Supplementary Information for Hydrophobicity of proteins and nanostructured solutes is governed by their topographical and chemical context

Erte Xi,¹ Vasudevan Venkateshwaran,² Lijuan Li,²

Nicholas Rego,¹ Amish J. Patel,¹ and Shekhar Garde²

¹Department of Chemical & Biomolecular Engineering, University of Pennsylvania, Philadelphia, PA 19104 ²Howard P. Isermann Department of Chemical & Biological Engineering and Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, NY 12180

In the main text we have presented results from molecular simulations of three different systems – hemi-cylindrical nanotubes (hCNTs), patterned SAM surfaces, and mutants of the protein hydrophobin-II. MD simulations were performed using GROMACS [1] with 3D periodic boundary conditions. The SPC/E water model [2] was used in all systems. The Leapfrog algorithm with a time-step of 2 fs was used to integrate the equations of motion. Bonds involving hydrogens were constrained using the LINCS algorithm [3]. Electrostatic interactions were calculated using the particle mesh Ewald algorithm [4]. Parameters for cross interactions were calculated using the Lorentz-Berthelot mixing rules. Configurations were stored every 1 ps for analysis. Here we provide additional information and results for each of these systems. We begin with the decription of hCNT systems.

I. NANOTUBE SIMULATIONS

A. Water droplet contact angle dependence on carbon-water attractions

To tune the hydrophobicity of the reference graphene surface, we varied graphene-water interactions, u(r), using the Weeks-Chandler-Andersen scheme [5] as follows:

$$u(r) = \begin{cases} 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 + \frac{1-\lambda}{4} \right] : r < 2^{1/6}\sigma \\ 4\epsilon\lambda \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 \right] : r \ge 2^{1/6}\sigma \end{cases}$$
(1)

where σ and ϵ are graphene-carbon-water-oxygen Lennard-Jones parameters (see Materials and Methods) and λ is a scaling factor.

We characterized the wettability (i.e., hydrophobicity) of the graphene sheet as a function of the scaling factor λ by performing 10 ns long MD simulations of water droplets containing 2500 water molecules placed on a 16×16 nm² sheet and measuring the contact angles (see [6] and [7]), as shown in Figure S1. As λ decreases, the surface becomes less attractive to water, and the contact angle θ increases.



FIG. S1. Contact angle of a water droplet on a flat graphene surface as a function of the scaling parameter, λ , used to vary carbon-water oxygen potential. The dashed line serves as a guide to the eye.

B. Details of the hCNT MD simulations

The hCNT was constrained with its axis aligned with the z-axis of the simulation box. Sizes of hCNTs, simulation boxes, and additional details are given in SI. Carbon-water interactions are Lennard-Jones type with $\sigma_{\rm CO} = 0.3275$ nm and $\epsilon_{\rm CO} = 0.478$ kJ/mol [8, 9]), which were scaled to study the effects of attractions as described in SI. Table I lists the radius, r, curvature, κ , and the average box dimesions in the x and y directions for each of our hCNT simulation systems. The length of the nanotube, L_{hCNT} , and the simulation box length in the z direction depended on the dimensions of the observation volume needed to sample the requisite quantities. These are presented in detail in the next section. The hCNTs was solvated in sufficient number of water molecules such that the bulk water density was 33.3 molecules per nm³ at the 1 atm pressure state point. Thus, the number of water molecules ranged from 3324 to 9112 depending on the length of the tube.

hCNT	$r_{\rm hCNT}$ (nm) ^a .	$ \kappa \ (\mathrm{nm}^{-1})^{\mathrm{\ b}}$	box cross-section (nm^2)	
(12, 12)	0.814	1.229	4.7×4.7	
(16, 16)	1.085	0.922	5.3 imes 5.3	
(24, 24)	1.627	0.614	6.4 imes 6.4	

TABLE I. Systems sizes for simulations of hCNT hydration

^a The radius of the armchair carbon nanotube is given by $r_{\rm hCNT} = \frac{a}{2\pi}\sqrt{n^2 + nm + m^2}$, where the lattice constant a = 0.246 nm

^b $|\kappa| = 1/r$. Concave and convex curvatures are given negative and positive signs, respectively.

