Supporting Information - Computational discovery of chemically patterned surfaces that effect unique hydration water dynamics

Jacob I. Monroe[a], M. Scott Shell[a*]

[a] Department of Chemical Engineering, University of California – Santa Barbara
* Corresponding author

1) Additivity of dynamics in simple Markov models

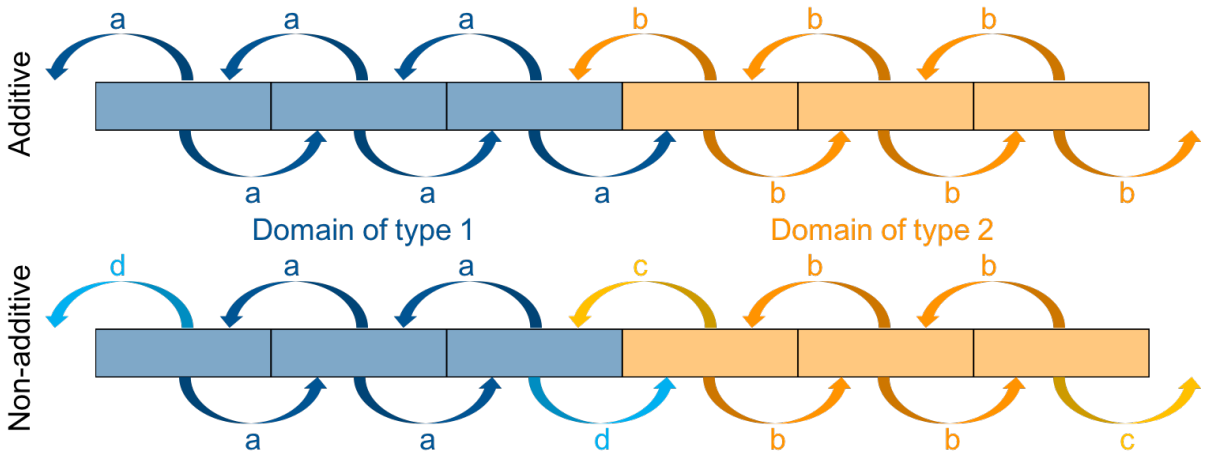

Figure S1: Schematics of 1D Markov models with distinct transition rates that result in average diffusivities that are additive (top) and non-additive (bottom).

From a continuum point of view, the dynamics of heterogeneous systems can be explored with models as simple as continuous time Markov chains. Imagine a one-dimensional system with two types of sites, either 1 or 2, that can be arbitrarily arranged along a line that wraps periodically in space (**Figure S1**). Note that the dimension and boundary conditions of the problem are chosen for convenience of illustration, though in principle the results are more general. Let us assume that if a random walker is at a type 1 site, the rate with which it leaves to either of the neighboring states is a, while this rate is b for type 2 sites. This results in a transition rate matrix corresponding to **Figure S1** (top) as

$$\boldsymbol{R} = \begin{bmatrix} -2a & a & 0 & 0 & 0 & a \\ a & -2a & a & 0 & 0 & 0 \\ 0 & a & -2a & a & 0 & 0 \\ 0 & 0 & b & -2b & b & 0 \\ 0 & 0 & 0 & b & -2b & b \\ b & 0 & 0 & 0 & b & -2b \end{bmatrix}$$

If we require that detailed balance be satisfied, the equilibrium occupation probability of any site 1 is related to the probability of any site 2 via

$$\wp_2 = \frac{a}{b}\wp_1$$

which fully specifies the problem if we assume that we start from an ensemble in which the equilibrium probabilities of each state are satisfied. For any time t, the time evolution of the occupation probabilities follows

$$\frac{d\vec{P}(t)}{dt} = \vec{P} \cdot \boldsymbol{R}$$

which is formally solved by

$$\boldsymbol{P}(t) = e^{\boldsymbol{R}t}$$

If we then measure the slope of the average mean-squared displacement at long times, the diffusivity is given by

$$D = \frac{\langle d^2(t+\delta t)\rangle - \langle d^2(t)\rangle}{2\delta t} = \frac{\sum_i \wp_i \sum_j \boldsymbol{P}(t+\delta t)_{i,j}(j-i)^2 - \sum_i \wp_i \sum_j \boldsymbol{P}(t)_{i,j}(j-i)^2}{2\delta t}$$

Note that the time must be long enough that a walker visits domains of both type 1 and 2, but short enough that the periodicity scale of the system remains comparatively large (this difficulty vanishes for infinitely sized systems). Though the above expressions are analytical, numerical solutions for arbitrary arrangements of sites 1 and 2 that address finite-size effects are easiest. Such results show that the diffusivity does not depend on the arrangement of sites of type 1 and 2, illustrating that a macroscopic model holds in which surface types have additive effects (i.e., the influence of a surface type is independent of neighboring site identities).

Additivity breaks down, however, as soon as we introduce edge effects between the domains, as shown in **Figure S1** (bottom). This is the equivalent of creating new types of sites with asymmetric rates of transition to adjacent sites. The asymmetry allows specific patterning of domains to influence the average diffusivity since the number of such edges can vary from one arrangement to another. The effect is similar to what is observed for interfaces with spatial chemical heterogeneity as described in the main text. At each surface site, one can consider analogous rates for moving to any neighboring surface site. If these rates are all the same, depending only on the nature of the site itself and not where the walker is moving, then additivity will hold. However, it is obvious that the arrangement of other nearby chemical groups can produce rate asymmetry, with subtle effects appearing as the free energy landscape changes across the surface. In this way, molecular scale heterogeneities likely make additive models for dynamical behavior break down, instead requiring brute force atomistic molecular dynamics simulations to evaluate the detailed dynamics.

2) Additional details of model systems

Our SAM systems are similar to the work of Garde and coworkers (1–3) and to Levine et al (4), but with some small modifications. In the work of Mondello (5) referenced by the Garde group, all bonds between united atoms were constrained, whereas here we only constrain bonds involving hydrogens. In previous parametrizations of the SAM surface, it appears based on prior descriptions of methodology that no angle potential energy function was applied to the C-O-H angle of hydroxyl-terminated chains, with the OPLS/AMBER parameters (6) for such an angle instead applied to the C-C-O angle. We apply the usual OPLS/AMBER C-C-O and C-O-H angle terms to these atoms, which somewhat limits the flexibility and hydrogen bonding ability of the hydroxyl head group. Chains are arranged with their sulfur atoms in a hexagonal close-packed

(HCP) pattern with nearest-neighbor spacing of 4.97 Å. We place a total of 72 chains on this lattice consisting of 9 rows (along the y dimension) and 8 columns (along the x dimension), resulting in surfaces of size 39.7600 x 38.7373 $Å^2$. An identical lattice defines the locations of LJ particles in those systems, with each layer along the direction perpendicular to the interface translated appropriately to maintain an overall HCP crystal structure.

The original SAM and LJ surfaces developed in our work entail a defect in the HCP lattice due to the use of an odd number of rows. This is manifested more obviously for the LJ surfaces, for which the particles are restrained in place whereas the flexible chains of the SAM are able to re-arrange so that the surface-water interface is still approximately HCP. Re-running the genetic algorithm at a hydrophilic fraction of 0.25 for re-constructed SAM and LJ surfaces with 10 rows and no defect (dimensions of 39.7600 x 43.0415 $Å^2$) makes no difference to the diffusivity values or range measured, and thus leads to no change in the conclusions presented.

3) Additional simulation details and analyses of molecular dynamics trajectories

Additional simulation details

For all MD simulations, the velocity Verlet method (7) is used to propagate the equations of motion with constraints placed on all bonds involving hydrogens through LINCS (8). The Smooth Particle Mesh Ewald algorithm, with default GROMACS parameters, treats long-range electrostatics (9).

Prior to each MD run, we use the following equilibration procedure: (1) The GROMACS *solvate* command centers and solvates a surface, followed by relaxation with steepest descent energy minimization and a very short NVT simulation with the Berendsen thermostat (10). (2) Short NPT equilibration with a Berendsen thermostat and semi-isotropic barostat (volume fluctuations only perpendicular to the surface) equilibrates the box size. (3) A second short pre-production NPT simulation evaluates average box size, which is then fixed. (4) Finally, a short NVT simulation utilizing a weak Nose-Hoover thermostat (11, 12) with 10 chains (13) precedes a 10 ns production run with the same settings and frames saved every 0.5 ps.

