# Communication: On the locality of Hydrogen bond networks at hydrophobic interfaces

Bradley P. Lambeth, Jr.,[1] Christoph Junghans,[2] Kurt Kremer,[2] Cecilia Clementi,[1,a)] and Luigi Delle Site[2,b)]
[1]*Department of Chemistry and Department of Chemical and Biomolecular Engineering, Rice University, Houston, Texas 77005, USA*
[2]*Max Planck Institut für Polymerforschung, Ackermannweg 10, D-55128 Mainz, Germany*

(Received 15 July 2010; accepted 10 November 2010; published online 9 December 2010)

The formation of structured hydrogen bond networks in the solvation shells immediate to hydrophobic solutes is crucial for a large number of water mediated processes. A long lasting debate in this context regards the mutual influence of the hydrophobic solute into the bulk water and the role of the hydrogen bond network of the bulk in supporting the solvation structure around a hydrophobic molecule. In this context we present a molecular dynamics study of the solvation of various hydrophobic molecules where the effect of different regions around the solvent can be analyzed by employing an adaptive resolution method, which can systematically separate local and nonlocal factors in the structure of water around a hydrophobic molecule. A number of hydrophobic solutes of different sizes and two different model potential interactions between the water and the solute are investigated. © *2010 American Institute of Physics*. [doi:10.1063/1.3522773]

## I. INTRODUCTION

Liquid water is capable of forming highly complex hydrogen bond networks which directly affect the way in which biological molecules move and function.[1,2] In this context understanding the solvation of hydrophobic molecules is a key to understand crucial processes occurring in (bio-)molecular systems.[3–5] In general, the molecular structure of liquid water around a solute molecule results from a competition between the disruption of the local tetrahedral order of bulk water and the formation of a two dimensional surfacelike order at the solute interface. For very small hydrophobic solutes, such as methane, water molecules can encapsulate the guest molecule.[6] For increasingly larger solutes, the structure of water close to the solute eventually approaches the limiting case of an infinite hydrophobic surface.[7–9] An important question related to the structure of water around a hydrophobic solute concerns the locality of the hydrogen bond network, that is, whether or not the solvation structure is determined by the surrounding bulk. In order to address this computationally a tool is needed that can slowly switch on and off the hydrogen bonds in a well defined region around the solute, so that their relevance on the rest of the network can be determined in an unequivocal way. The switching process must occur without affecting the thermodynamic equilibrium of the whole system and, if the switching off occurs only in the bulk, thermodynamic equilibrium must be assured between the regions of different resolution. Such a tool is provided by the adaptive resolution scheme (AdResS),[10] that allows to interface regions with different molecular representations (e.g., atomistic and coarse-grained) while maintaining free exchange of particles and equilibrium between the regimes, that is, the two different regions maintain equal temperature, pressure, and density.[11] For the problem treated here AdResS is used to interface an atomistic model of water which explicitly forms hydrogen bonds with a coarse-grained (CG) spherical representation of water, which does not have any directional interactions. This allows to systematically determine the role of hydrogen bonds of the bulk onto the structure of water around hydrophobic solutes of different sizes. In particular we investigate the role of two common water-solute interactions, a Lennard–Jones (LJ) as well as a short ranged purely repulsive potential. For both interaction types the structure of the water layer close to the solute has to accommodate the geometrical constraints induced by the solvated molecule. To systematically test the locality of hydrophobic hydration, the $60n^2$ series of icosahedral fullerenes ($C_{60}$ to $C_{2160}$) were studied by AdResS simulations with varying thickness of the layer of atomistic molecular representation. Figure 1 illustrates the computational setup (see below for details). If adaptive resolution simulations with a minimal all-atom region can reproduce the results of fully all-atom simulations, then the effect of solute on the water structure is deemed local: it is not significantly supported/influenced by the bulk hydrogen bond network, as the bulk is modeled by CG water, unable to form directional hydrogen bonds.[12,13] In this case the role of the CG bulk water for the surface layer is minimal, that is, it is enough to ensure that the CG water acts as a thermodynamic bath with the same temperature, pressure, and density as the all-atom model. If the structural properties are perfectly local, there is no need for the CG bulk to provide any structural information. Regarding the specific character of the water-solute interaction, the explicit form of the potential has been long debated in the literature.[14–16] Some suggest the use of a Lennard–Jones type potential for the C-O nonbonded interaction[17] others a purely repulsive interaction.[18,19]

a)Electronic mail: cecilia@rice.edu.
b)Electronic mail: dellsite@mpip-mainz.mpg.de.