C. Observation volume dimensions

The observation volume, v, for hCNTs includes a cylindrical shell volume that hugs the concave or convex surface of hCNT bounded in radial, axial, and angular directions. The bottom panel of Figure 1A shows the observation volume for the concave surface as an example. The first minimum in the radial density profile of water with respect to the surface of the hCNT is taken as one of the radial bounds for v. Angle of 120 degrees was consistently used in the angular direction.



FIG. S2. Radial density profile of water with respect to the hCNT. The hydration shell on the concave and convex sides are shaded blue and red, respectively, bounded on the water side by the distance corresponding to the first minima in the density. (Inset) For $\lambda = 0$ (i.e., for purely repulsive WCA hCNTs) the density profile is sigmoidal in nature. In such cases the first hydration shell is defined as the distance at which the density reaches 95% of the bulk water density.

Figure S2 shows the radial water density profile near the concave and convex sides with various curvatures for $\lambda = 0.5$. That the profiles for different curvature are essentially the same highlights



FIG. S3. Schematic illustrating the observation volume, v, on the concave side of hCNT

(112) $(112$									
hCNT	Concave			Convex					
	$r_i \ (nm)$	$r_o (\rm nm)$	$L \ (nm)^a$	$r_i (\mathrm{nm})$	$r_o (\rm nm)$	L (nm)			
(12, 12)	0.264	0.814	5.35	0.814	1.364	1.62			
(16, 16)	0.535	1.085	3.14	1.085	1.635	1.36			
(24, 24)	1.077	1.627	1.71	1.627	2.177	1.00			

TABLE II. Observation volume dimensions for compressibility and INDUS calculations for $\lambda = 0.5$

^a For all simulations, L_{hCNT} was chosen to be L + 0.5 nm. Correspondingly the simulation box dimension along the axis of the hCNT was $L_{hCNT} + 3.0$ nm

the fact that water density profiles does not provide a signature of the hydrophobicity of the curved surface. In contrast, as shown in the main text, density fluctuations of water near these surfaces are significantly different, providing a robust signature of hydrophobicity.

The first hydration shells on the concave and convex sides are highlighted using blue and red shading, respectively. Table II lists the values of the system sizes and the relevant dimensions that define the observation volume for both the concave and convex sides.

For calculation of hydration water compressibility, the length of the shell along the axial direction, L, was set to 1.2 nm for all hCNT systems. For INDUS calculations, the length along the axial direction was selected to ensure the same average number of waters in v for all hCNTs to enable comparisons on the same footing as described in the main text. The values of the length of observation volume, L, for INDUS calculations are listed in Table II. To avoid edge effects, the length of the nanotube, L_{hCNT} , was chosen to be L + 0.5 nm. Correspondingly, the simulation box dimensions in the z direction was increased to $L_{hCNT} + 3.0$ nm to accomadate the larger nanotube.



FIG. S4. The normalized compressibility of water molecules in the first hydration shell, $\chi/\chi_{\text{graphene}}$ as a function of hCNT curvature, κ , for $\lambda = 0.0, 0.5$, and 1.0. The dashed line serves as a guide to the eye.

D. hCNT compressibility calculations

The hCNT compressibility calculations were performed in a isothermal-isobaric, (N, P, T), ensemble with temperature of 300 K maintained using a stochastic velocity rescaling thermostat [10] and pressures ranging from 1 atm to 4000 atm maintained using a Berendsen barostat [11].

E. Effect of attractions on the hydration shell compressibility

The main text discussed results for hCNTS with $\lambda = 0.5$. This scaling was chosen because a flat graphene surface with $\lambda = 0.5$ shows wettability similar to a -CH₃ SAM surface. Figure S4 shows the normalized compressibility, $\chi/\chi_{\text{graphene}}$, as a function of κ for two other attractions as well. As shown previously, the hydration shell compressibility near a flat surface decreases with increasing surface-water attractions [12–15]. Such a trend is also observed for the curved surfaces of hCNTs. For a given curvature, $|\kappa|$, on the concave side, the decrease in χ with increasing attractions is more drastic compared to that near a flat graphene sheet, whereas on the convex side, the decrease in χ is smaller compared to a flat graphene sheet, resulting in the trends observed in Figure S4. These results are consistent with previous simulation studies on spherical solutes [16, 17]



FIG. S5. The normalized compressibility of water molecules in the first hydration shell, $\chi/\chi_{\text{graphene}}$ as a function of liquid interface curvature, κ' , for $\lambda = 0.5$. The dashed line serves as a guide to the eye.