We separately remove center of mass momentum from the water and surface atoms every 30 time steps. This is sufficient to keep the α–cristobalite surfaces motionless relative to water. For the SAM surface, all united sulfur atoms are harmonically restrained in space, with each sulfur reference position scaled separately during NPT runs. LJ surface particles are also harmonically restrained, with the center of mass reference position of all particles rescaled during NPT simulations.

Calculation of diffusivities

To assess hydration water mobility, we determine the diffusivity via linear fits to the 2D mean-squared displacement (MSD) of water oxygen atoms parallel to the plane of the surface. Similar to earlier simulation studies (14, 15), we collect MSD data from water oxygen atoms as long as they remain within 8 Å of the surface, as defined by the most extreme surface heavy atom coordinates. For all systems, this includes approximately 2 hydration layers. We focus this calculation to a time window of 2-10 ps, which we find to be optimal for assessing lateral water dynamics near interfaces. At times shorter than this, waters do not exhibit diffusive behavior, while

at times beyond 10 ps, too many waters have left the surface and the slope of the MSD curve decreases as only the slowest-moving waters remain. Profiles of water lateral diffusivity in the direction perpendicular to the interface are obtained by the same procedure but using 3 Å slices with their centers spaced 0.5 Å apart along the interface normal.

Calculation of excess hard-sphere chemical potentials

We compute the spatially-varying excess chemical potential for hard-spheres of radius 3.3 Å using the Widom insertion technique (16). We discretize the simulation box into cells and count the number of times a randomly-inserted sphere in any cell is empty of water oxygens or surface heavy atoms. Profiles of μ_{ex}^{HS} result from placing spheres randomly within 0.5 Å slices perpendicular to the interface, with at least 5.04 x 10^8 insertions attempted for each slice. The negative log probability for a successful insertion (main text **Equation 1**) then provides the hard-sphere chemical potential and its spatial dependence. In the results presented, we reference hard-sphere excess chemical potentials to the case in bulk TIP4P-Ew water at 300 K and 1 bar, where $\mu_{ex}^{HS} \approx 10.476\ k_B T$.

Determination of interfaces and interface fluctuations

We use two definitions for the surface-water interface to spatially map surface hydrophobicity and hydrophilicity. The "mean" interface is defined by the box Z-coordinate where the average density of water becomes 0.3 of its bulk value (0.0332 $Å^{-3}$ for TIP4P-Ew water). "Instantaneous" interfaces are defined according to the procedure of Willard and Chandler (17), utilizing the same parameters except for a density fraction cut-off of 0.3 instead of 0.5. In short, each water oxygen is assigned a truncated and shifted Gaussian function, with the density field represented by the sum of such functions. The interface is the iso-surface within this field where the density is 0.3 of the bulk value. The *marching_cubes_lewiner* method (18) provided by the scikit-image package (19) identifies the iso-surface after first evaluating the density field on a 1.0 Å grid. For each X-Y bin on the grid (i.e. bins perpendicular to the surface slab), we prune the density field as necessary such that only two Z-coordinate iso-surface points are kept representing an upper and lower interface. Distances of waters to such an interface are defined by projecting the vector between a water oxygen and the nearest surface point onto the surface unit normal of said surface point.

Calculation of water re-orientation times

Water re-orientation times are determined by fitting a stretched exponential function to the time-decay of the cosine of the angle between the water dipole vector at an initial and later time. This was done only for waters remaining in the surface-water distance range of interest, as described by Debenedetti and coworkers (14). Profiles result from using 3 Å slices in the direction perpendicular to the interface with slice centers spaced 0.5 Å apart.

Calculation of contact angles

We calculate water contact angles by placing an equilibrated water box of 2000 TIP4P-Ew water molecules over SAM or LJ surfaces with dimensions of about 119 x 116 $Å^2$. Drops are

equilibrated under NVT conditions with molecular dynamics parameters as described in other sections. Once the drop shape stops changing (~1 ns for hydrophobic surfaces, and 5-10 ns for hydrophilic surfaces), a 10 ns production simulation is used to calculate the average drop geometry. The direction perpendicular to the interface is discretized into 1.0 Å bins, with the maximum water oxygen radial distance from the drop center of mass recorded at each timestep for each bin. At each frame, a fixed cut-off distance from the drop center is used to ensure that gas-phase waters were not considered. The average maximum radial distance versus height from the interface is fit to a second order polynomial, with the coefficient on the term linear in radius providing the inverse tangent of the contact angle.

4) Supporting Data and Figures

Table S1: Performance metrics for the surface design genetic algorithm

Surface	P-value for hypothesis that min diffusivity is greater than max using Welch's t-test*	Probability to observe min diffusivity surface pattern at least once through random sampling alone**	Probability to observe max diffusivity surface pattern at least once through random sampling alone**
cristobalite-$10\bar{1}$, $f_{Hyd} = 0.25$	5.20×10^{-12}	9.85×10^{-1}	1.07×10^{-2}
cristobalite-$10\bar{1}$, $f_{Hyd} = 0.375$	0.00×10^{0}	5.45×10^{-1}	5.95×10^{-1}
cristobalite-$10\bar{1}$, $f_{Hyd} = 0.50$	1.5×10^{-11}	9.98×10^{-1}	1.61×10^{-1}
cristobalite-$10\bar{1}$, $f_{Hyd} = 0.75$	5.90×10^{-3}	9.02×10^{-1}	1.00×10^{0}
SAM, $f_{Hyd} = 0.25$	5.20×10^{-7}	9.66×10^{-1}	5.21×10^{-6}
SAM, $f_{Hyd} = 0.50$	8.90×10^{-6}	9.83×10^{-1}	1.05×10^{-3}
LJ, $f_{Hyd} = 0.25$	1.80×10^{-1}	1.00×10^{0}	1.00×10^{0}
LJ, $f_{Hyd} = 0.50$	6.70×10^{-2}	1.00×10^{0}	1.00×10^{0}

*Uses Student's t-distribution, but does not assume equal variances of the two sets of samples

**Assumes a Gaussian distribution of diffusivities for randomly generated surfaces, with the binomial distribution then providing the probability of obtaining at least one equally or more anomalous diffusivity from a draw of the same size as the sample generated by the genetic algorithm optimization