 

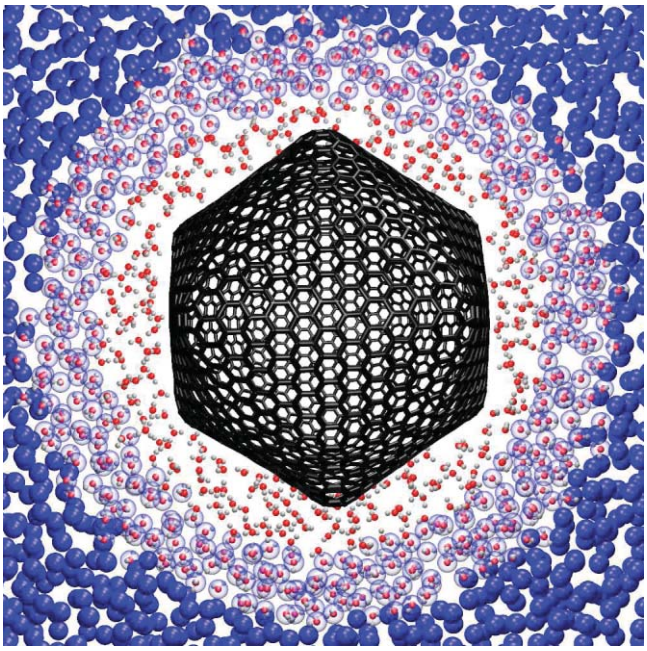

FIG. 1. Adaptive resolution simulation scheme for hydrophobic solutes, illustrated for the case of $C_{2160}$ icosahedral fullerene.

## II. SIMULATION SETUP

The application of AdResS to solvation processes requires the definition of an atomistic region around the solute, immersed in a "thermodynamic bath" of CG bulk water. In AdResS two molecules $\alpha$ and $\beta$ interact with a total force $F_{\alpha\beta}$, evaluated via a space dependent interpolation of the atomistic force $F_{\alpha\beta}^{\text{atom}}$ and CG force $F_{\alpha\beta}^{\text{cg}}$ as $\mathbf{F}_{\alpha\beta} = w(R_\alpha)w(R_\beta)\mathbf{F}_{\alpha\beta}^{\text{atom}} + [1 - w(R_\alpha)w(R_\beta)]\mathbf{F}_{\alpha\beta}^{\text{cg}}$, where $w(R)$ is a switching function; $w(R)$ is zero in the CG region, one in the spherical atomistic region (see Fig. 1), and continuous and monotonic in a transition region in between. The all-atom sphere is located at the center of the simulation volume, to which also the center of mass of the solute is constrained, ensuring that the solute is always expressed in full all-atom detail and surrounded by a layer of all-atom water.[20] A 1 nm transition region of hybrid molecules smoothly couples the layer of all-atom water around the solute with a bulk CG water region. As is evident from Fig. 1, the spherical AdResS setup creates an all-atom region of nonuniform thickness around the larger fullerenes. The distribution of the distances of the C atoms from the solute center of mass is spread over an interval $\Delta r$, with $\Delta r$ ranging from less than 0.05 nm for $C_{60}$ to about 0.5 nm for $C_{2160}$.[21,22] Nevertheless one can calculate a radial distribution function for the water molecules measured from the surface of the solute and define a first and second solvation shell (see Table I), even though this is somewhat smeared out for the larger solutes. The AdResS setup for the thinnest all-atom layer studied is illustrated in the inset of Fig. 2 for the largest solute. The green circle marks the average distance of the C atoms from the center of $C_{2160}$. The red circle marks the boundary of the all-atom region, which in this case also coincides with the minimum between the first and second peak of the solute-center water $g(r)$. The purple circle marks the border of the transition regime to the fully CG region. For both potentials we study two cases, namely an atomistic region up to the first and second minimum in the solute water $g(r)$, keeping in mind that the position of these minima is slightly different for the two interaction potentials. To actually analyze the solvation shell itself we calculate the number of water molecules $S(d)$, where $d$ is the distance from the solute surface and *not* from the center of the solvation shell (see definition below). All of the results presented were obtained for canonical ensemble simulations