F. Impact of alternate definition of curvature

In the main text and in Figure S4, we defined curvature, $\kappa = 1/r$, using the radius of curvature of the surface that contains carbon centers. Thus, for a given nanotube, concave and convex curvatures are identical in magnitude, but opposite in sign. One might argue, however, that surface curvature should be defined using the position of liquid water interface. We used the locations of the first peaks of radial density profile (see Fig. S2) to obtain the shift in the radius of curvature, and redefined $\kappa' = 1/(r + \delta)$ where $\delta = 0.37$ nm. (Note that for concave surfaces, r, is negative). Figure S5 shows the variation of water compressibility (for $\lambda = 0.5$ system) using this alternate definition of κ' , which now spans a larger range of concave curvatures. While the results are quantitatively different from those shown in the main text, the key result that compressibility of hydration water is higher near concave surfaces compared to convex ones still holds true.

G. hCNT INDUS calculations

INDUS calculations to obtain $P_v(N)$ were performed in the canonical, (N, V, T), ensemble with temperature of 300 K maintained using a stochastic velocity rescaling thermostat [10]. Each umbrella window was equilibrated for 1 ns, followed by a production run of at least 3 ns. The unbiased distribution was calculated using the WHAM method [18]. The simulation system also included a repulsive wall and a vapor buffer layer to avoid suppression of density fluctuations. The details and advantages of such a construction have been discussed previously by Patel et. al. [13]. The coarse-grained particle number was calculated by setting $\xi = 0.01$ nm and $r_c=0.02$ nm. A harmonic potential was applied to the coarse-grained particle number with a spring constant of k=1 kJ/mol [19].

II. PATTERNED SAM SURFACES

The construction of SAMs and the force-field parameters have been described in detail previously [12]. They comprise 10-carbon alkane chains terminated with hydrophobic ($-CH_3$) and hydrophilic (-OH) head-groups. SAMs used here have a cross-sectional area of 7 nm × 7 nm (with 224 total chains). Chemical patterns were created by replacing the $-CH_3$ head-groups with -OH. A ~4 nm thick slab of water containing 6464 water molecules was used to solvate the SAM surface.

A. INDUS calculations in SAM systems

The INDUS calculations for patterned SAM systems were performed in the canonical, (N, V, T), ensemble with temperature of 300 K maintained using a Berendsen thermostat [20]. Each umbrella window was equilibrated for 1 ns, followed by a production run of at least 3 ns. The unbiased distribution was calculated using the WHAM method [18]. As described in the previous section, the simulations included a repulsive wall and a vapor buffer layer. The coarse-grained particle number was calculated by setting $\xi = 0.01$ nm, and $r_c=0.02$ nm. A harmonic potential on the coarse-grained particle number with a spring constant of k = 0.25 kJ/mol was used.

B. Free energy profiles of patterned SAM surfaces



FIG. S6. Free energy as a function of number of water molecules for SAM surface (A) patterned with 4 -OH head groups. (B) patterned with 7 -OH head groups. The right hand column in each row shows a close-up of the region near N = 0.

C. Interface configurations for patterned SAM surfaces

In Figure 2 of the main text we discussed the behavior of the SAM–water interface, z(x, y), on patterned SAM surfaces. Figure 2 showed a subset of our results that highlighted the key insights gained from the data. Figures S7 (pattern with 4 -OH head groups) and S8 (pattern with 7 -OH head groups) show the behavior of the interface, z(x, y), for a range of ρ_v and s values.



FIG. S7. Interfaces for patterned SAM surface with 4 -OH head groups.



FIG. S8. Interfaces for patterned SAM surface with 7 -OH head groups.

III. HYDROPHOBIN-II SIMULATIONS

Hydrophobin-II was represented using the AMBER-94 force field [21] and solvated in a periodic box of dimensions $6.3 \times 6.3 \times 6.3 \text{ mm}^3$ with ~8000 water molecules. Protein heavy atoms were position constrained harmonically with a spring constant of 1000 kJ/mol/nm.