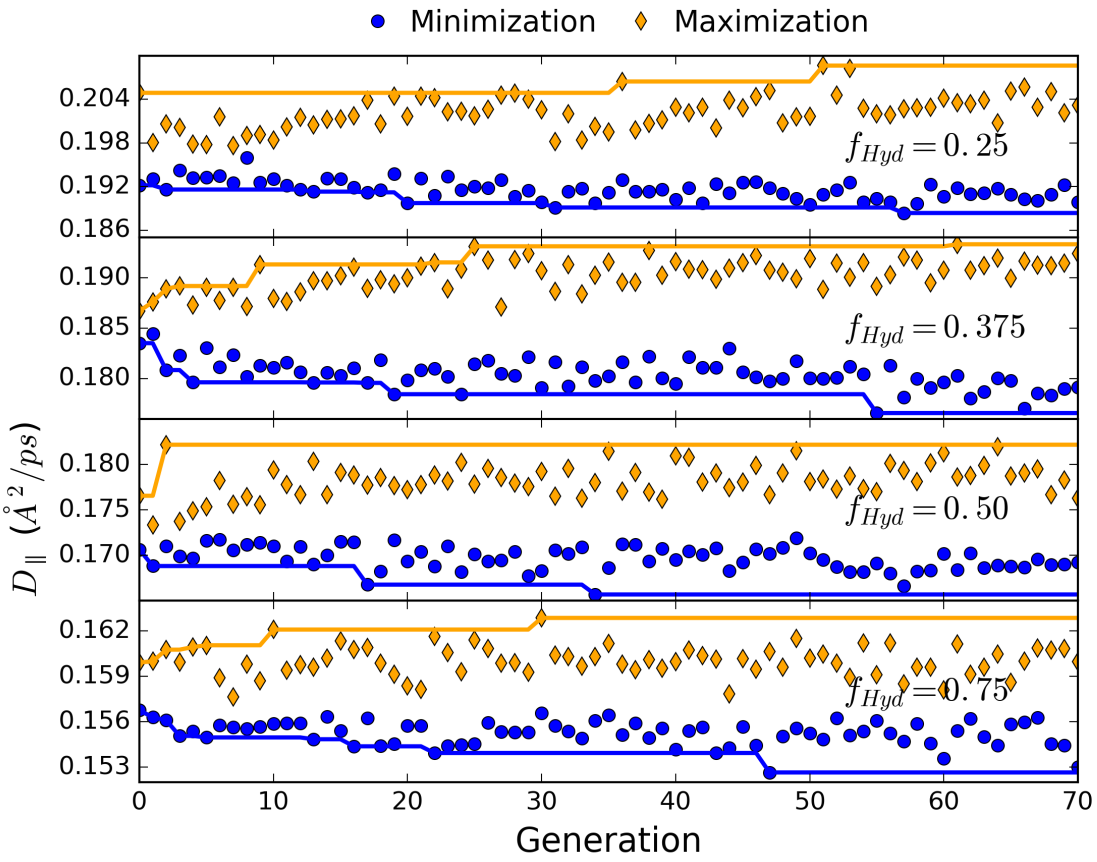

Figure S2: Evolution of the genetic algorithm for various densities of the cristobalite surface. Blue points are minimum diffusivity surfaces for a given generation during minimization runs of the algorithm while orange points are maximum diffusivity surfaces from a separate maximization run. Blue and orange lines represent convex hulls. Generations 0 through 10 represent surfaces that were randomly generated without any optimization. For these early points, the blue and orange give the minimum and maximum diffusivities in the set of initial random surfaces.

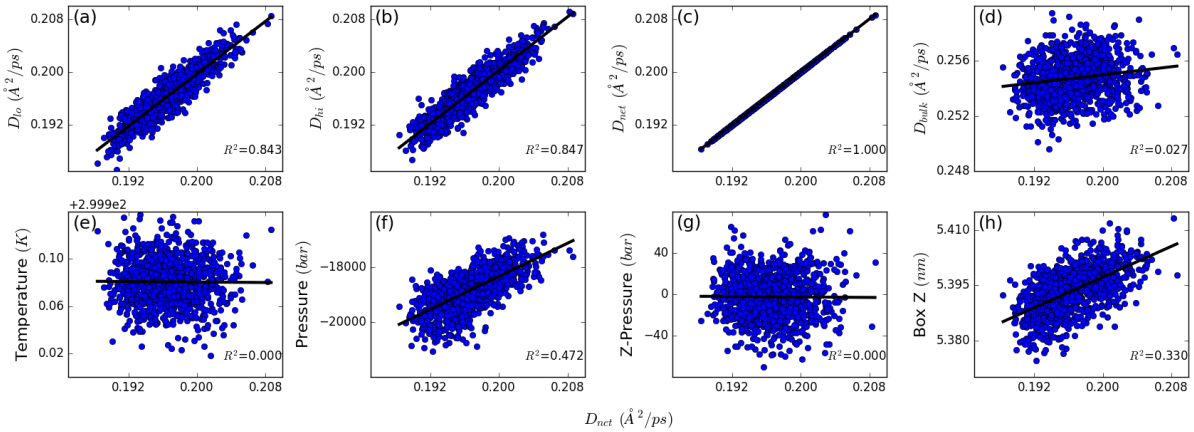

Figure S3: For cristobalite at a fraction of 0.25 of the maximum possible number of silanol groups, correlations are shown between the measured diffusivity parallel to the plane of the surface (D_{net} on the x-axis of all panels) and (a) the diffusivity measured below the surface (b) the diffusivity measured above the surface (c) D_{net} itself (d) the measured diffusivity far from the surface (e) the average simulation temperature (f) the total simulation pressure (g) the component of simulation pressure perpendicular to the surface and (h) the equilibrated box dimension perpendicular to the surface. Black lines are linear regressions to the data with R^2 values shown on each plot. While the average simulation box length perpendicular to the interface weakly correlates with diffusivity (larger volumes for faster diffusion), this is due to variation in effective surface hydrophobicity that modulates interfacial width at constant pressure. For cristobalite surfaces at low density (here and in **Figure S4**), it appears that the total system pressure (including all Cartesian directions) correlates weakly with the mobility of hydration water. Noting that the surface area is fixed in the plane of the interface and not relaxed during constant-pressure equilibration, it is clear that arrangements of silanol groups can affect the in-plane strain of these surfaces by changing the bonding pattern. Thus, a weak correlation is observed between measured diffusivity and total simulation pressure as the pattern of functional groups in cristobalite is adjusted.

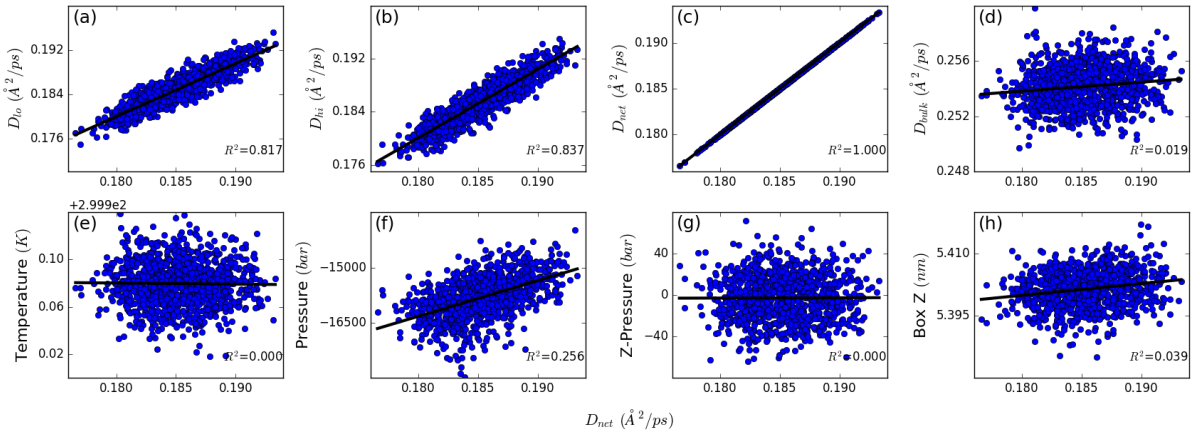

Figure S4: Same as **Figure S3**, but for cristobalite surfaces at a fraction of 0.375 of the maximum possible number of silanol groups.

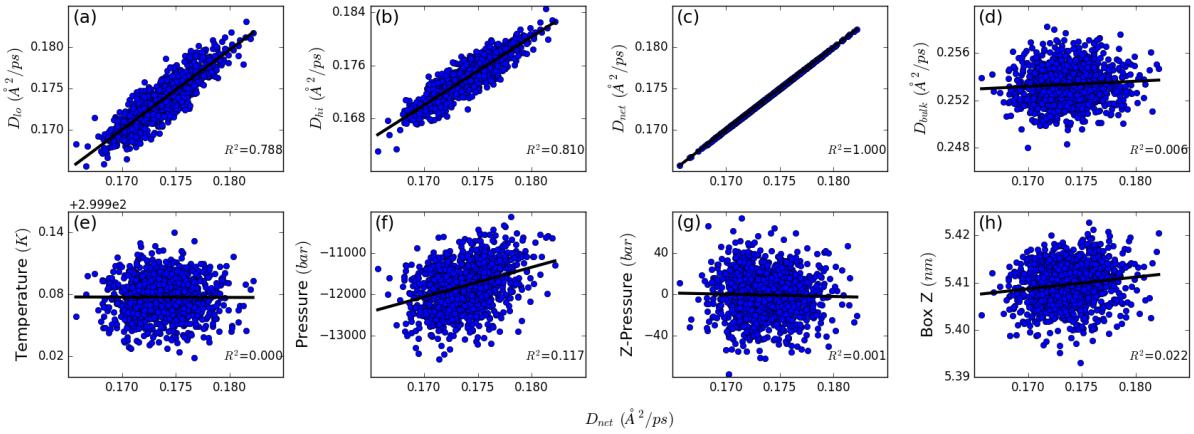

Figure S5: Same as **Figure S3**, but for cristobalite surfaces at a fraction of 0.50 of the maximum possible number of silanol groups.