TABLE I. The distance in nm of the first and second hydration shells measured from the center of mass of the solute, where we observe slight deviations between the two surface potentials.

| Interaction | Shell | $C_{60}$ | $C_{240}$ | $C_{540}$ | $C_{960}$ | $C_{1500}$ | $C_{2160}$ |
|---|---|---|---|---|---|---|---|
| Lennard–Jones | 1st | 0.85 | 1.20 | 1.53 | 1.90 | 2.25 | 2.55 |
| | 2nd | 1.05 | 1.40 | 1.73 | 2.10 | 2.45 | 2.75 |
| Purely repulsive | 1st | 0.95 | 1.30 | 1.63 | 2.00 | 2.35 | 2.65 |
| | 2nd | 1.25 | 1.60 | 1.93 | 2.30 | 2.65 | 2.95 |

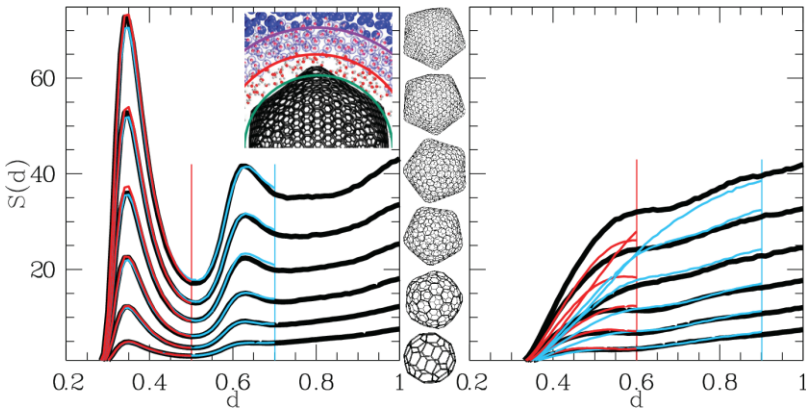


FIG. 2. Distribution $S(d)$ of the number of water O atoms at a distance $d$ (in nm, 0.01 nm bin size) from the surface of the solute, for the LJ (left), and the purely repulsive solutes (right). Each set of curves corresponds to the solutes in the order of their cartoon representation. Vertical lines indicate the average size of the all-atom region in each AdResS simulation. Colored curves correspond to different sizes of the all-atom region: up to the first hydration shell (red), or to the second hydration shell (light blue) as measured by $g(r)$. The thick black line represents the result for the reference all-atom simulation. The inset illustrates the simulation setup for the smallest all-atom layer for the case of the $C_{2160}$.

using a modified version of the GROMACS package.[23,24] Stochastic dynamics with a friction constant of $\Gamma = 5$ ps$^{-1}$ and a time step of 2.0 fs was used, and the electrostatics were modeled by a reaction-field method. The volume of the system was obtained from all-atom NPT simulations with $P_{\mathrm{ref}} = 1$ atm using the Berendsen barostat. A cutoff of 0.9 nm was used for the nonbonded interactions. After warm-up and equilibration of 1 ns, several trajectories of at least 2 ns were collected for each AdResS setup. To ensure that the results are not influenced by the choice of water model, simulations were tested for rigid and flexible versions of the TIP3P and SPC/E point charge models.[25,26] A single-site isotropic coarse-grained potential was obtained for each water model to match the corresponding O-O $g(r)$ using inverse Monte Carlo iterations with the VOTCA package.[27] The CG interaction potentials are centered on the O atom and are comparable in shape to other single-site water models.[28–30] This means that the CG model satisfies only a minimal structural requirement of a two body form; the local average tetrahedral structure is not reproduced.[28] A previously described optimized potentials for liquid simulations model for buckminsterfullerene was adapted for the $60n^2$ fullerenes.[19] The functional form of the nonbonded potentials (for C-O) was $U = 4\epsilon[(\sigma/r)^{12} - a(\sigma/r)^6]$ with $a = 1$ for Lennard–Jones and $a = 0$ for purely repulsive interactions; the corresponding parameters were determined by Lorentz–Berthelot mixing rules. It is important to note that the typical energy of a hydrogen bond (on average 20.41 kJ/mol for SPC/E), is roughly 50 times larger than the optimal C-O Lennard–Jones interaction.[31]