A. Comparing hydrophobin-II mutant structures

To ensure that swapping amino acids does not perturb the protein structure significantly, we performed 3 ns long MD simulations of the wild-type hydrophobin-II and its swap mutant in water starting from the crystal structure. Simulations were performed in the (N, P, T) ensemble without any position constraints on the protein heavy atoms. Figure S9 shows that both the wild-type and



FIG. S9. RMSD for Free energies as a function of number of water molecules for protein hydrophobin II.

swap mutant maintain a stable folded structure with RMSD of ≈ 0.2 nm from the crystal structure. Additionally, the RMSD between the a representative wild-type structure (extracted after the 3 ns simulation) and swap mutant structures sampled is also ≈ 0.2 nm confirming that the perturbation introduced by the swapping of amino acids is minimal and within thermal fluctuations that are observed during a MD simulation.

B. Protein hydrophobicity at atomic level



FIG. S10. Protein hydrophobin II wild type and mutant showing non-polar atoms in red and polar atoms in blue. The orientations of the proteins are the same as in the main text.

Atomic-level details within each amino acid can be important to better understand protein interactions. As shown in the recent publication by Kapcha and Rossky [22], an atomic-level description of protein hydrophobicity can provide rationales to protein functions that residuebased scales fail to explain. In Figure S10 we show the large hydrophobic patch of the protein hydrophobin II for both wild type and mutant, but with the atomic-level hydrophobicity scale. It is clear that this large hydrophobic patch of interest is actually quite heterogeneous.

C. Defining the observation volume near protein patches

The observation volume near the hydrophobic patch on the surface of hydrophobin-II includes a contiguous region encompassing 10 amino acid residues (LEU-7, PHE-8, LEU-19, LEU-21, ILE-22, VAL-54, ALA-61, LEU-62, LEU-63, and ASP-59) shown in Figures 3A and 3B. This volume is created by defining spherical sub-volumes of diameter 0.6 nm centered on each heavy atom of the 10 amino acids that form the hydrophobic patch. This results in 79 spherical sub-volumes, the union of which defines the observation volume for INDUS calculations.

D. INDUS calculations for hydrophobin-II systems

Hydrophibin-II simulations were performed in the isothermal-isobaric, (N, P, T), ensemble with temperature of 300 K maintained using a stochastic velocity rescaling thermostat [10] and pressure of 1 bar maintained using a Parrinello-Rahman barostat [23]. Each umbrella window was equilibrated for 1 ns, followed by a production run of at least 3 ns. The unbiased distribution was calculated using the WHAM method [18]. The INDUS algorithm was modified to include contributions of the forces from the umbrella potential in the calculation of the system pressure. The coarse-grained particle number was calculated by setting $\xi = 0.01$ nm and $r_c=0.02$ nm. A harmonic potential on the coarse-grained particle number with a spring constant of k = 0.24 kJ/mol was used.

E. The effect of the initial hydration state on the free energies

The simulations of hydrophobin-II discussed in the main text were prepared in a "wet state", i.e., the observation volume was filled with water and contained $\approx \langle N \rangle_v$ water molecules to begin with. In order to eliminate any artifacts of this initial condition, we also performed simulations of our protein systems where the initial coordinates were prepared in a "dry state", *i.e.*, the observation volume contained no water molecules in it. Figure S11 compares the free energies computed from both the dry and wet initial states. We note that for both the wild type and the swap mutant, the free energy differences calculated from the dry state is within 3% of the value calculated from



FIG. S11. Free energies as a function of number of water molecules for protein hydrophobin II.

the wet state, confirming the accuracy of the reported free energies, and allowing us to distinguish between the hydrophobicities of wild type and mutant systems.