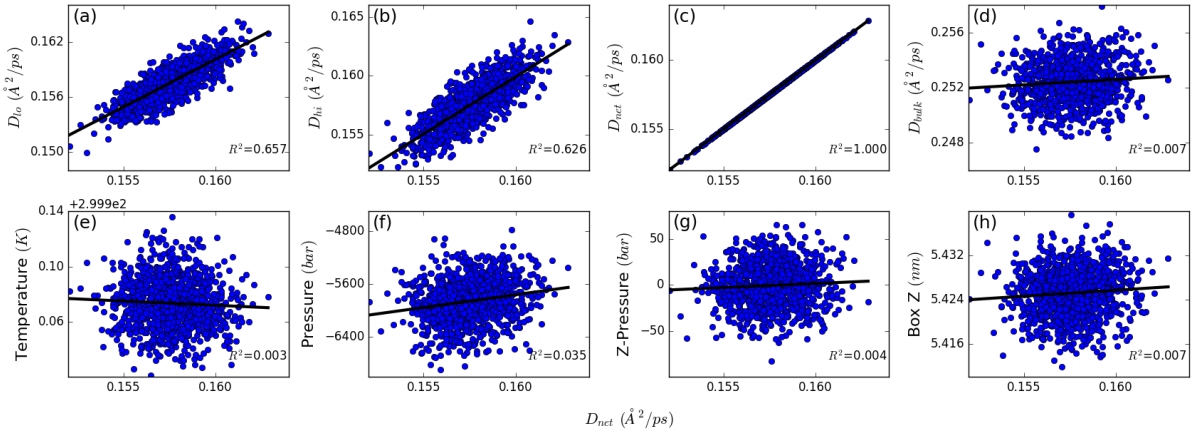

Figure S6: Same as **Figure S3**, but for cristobalite surfaces at a fraction of 0.75 of the maximum possible number of silanol groups.

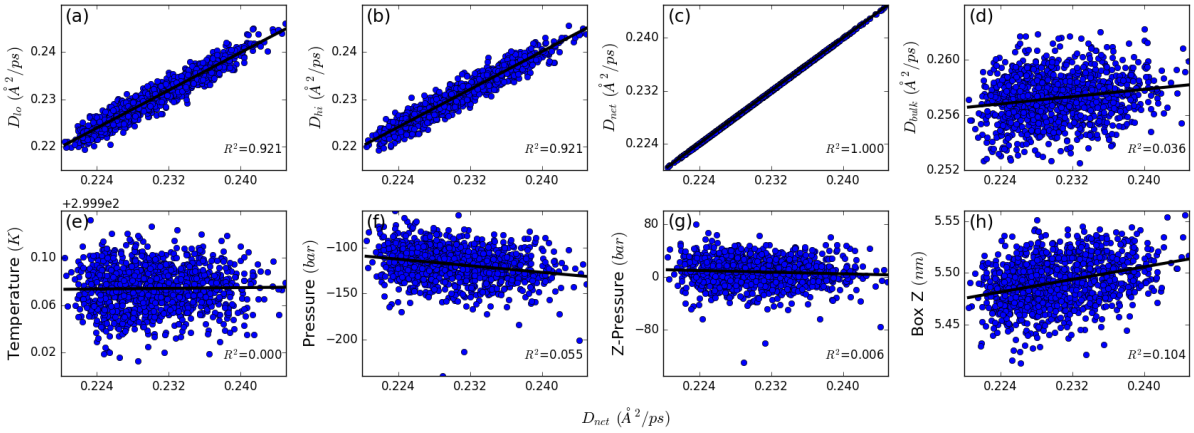

Figure S7: Same as **Figure S3**, but for the SAM surface at an OH-terminated chain fraction of 0.25.

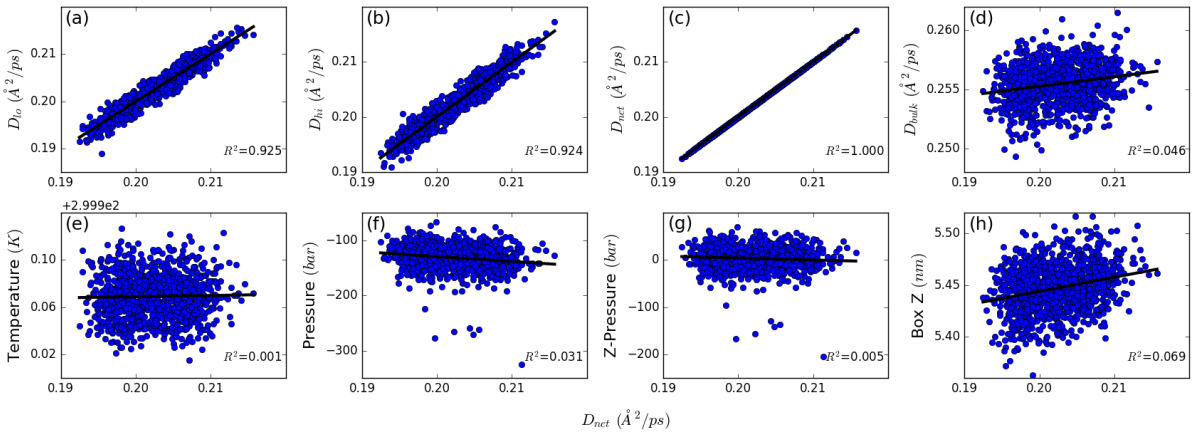

Figure S8: Same as **Figure S3**, but for the SAM surface at an OH-terminated chain fraction of 0.5.

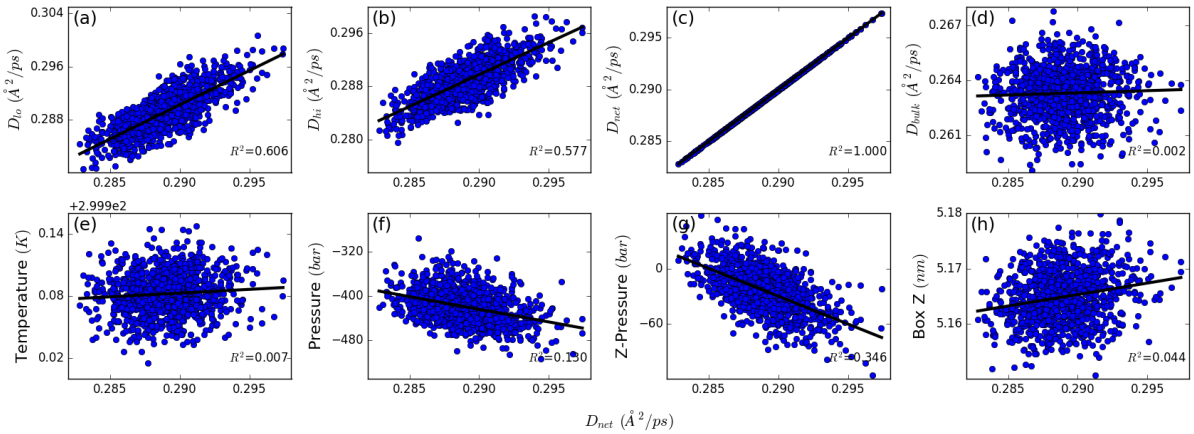

Figure S9: Same as **Figure S3**, but for the LJ surface at an attractive particle fraction of 0.25. For LJ surfaces at all coverages, the pressure normal to the interface correlates with the diffusivity. It should be noted, however, that the range of Z-pressures observed is the same as for other surface types. With surface patterning playing little to no role in determining diffusivity for the LJ surfaces (see main text), system properties that vary with system set-up and equilibration, like the Z-pressure, are more statistically significant, whereas these same fluctuations primarily introduce statistical noise to diffusivities on SAMs or cristobalite.