To quantify the change in water structure around a solute, the average number of water molecules $S(d)$ and the average tetrahedral order parameter $q(d)$ are measured as a function of $d$ from the closest C atom of the solute for a bin size of 0.01 nm. For one single water molecule $i$, the order parameter $q_i$ is defined as $q_i = 1 - \frac{3}{8}\sum_{j=1}^{3}\sum_{k=j+1}^{4}(\cos\,\psi_{ijk} + \frac{1}{3})^2$, where $\psi_{ijk}$ is the angle formed by the oxygen atoms of two neighbor water molecules $j$ and $k$ with the oxygen atom of molecule $i$, and the sum runs over the four nearest neighbor molecules of molecule $i$.[32] The function $q(d)$ is the average $q_i$ over all water molecules at a given distance $d$ from the solute surface. For perfectly tetrahedral systems $q = 1$, while for disordered systems $q = 0$ on average.

## III. RESULTS

Figures 2 and 3 show the relative ability of each adaptive resolution simulation to reproduce the local density and structure of water around the solute. Comparison of the right and left panels of Fig. 2 immediately reveals that while limiting the size of the all-atom region to the first hydration shell is sufficient to reproduce the density of water around the LJ solutes (left panel), the situation is very different for the purely repulsive solutes, where the radial distribution of the water is greatly disturbed by the coarse-graining of the bulk (right panel). We notice an improvement in the results for much larger all-atom regions (not shown), however, all-atom region sizes comparable to the first and second hydration shells are insufficient to accurately reproduce the results obtained in fully all-atom simulation. Thus for purely repulsive hydrophobes, structural changes in the bulk directly affect the local water density in the first hydration shell.

This difference between LJ and purely repulsive solutes is also immediately evident from the distribution of the tetrahedral order parameter $q(d)$ around the solutes, shown in Fig. 3. The repulsive solutes (right panel) exhibit almost negligible tetrahedral order close to the solute surface. In particular, right next to the repulsive solute surface the packing appears to be increasingly more random with q→0 for increasing solute size. Surprisingly, this appears rather independent of the thickness of the all-atom layer, indicating the locality of the tetrahedral order, while the water density in the first

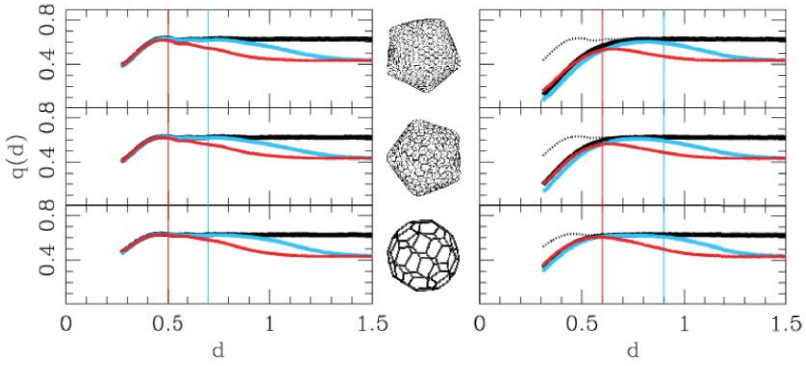


FIG. 3. The average tetrahedral order parameter $q(d)$ as a function of the distance $d$ (in nm) from the surface of the solute, for the LJ (left) and purely repulsive solutes (right). Vertical lines and colors are the same as in Fig. 2. Results are reported for three representative solutes (see also the corresponding cartoon): $C_{60}$ (bottom plots in both right and left panels), $C_{960}$ (middle plots), and $C_{2160}$ (top plots). On the plots for the repulsive solutes the curves for the fully all-atom simulations of the corresponding LJ solutes are also reported as a dotted black line, to provide a direct comparison of the different cases.

