Chem. Chem. Phys.* **2010**, *12*, 5203–5205.

(7) Biscay, F.; Ghoufi, A.; Malfreyt, P. Adsorption of n-alkane vapours at the water surface. *Phys. Chem. Chem. Phys.* **2011**, *13*, 11308–11316.

(8) Nicolas, J.; de Souza, N. Molecular dynamics study of the n-hexane-water interface: towards a better understanding of the liquid-liquid interfacial broadening. *The Journal of chemical physics* **2004**, *120*, 2464–2469.

(9) Marshall, B. D.; Cox, K. R.; Chapman, W. G. A Classical Density Functional Theory Study of the Neat n-Alkane/Water Interface. *The Journal of Physical Chemistry C* **2012**, *116*, 17641–17649.

(10) Kästner, J. Umbrella sampling. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011**, *1*, 932–942.

(11) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*; Academic Press, 2002.

(12) Ferguson, A. L.; Debenedetti, P. G.; Panagiotopoulos, A. Z. Solubility and Molecular Conformations of n-Alkane Chains in Water. *The Journal of Physical Chemistry B* **2009**, *113*, 6405–6414, PMID: 19361179.

(13) Tolls, J.; van Dijk, J.; Verbruggen, E. J. M.; Hermens, J. L. M.; Loeprecht, B.; Schüürmann, G. Aqueous Solubility-Molecular Size Relationships: A Mechanistic Case Study Using C10 to C19 Alkanes. *The Journal of Physical Chemistry A* **2002**, *106*, 2760–2765.

(14) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms

for Highly Efficient, Load-Balanced and Scalable Molecular Simulation. *Journal of Chemical Theory and Computation* **2008**, *4*, 435–447.

(15) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *Journal of Computational Chemistry* **2011**, *32*, 2319–2327.

(16) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The missing term in effective pair potentials. *The Journal of Physical Chemistry* **1987**, *91*, 6269–6271.

(17) Miyamoto, S.; Kollman, P. A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of Computational Chemistry* **1992**, *13*, 952–962.

(18) Martin, M. G.; Siepmann, J. I. Transferable Potentials for Phase Equilibria. 1. United-Atom Description of n-Alkanes. *The Journal of Physical Chemistry B* **1998**, *102*, 2569–2577.

(19) Sun, L.; Siepmann, J. I.; Schure, M. R. Conformation and Solvation Structure for an Isolated n-Octadecane Chain in Water, Methanol, and Their Mixtures. *The Journal of Physical Chemistry B* **2006**, *110*, 10519–10525, PMID: 16722762.

(20) Ferguson, A. L.; Debenedetti, P. G.; Panagiotopoulos, A. Z. Solubility and Molecular Conformations of n-Alkane Chains in Water. *The Journal of Physical Chemistry B* **2009**, *113*, 6405–6414.

(21) Hünenberger, P. Thermostat algorithms for molecular simulations. *Adv. Polym. Sci.* **2005**, *173*, 105.

(22) Vembanur, S.; Venkateshwaran, V.; Garde, S. Structure and Dynamics of Single Hy-

drophobic/Ionic Heteropolymers at the Vapor-Liquid Interface of Water. *Langmuir* **2014**, *30*, 4654–4661, PMID: 24689358.

(23) Deserno, M.; Holm, C. How to mesh up Ewald sums. 1. a theoretical and numerical comparison of various particle mesh routines. *J. Chem. Phys.* **1988**, *109*, 7678.

(24) Hockney, R.; Eastwood, J. *Computer Simulation using Particles*; IOP, London, 1988.

(25) Souaille, M.; Roux, B. Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations. *Computer Physics Communications* **2001**, *135*, 40 – 57.

(26) Chipot, C.; Wilson, M. A.; Pohorille, A. Interactions of Anesthetics with the Water-Hexane Interface. A Molecular Dynamics Study. *The Journal of Physical Chemistry B* **1997**, *101*, 782–791, PMID: 11542402.

(27) Plyasunov, A. V.; Shock, E. L. Thermodynamic functions of hydration of hydrocarbons at 298.15 K and 0.1 {MPa}. *Geochimica et Cosmochimica Acta* **2000**, *64*, 439 – 468.

(28) Chakrabarty, S.; Bagchi, B. Self-Organization of n-Alkane Chains in Water: Length Dependent Crossover from Helix and Toroid to Molten Globule. *The Journal of Physical Chemistry B* **2009**, *113*, 8446–8448, PMID: 19485315.

(29) Mountain, R.; Thirumalai, D. Hydration for a series of hydrocarbons. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 8436–8440.

(30) Mountain, R. D.; Thirumalai, D. Molecular Dynamics Simulations of End-to-End Contact Formation in Hydrocarbon Chains in Water and Aqueous Urea Solution. *Journal of the American Chemical Society* **2003**, *125*, 1950–1957, PMID: 12580622.