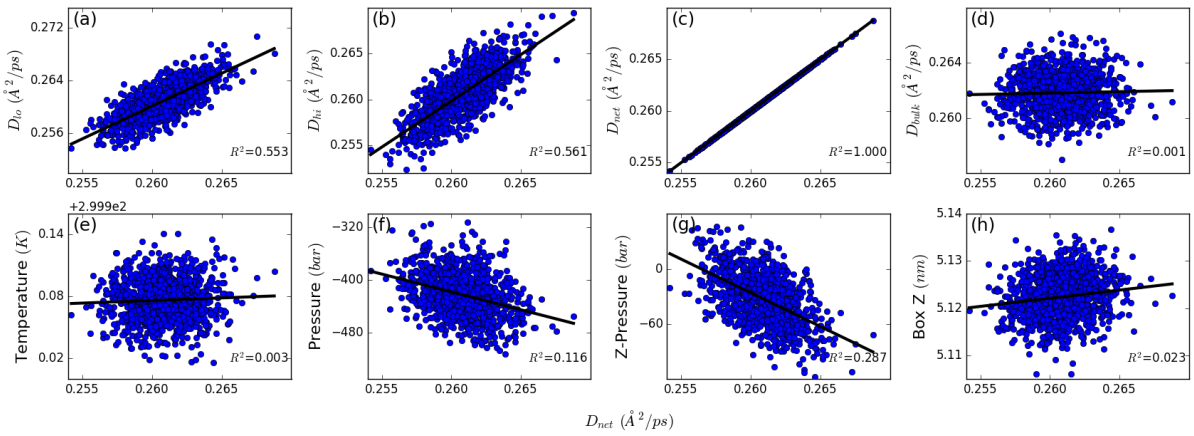

Figure S10: Same as **Figure S3**, but for the LJ surface at an attractive particle fraction of 0.50.

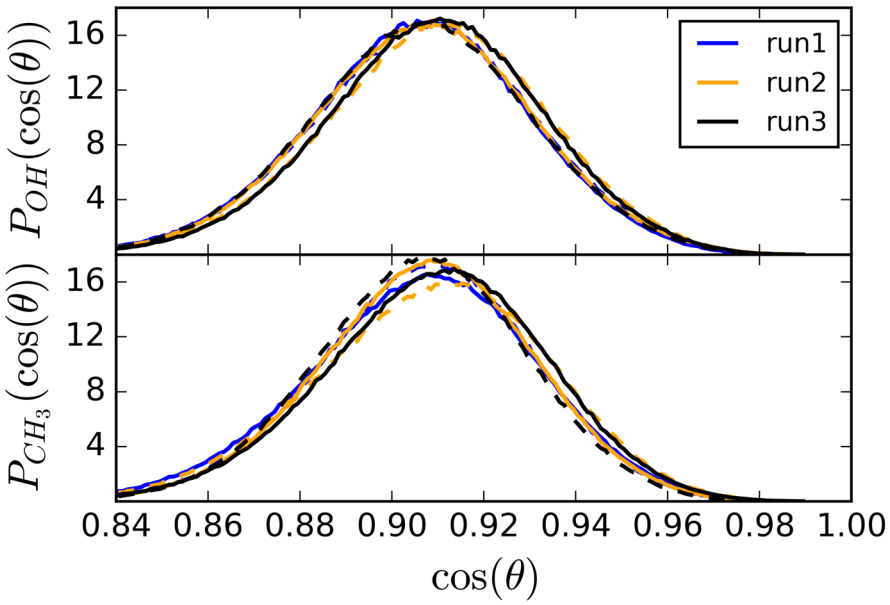

Figure S11: Probability density distributions for the cosine of the tilt angle for SAM surfaces (the angle between the direction perpendicular to the interface and the vector between the restrained sulfur atom and the last united atom before the head group). The top panel shows the distribution for OH-terminated chains while the bottom for CH_3-terminated ones. Each color represents an independent simulation, with solid lines corresponding to the minimum-diffusivity surface pattern and dashed lines to the maximum-diffusivity surface. Variations in tilt angle distributions between independent simulation runs are larger than any differences observed between the minimum and maximum diffusivity patterns, implying negligible changes in average surface structure. Additionally, freezing all but the head-group atoms of SAM chains results in nearly identical diffusivity ranges in **Figure 2** (also see **Figure S14**), although the diffusivity is slightly reduced overall (by 0.014 to 0.032 Å^2/ps). The decreases in diffusivity are more pronounced at higher surface densities and stem from the reduction, due to chain rigidification, in hydrogen bonding between surface hydroxyls, increasing the availability of water-surface hydrogen bonds.

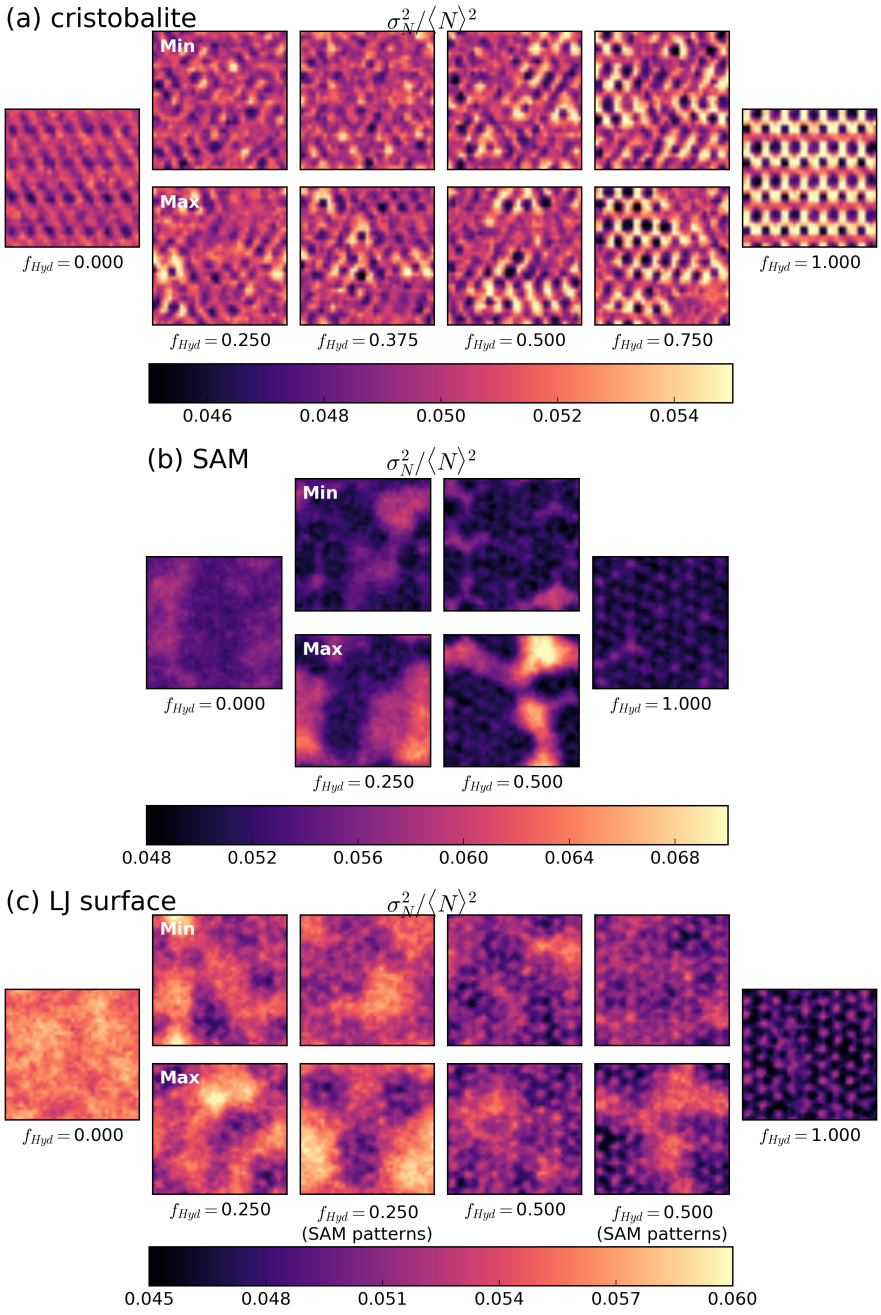

Figure S12: Top-views for the normalized fluctuations in the number of waters within probes of radius 3.3 Å sitting with their edge at the mean interface of (a) cristobalite, (b) SAM, and (c) LJ surfaces. The mean interface is defined as the location at which the average water density is 0.3 of the bulk value for TIP4P-Ew of 0.0332 $Å^{-3}$. For LJ surfaces, the middle column considers the patterns determined by optimization of the SAM surface at the same density. It is clear that fluctuations are enhanced near larger patches of more hydrophobic groups on both SAM and LJ surfaces. For cristobalite, both very small and large fluctuations are observed near hydroxyl groups. This in turn leads to higher excess chemical potentials for hard-sphere probes near the surface.