TABLE II. The fluctuations, expressed as $(\langle N^2\rangle - \langle N\rangle^2)/\langle N\rangle$, for each solute in the first layer as defined for Fig. 2 and all-atom region size (d1 for red, d2 for light blue).

| Interaction | Simulation | $C_{60}$ | $C_{240}$ | $C_{540}$ | $C_{960}$ | $C_{1500}$ | $C_{2160}$ |
|---|---|---|---|---|---|---|---|
| Lennard–Jones | AdResS$_{d1}$ | 0.130 | 0.116 | 0.116 | 0.119 | 0.116 | 0.115 |
| | AdResS$_{d2}$ | 0.152 | 0.140 | 0.133 | 0.133 | 0.130 | 0.124 |
| | All-atom | 0.138 | 0.131 | 0.129 | 0.130 | 0.146 | 0.134 |
| Purely repulsive | AdResS$_{d1}$ | 0.379 | 0.643 | 0.823 | 1.00 | 1.22 | 1.30 |
| | AdResS$_{d2}$ | 0.585 | 0.961 | 1.228 | 1.45 | 1.62 | 1.49 |
| | All-atom | 0.463 | 0.702 | 0.812 | 0.911 | 1.12 | 1.24 |

layers exhibits the characteristics of a property which needs the support of the bulk. On the contrary, the LJ solutes (left panel) present a smaller, albeit significant, decrease in the parameter $q$ (from the bulk value of $q \simeq 0.6$ to $q \simeq 0.3$–$0.4$) closer to the solute surface. These results suggest that LJ solutes induce a locally ordered yet very flexible hydrogen bond network, completely consistent with recent vibrational sum frequency spectroscopy results for water.[33] In addition, the height of the first peak of the radial distribution function (with respect to the surface) $g(r)$ (not shown) for the LJ solutes does not decrease for increasing solute size, consistently with available literature.[34] On the contrary, for the repulsive solutes a complete "nonwetting" is observed, as the $g(r)$ function for two largest repulsive solutes has essentially no first peak.[35] The difference in nonwetting behavior observed with the two interaction types is fully consistent with previous studies for attractive and repulsive hydrophobic sheets.[9,36]

These results show how adaptive simulations are capable of shedding some light on interface problems. Adaptive simulations that approximate the bulk with a CG model cannot match the local density of water around a repulsive solute, while they do for a LJ solute. In contrast, the tetrahedral order in both cases seems to be only weakly affected. As a consequence, contrary to the radial distribution function, the tetrahedral order parameter does not allow for a clear distinction between local and nonlocal effects for the hydrophobic solute-water interactions as defined in this work. Although these static quantities are among the most relevant for the structure at the hydrophobic interface, they are not sufficient to determine whether AdResS has adequately captured the essential physics of the solvation process. A complementary dynamic quantity which can support the validity of our conclusions is the particle number fluctuation in the first hydration shell,[9,14,34] see Table II. For each LJ solute the adaptive resolution simulations match the all-atom simulation, demonstrating that free exchange of particles occurs between the all-atom and CG regions. A small discrepancy can be observed for the repulsive solutes, reflecting the lack of locality. In summary, the level of locality in the hydrogen bond network at hydrophobic interfaces is primarily controlled by the nature of the interaction between the hydrophobic solute and water. For purely repulsive solutes, perturbations in the bulk (such as the CG approximation) affect the density but not the tetrahedral order of water at the solute interface. For a weak LJ solute, bulk perturbations tested here do not affect the first water layers. The size of the solute does not appear to affect the locality of the solvation significantly, at least for the range of solutes considered here.

## ACKNOWLEDGMENTS

We thank Matej Praprotnik for the technical support. C.C. acknowledge the financial support of the Welch Foundation grants and NSF. The computational resources of the Rice Computational Research Clusters were used. C.J., K.K., V.S.F. and L.D.S. acknowledge the financial support of the MMM initiative of the Max-Planck Society.