IV. A NOTE ON THE COMPUTATION OF INSTANTANEOUS INTERFACES

In Figure 2 and Figure 3 of the main text we presented results that visualized interfaces near a patterned SAM and protein surface. We calculate these density isosurfaces following the method developed by Willard and Chandler [24]. Briefly, the method involves defining a coarse-grained density field :

$$\tilde{\rho}_{\alpha}(x, y, z; \overline{\mathbf{R}}) = \sum_{i=1}^{N_{\alpha}} \phi(x - x_i) \phi(y - y_i) \phi(z - z_i)$$
(2)

where $\tilde{\rho}_{\alpha}(x, y, z; \mathbf{\overline{R}})$ is the coarse-grained density of type α atoms at coordinates (x, y, z), as a function of all the atom positions $\mathbf{\overline{R}}$. N_{α} is the number of atoms with type α . $\phi(x)$ is the coarsegraining function which was chosen to be a Gaussian with a width of 0.24 nm, which is truncated at 0.7 nm and shifted down. Depending on the atomic coordinates included in the computation of the coarse-grained density field one could compute and visualize iso-surfaces, $\tilde{\rho} = c$, that represent various interfaces within the system. See for example [24] (an air–water interface) and [25] (cavities in a confined liquid phase).

Patterned SAM systems: Our definition of the interface, z(x, y), as presented in the discussion of the patterned SAM systems was computed on a coarse-grained density filed that included oxygen atoms from all water molecules, and normalized by the bulk water density,

$$\tilde{\rho}_{\text{total}}(x, y, z; \overline{\mathbf{R}}) = \frac{\tilde{\rho}_{\text{water}}(x, y, z; \overline{\mathbf{R}})}{\rho_{\text{water}}^{\text{bulk}}}$$
(3)

where $\tilde{\rho}_{\text{total}}$ will be approximately unity everywhere within the bulk water phase and will approach zero as we cross into the SAM surface region. We compute the ensemble average of the coarse-grained density field, $\langle \tilde{\rho}_{\text{total}} \rangle$, over the entire simulation trajectory, and define $\langle \tilde{\rho}_{\text{total}} \rangle = 0.5$ iso-surface as the SAM-water interface, z(x, y).

Hydrophobin-II mutants: For our hydrophobin-II mutants, our goal is to compute and visualize the dewetted regions (cavities) on the protein surface. For this purpose we include in our computation the coarse-grained density arising from the protein heavy atoms to define a overall normalized coarse-grained density,

$$\tilde{\rho}_{\text{total}}(x, y, z; \overline{\mathbf{R}}) = \frac{\tilde{\rho}_{\text{water}}(x, y, z; \overline{\mathbf{R}})}{\rho_{\text{water}}^{\text{bulk}}} + \frac{\tilde{\rho}_{\text{protein}}(x, y, z; \overline{\mathbf{R}})}{\rho_{\text{protein}}^{\text{max}}}$$
(4)

where we normalized the protein coarse-grained density by the maximum protein density occurs at the center of the protein. Under this definition, $\tilde{\rho}_{\text{total}}$ will be approximately equal to unity everywhere in the system, except at dewetted regions which form a cavity. We define the $\tilde{\rho}_{\text{total}} = 0.5$ iso-surface here to represent those cavities.

- Berk Hess, Carsten Kutzner, David van der Spoel, and Erik Lindahl. Gromacs 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. J. Chem. Theory Comput., 4(3):435–447, 2008.
- [2] Herman J.C. Berendsen, J. Raul Grigera, and T. P. Straatsma. The missing term in effective pair potentials. J. Phys. Chem., 91(24):6269–6271, 1987.
- [3] Berk Hess, Henk Bekker, Herman J. C. Berendsen, and Johannes G. E. M. Fraaije. Lincs: A linear constraint solver for molecular simulations. J. Comput. Chem., 18(12):1463–1472, 1997.
- [4] Tom Darden, Darrin York, and Lee Pedersen. Particle mesh ewald an n.log(n) method for ewald sums in large systems. J. Chem. Phys., 98(12):10089–10092, 1993.
- [5] John D. Weeks, David Chandler, and Hans C. Andersen. Role of repulsive forces in determining equilibrium structure of simple liquids. J. Chem. Phys., 54(12):5237+, 1971.
- [6] Joseph Hautman and Michael L Klein. Microscopic wetting phenomena. Phys. Rev. Lett., 67(13):1763, 1991.
- [7] Michel J De Ruijter, TD Blake, and Joël De Coninck. Dynamic wetting studied by molecular modeling simulations of droplet spreading. *Langmuir*, 15(22):7836–7847, 1999.