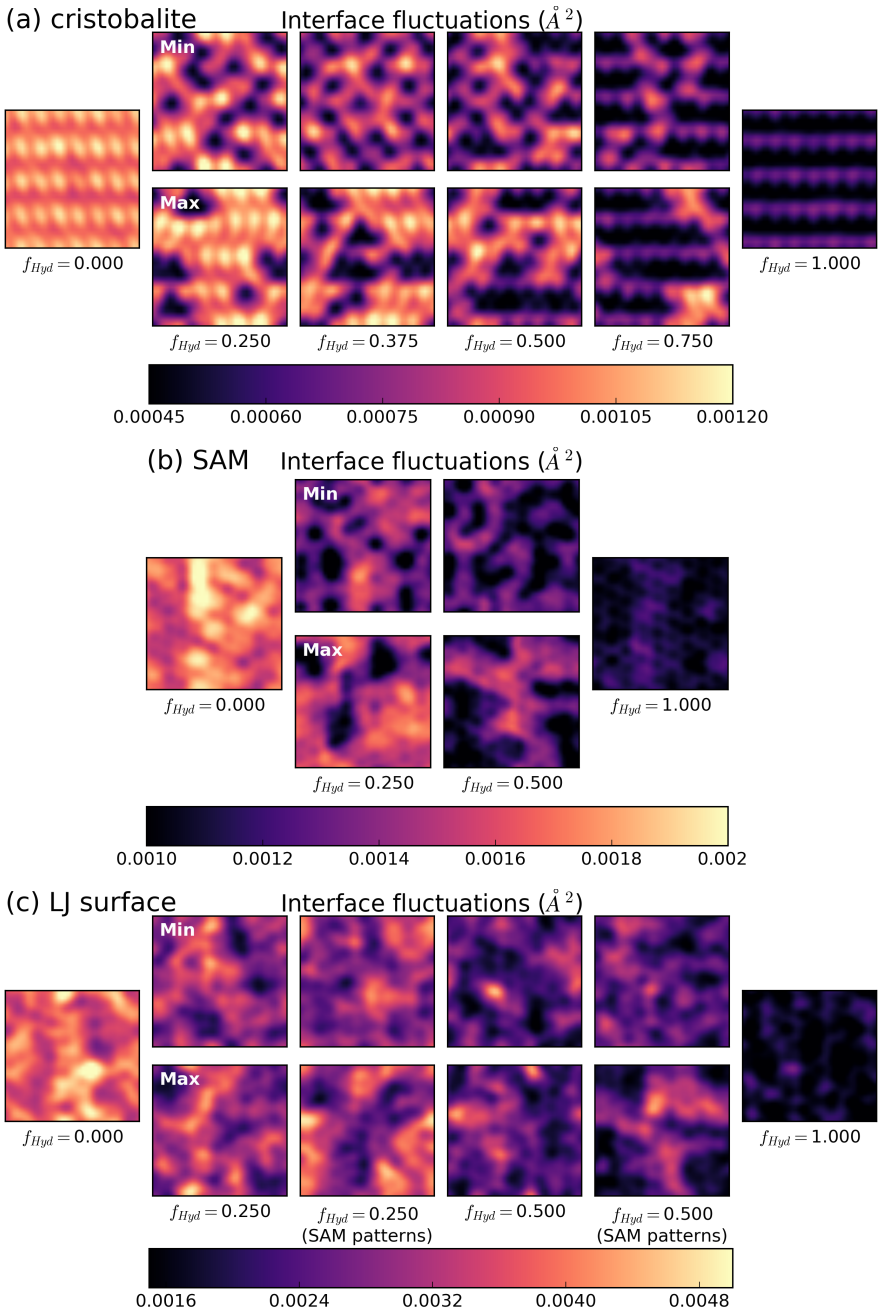

Figure S13: Top-views of height fluctuations for instantaneous interfaces for cristobalite (top), SAM (middle), and LJ surfaces (bottom). The instantaneous interface is defined according to the methodology of Willard and Chandler,(17) but with a surface density cut-off of 0.3 of the bulk value of 0.0332 Å^{-3} for TIP4P-Ew water.

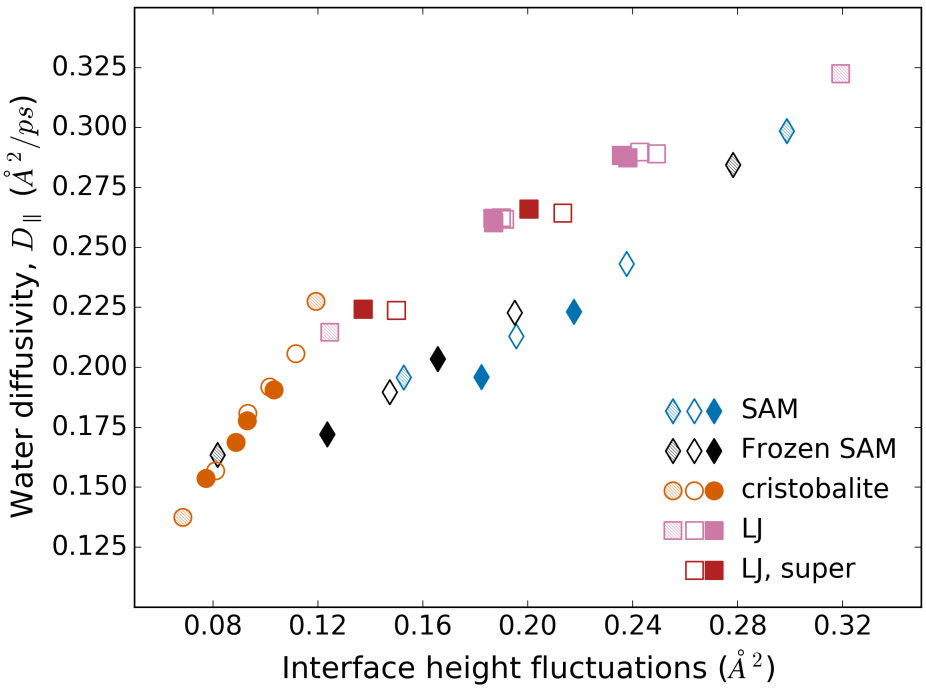

Figure S14: Correlation between lateral diffusivities of water near minimum and maximum diffusivity surfaces and the average interfacial height fluctuations using the method of Willard and Chandler (17) to define the instantaneous interface. Hatched points represent surfaces at hydrophilic fractions of zero or one, filled points minimum diffusivity surfaces, and open points maximum diffusivity surfaces. The correlation is enhanced by freezing the united atoms (non-head-group atoms) of the SAM chains, as shown by the black versus blue diamond symbols.