[1] D. Huang and D. Chandler, Proc. Natl. Acad. Sci. U.S.A. **97**, 8324 (2000).
[2] D. Chandler, Nature (London) **437**, 640 (2005).
[3] G. Papoian, J. Ulander, M. Eastwood, Z. Luthey-Schulten, and P. Wolynes, Proc. Natl. Acad. Sci. U.S.A. **101**, 3352 (2004).
[4] P. Ball, Chem. Rev. **108**, 74 (2008).
[5] C. Clementi, Curr. Opin. Struct. Biol. **18**, 10 (2008).
[6] T. Lazaridis and M. E. Paulaitis, J. Phys. Chem. **96**, 3847 (1992).
[7] C. Lee, J. A. McCammon, and P. J. Rossky, J. Chem. Phys. **80**, 4448 (1984).
[8] K. Lum, D. Chandler, and J. Weeks, J. Phys. Chem. B **103**, 4570 (1999).
[9] A. J. Patel, P. Varilly, and D. Chandler, J. Phys. Chem. B **114**, 1954 (2010).
[10] M. Praprotnik, L. Delle Site, and K. Kremer, J. Chem. Phys. **123**, 224106 (2005).
[11] S. Poblete, M. Praprotnik, K. Kremer, and L. Delle Site, J. Chem. Phys. **132**, 114101 (2010).
[12] S. Matysiak, M. Praprotnik, L. Delle Site, K. Kremer, and C. Clementi, J. Phys.: Condens. Matter **19**, 292201 (2007).
[13] S. Matysiak, C. Clementi, M. Praprotnik, K. Kremer, and L. Delle Site, J. Chem. Phys. **128**, 024503 (2008).
[14] S. Sarupria and S. Garde, Phys. Rev. Lett. **103**, 037803 (2009).
[15] N. Choudhury and B. Pettitt, J. Am. Chem. Soc. **129**, 4847 (2007).
[16] H. Ashbaugh and M. Paulaitis, J. Am. Chem. Soc. **123**, 10721 (2001).
[17] J. Ma, D. Alfe, A. Michaelides, and E. Wang, J. Chem. Phys. **130**, 154303 (2009).
[18] D. Chandler, J. Weeks, and H. Andersen, Science **220**, 787 (1983).
[19] D. Weiss, T. Raschke, and M. Levitt, J. Phys. Chem. B **112**, 2981 (2008).
[20] M. Praprotnik, L. Delle Site, and K. Kremer, J. Chem. Phys. **126**, 134902 (2007).
[21] C. Xu and G. Scuseria, Chem. Phys. Lett. **262**, 219 (1996).
[22] R. Zope, T. Baruah, M. Pederson, and B. Dunlap, Phys. Rev. B **77**, 115452 (2008).
[23] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, J. Chem. Theory Comput. **4**, 435 (2008).
[24] B. Lambeth, S. Fritsch, and C. Junghans, *AdressMacs*, http://repo.or.cz/w/gromacs/adressmacs.git (2010).
[25] W. Jorgensen, J. Chandrasekhar, J. Madura, R. Impey, and M. Klein, J. Chem. Phys. **79**, 926 (1983).
[26] H. Berendsen, J. Grigera, and T. Straatsma, J. Phys. Chem. **91**, 6269 (1987).
[27] V. Rühle, C. Junghans, A. Lukyanov, K. Kremer, and D. Andrienko, J. Chem. Theory Comput. **5**, 3211 (2009).
[28] H. Wang, C. Junghans, and K. Kremer, Eur. Phys. J. E **28**, 221 (2009).
[29] R. M. Lynden-Bell and T. Head-Gordon, Mol. Phys. **104**, 3593-3605 (2006).
[30] S. Izvekov and G. A. Voth, J. Chem. Phys. **123**, 134105 (2005).
[31] J. Zielkiewicz, J. Chem. Phys. **123**, 104501 (2005).
[32] J. Errington and P. Debenedetti, Nature (London) **409**, 318 (2001).
[33] L. Scatena, M. Brown, and G. Richmond, Science **292**, 908 (2001).
[34] J. Mittal and G. Hummer, Proc. Natl. Acad. Sci. U.S.A. **105**, 20130 (2008).
[35] F. H. Stillinger, J. Solution Chem. **2**, 141 (1973).
[36] N. Choudhury and B. Pettitt, Mol. Simul. **31**, 457 (2005).