- [8] Gerhard Hummer, Jayendran C. Rasaiah, and Jerzy P. Noworyta. Water conduction through the hydrophobic channel of a carbon nanotube. *Nature*, 414(6860):188–190, 2001.
- [9] Amrit Kalra, Shekhar Garde, and Gerhard Hummer. Osmotic water transport through carbon nanotube membranes. Proc. Natl. Acad. Sci. USA, 100(18):10175–10180, 2003.
- [10] Giovanni Bussi, Davide Donadio, and Michele Parrinello. Canonical sampling through velocity rescaling. J. Chem. Phys., 126(1), 2007.
- [11] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. Molecular dynamics with coupling to an external bath. J. Chem. Phys., 81(8):3684–3690, 1984.
- [12] Rahul Godawat, Sumanth N. Jamadagni, and Shekhar Garde. Characterizing hydrophobicity of interfaces by using cavity formation, solute binding, and water correlations. *Proc. Natl. Acad. Sci. USA*, 106(36):15119–15124, 2009.
- [13] Amish J. Patel, Patrick Varilly, and David Chandler. Fluctuations of water near extended hydrophobic and hydrophilic surfaces. J. Phys. Chem. B, 114(4):1632–1637, 2010.
- [14] R. Godawat, S. N. Jamadagni, V. Venkateshwaran, and S. Garde. Connecting water correlations, fluctuations, and wetting phenomena at hydrophobic and hydrophilic surfaces. ArXiv e-prints, September 2014.
- [15] Nicolas Giovambattista, Pablo G Debenedetti, and Peter J Rossky. Enhanced surface hydrophobicity by coupling of surface polarity and topography. Proc. Natl. Acad. Sci. USA, 106(36):15181–15185, 2009.
- [16] Sapna Sarupria and Shekhar Garde. Quantifying water density fluctuations and compressibility of hydration shells of hydrophobic solutes and proteins. *Phys. Rev. Lett.*, 103(3), 2009.
- [17] Jeetain Mittal and Gerhard Hummer. Static and dynamic correlations in water at hydrophobic interfaces. Proc. Natl. Acad. Sci. USA, 105:20130–20135, 2008.
- [18] Shankar Kumar, John M. Rosenberg, Djamal Bouzida, Robert H. Swendsen, and Peter A. Kollman. The weighted histogram analysis method for free-energy calculations on biomolecules. i. the method. J. Comput. Chem., 13(8):1011–1021, 1992.
- [19] Amish J Patel, Patrick Varilly, David Chandler, and Shekhar Garde. Quantifying density fluctuations in volumes of all shapes and sizes using indirect umbrella sampling. J. Stat. Phys., 145(2):265–275, 2011.
- [20] Herman JC Berendsen, JPM van Postma, Wilfred F van Gunsteren, ARHJ DiNola, and JR Haak. Molecular dynamics with coupling to an external bath. J. Chem. Phys., 81(8):3684–3690, 1984.
- [21] Wendy D. Cornell, Piotr Cieplak, Christopher I. Bayly, Ian R. Gould, Kenneth M. Merz, David M. Ferguson, David C. Spellmeyer, Thomas Fox, James W. Caldwell, and Peter A. Kollman. A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules. J. Am. Chem. Soc., 117(19):5179–5197, 1995.
- [22] Lauren H Kapcha and Peter J Rossky. A simple atomic-level hydrophobicity scale reveals protein interfacial structure. J. Mol. Biol., 426(2):484–498, 2014.

- [23] Michele Parrinello and Aneesur Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics*, 52(12):7182–7190, 1981.
- [24] A. P. Willard and D. Chandler. Instantaneous liquid interfaces. J. Phys. Chem. B, 114(5):1954–1958, 2010.
- [25] Richard C. Remsing, Erte Xi, Srivathsan Ranganathan, Sumit Sharma, Pablo G. Debenedetti, Shekhar Garde, and Amish J. Patel. Pathways to dewetting in hydrophobic confinement. Proc. Natl. Acad. Sci. USA, 112:8181–8186, 2015.