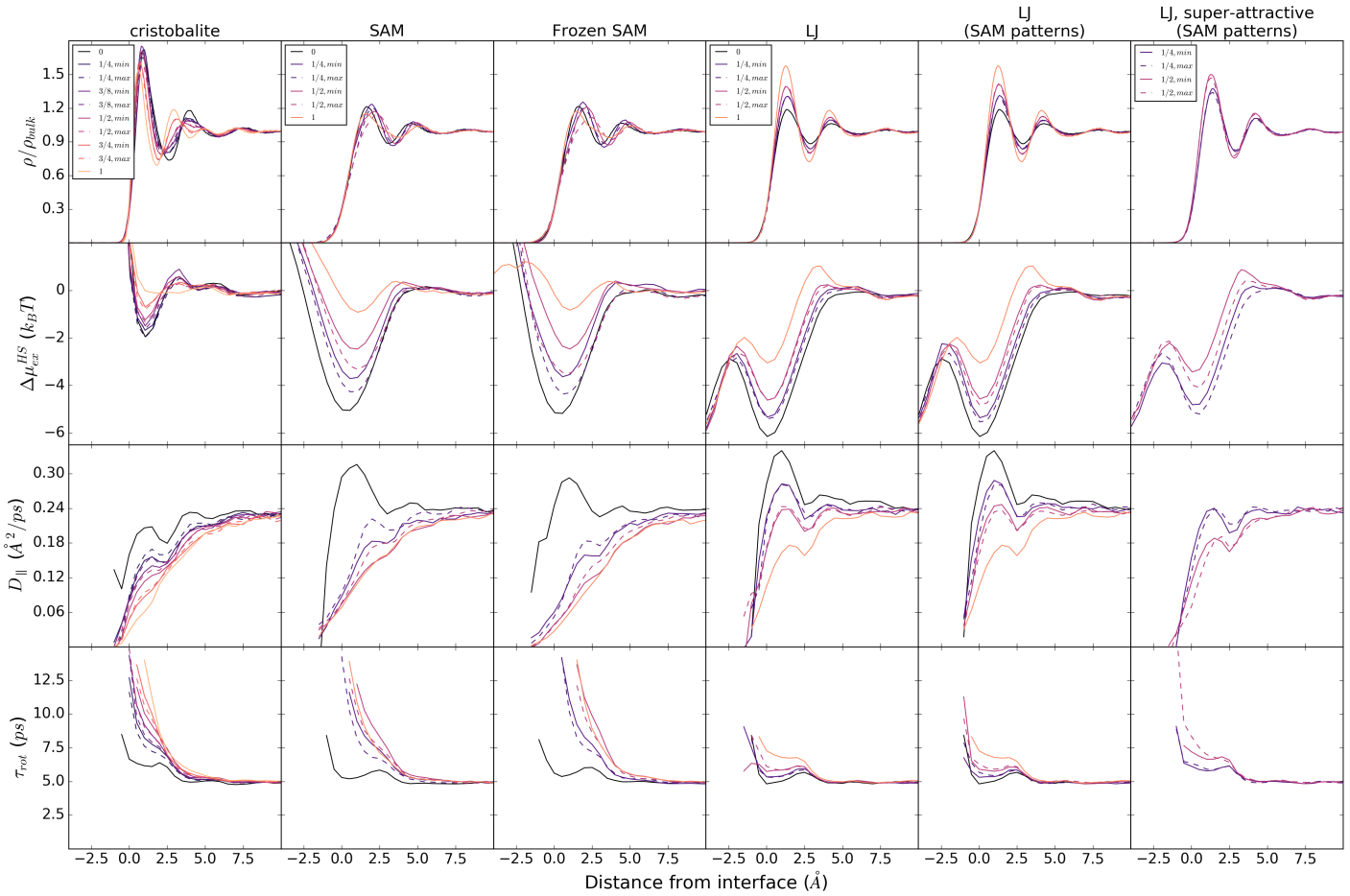

Figure S15: Profiles for water density, hard-sphere excess chemical potential, water diffusivity parallel to the interface, and water dipole re-orientation times as a function of distance from the mean interface (defined as the point where water reaches 0.3 of its bulk value). Dipole re-orientation times were computed by fitting the time decay of the cosine of the angle between the original dipole moment vector with a stretched exponential, as in the work of Debenedetti and co-workers (14). The qualitative trends are the same if an instantaneous interface definition is used instead. In general, hard-sphere excess chemical potentials are correlated with both lateral diffusivities and water re-orientation times. However, these spatial correlations seem to break down very close to the solid surfaces, where geometric constraints due to the presence of surface atoms seem to impact hard-sphere insertion probabilities and water mobility in an unrelated manner.

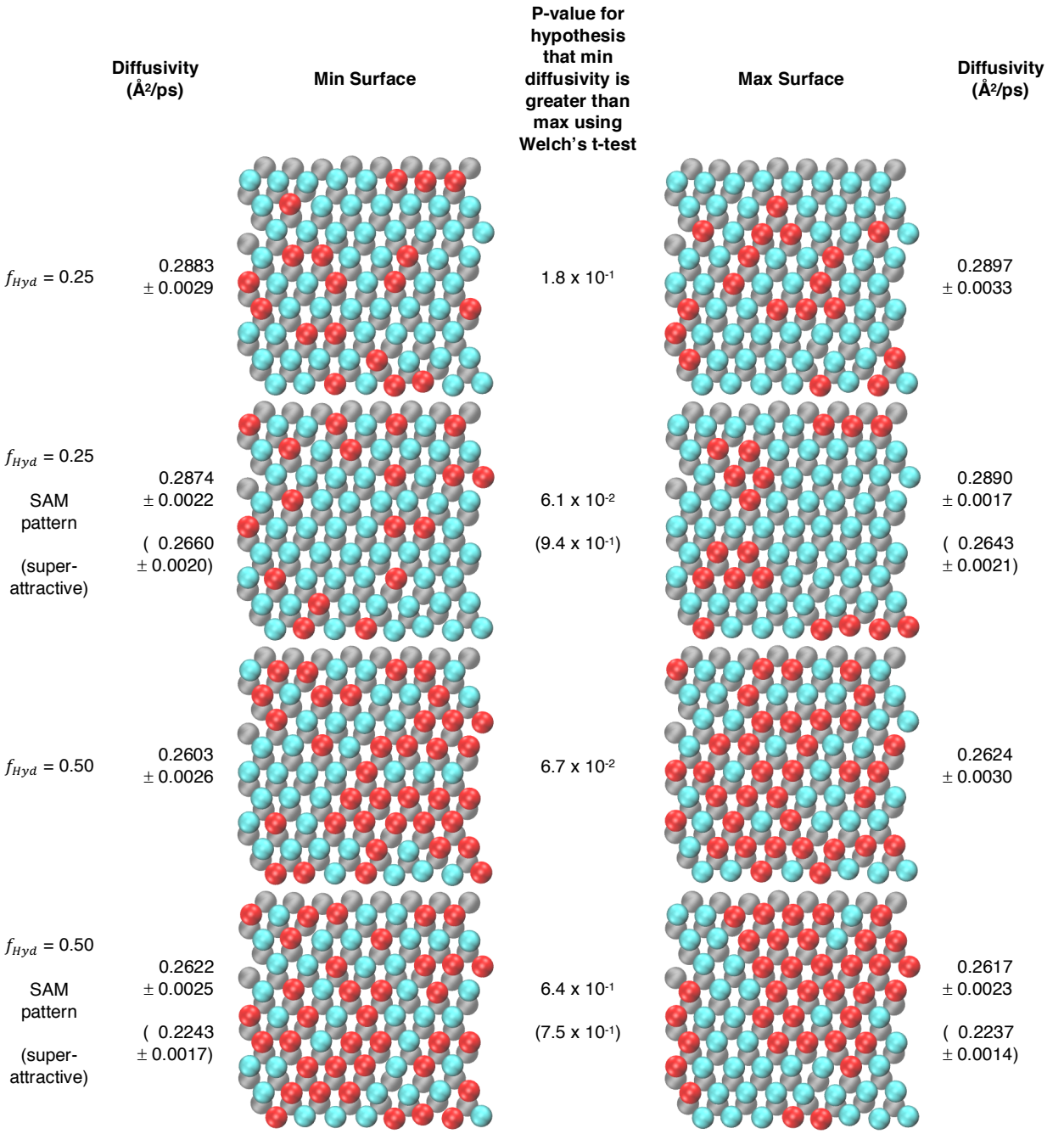

Figure S16: Minimum and maximum diffusivity LJ surfaces from the genetic algorithm optimization, as well as LJ surfaces with the optimum patterns from SAM surfaces (values for these surfaces with super-attractive particles are shown in parentheses). Lateral diffusivities are averaged over 8 independent runs. The central column shows the P-value for the hypothesis that the minimum value is greater than the maximum using Welch's t-test, which in all cases shows that there are no statistical differences between the minimum and maximum surfaces.

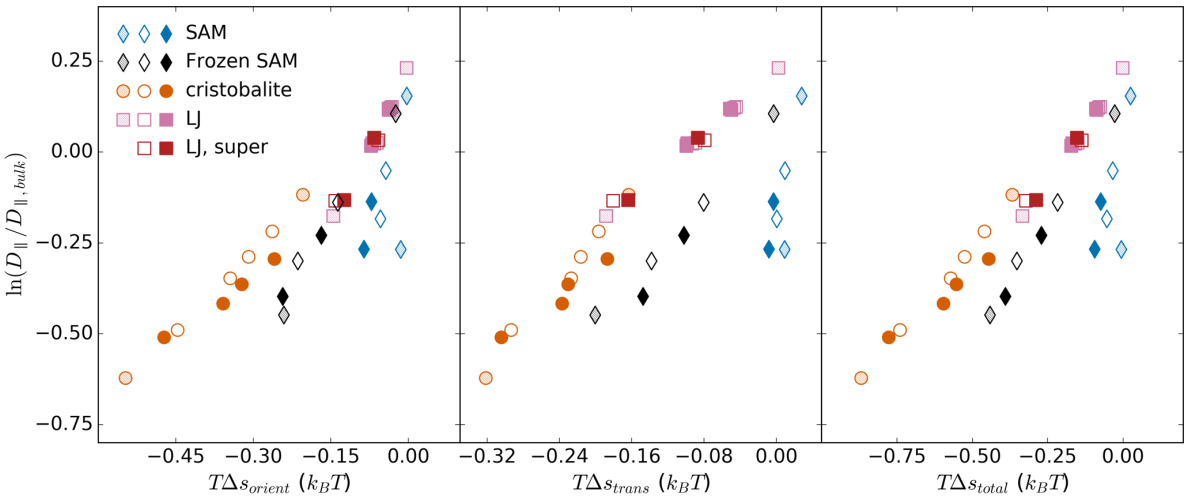

Figure S17: The natural logarithm of the ratio of lateral diffusivity to its bulk value is shown as a function of the orientational, translational, and total contributions to hydration water entropy 8 Å from the interface (reported relative to bulk waters). Filled symbols are minimum diffusivity surfaces (distributed hydrophilic groups), open symbols are maximum diffusivity surfaces (clumped hydrophilic groups), and hatched symbols are surfaces with hydrophilic fractions of 0.0 or 1.0. When all atoms in SAM chains except head-groups are frozen, all correlations are enhanced, mainly due to more negative translational contributions.

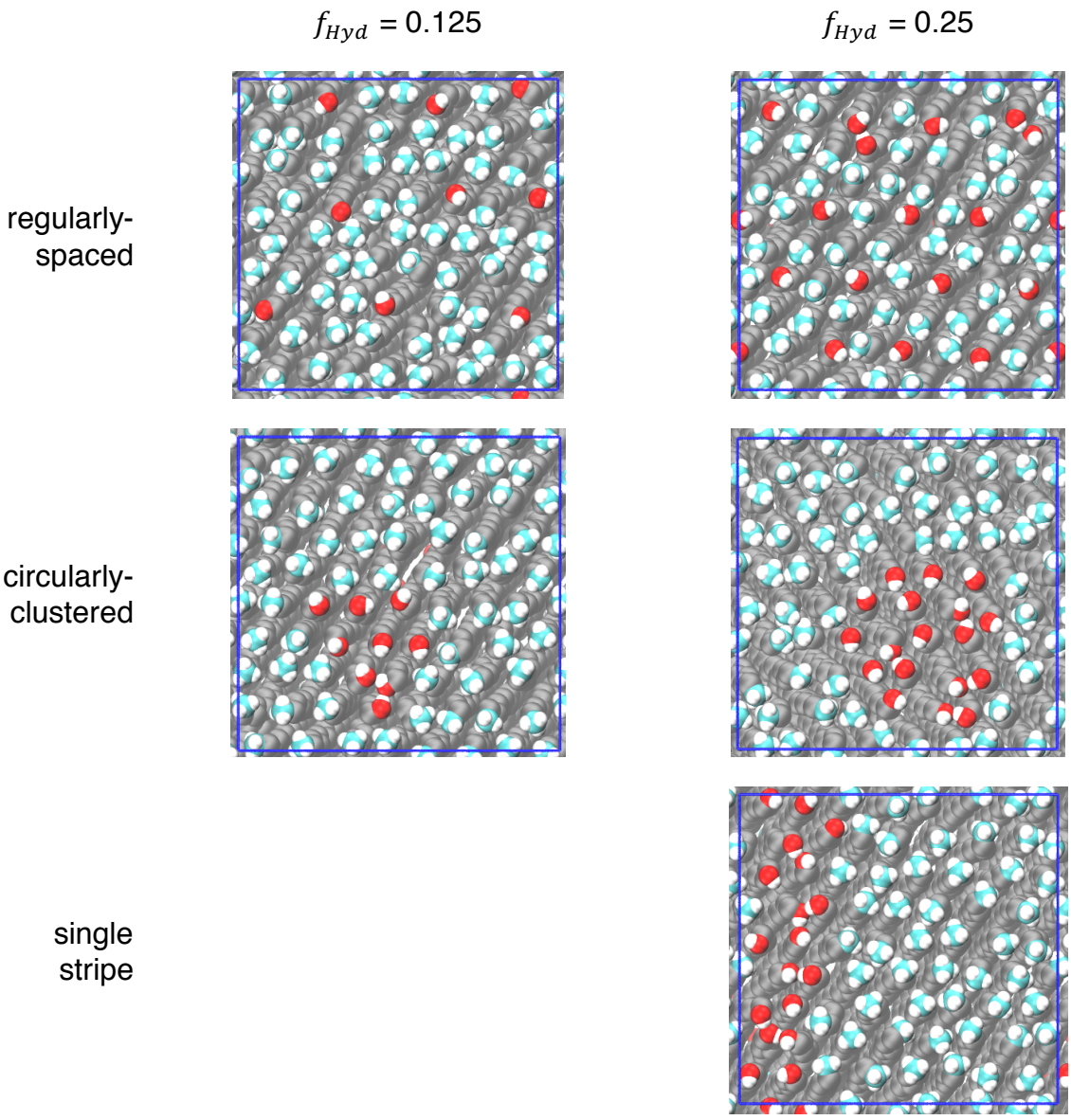

Figure S18: Top-views of surfaces with specific patterns to further test the genetic algorithm. Regularly-spaced and circularly-clustered hydroxyl groups represent the extremes of diffusivity suggested by populations of surfaces discovered by the genetic algorithm. The single stripe configuration for 25% hydroxyl-terminated chains represents a more exotic patterned to rudimentarily test for anisotropy in the diffusivity in directions parallel to and perpendicular to the stripe, as discussed in the main text.

5) Dataset file descriptions

Dataset S1: Example Cristobalite Structure File
Dataset S2: Example Cristobalite Topology File
Dataset S3: Example SAM Structure File
Dataset S4: Example SAM Topology File
Dataset S5: Example LJ Structure File
Dataset S6: Example LJ Topology File
Dataset S7: SAM Overall Parameter File
Dataset S8: SAM Bonded Parameter File
Dataset S9: SAM Non-bonded Parameter File
Dataset S10: Methyl-terminated SAM Parameter File
Dataset S11: Hydroxyl-terminated SAM Parameter File
Dataset S12: VDW Radii for Solvation
Dataset S13: Energy Minimization Input File
Dataset S14: Equilibration NVT1 Input File
Dataset S15: Equilibration NPT Input File
Dataset S16: Equilibration NVT2 Input File
Dataset S17: Production Input File
Dataset S18: Cristobalite Simulation Run File
Dataset S19: SAM Simulation Run File
Dataset S20: LJ Simulation Run File

References

1. Godawat R, Jamadagni SN, Garde S (2009) Characterizing hydrophobicity of interfaces by using cavity formation, solute binding, and water correlations. *Proc Natl Acad Sci* 106(36):15119–15124.
2. Shenogina N, Godawat R, Keblinski P, Garde S (2009) How wetting and adhesion affect thermal conductance of a range of hydrophobic to hydrophilic aqueous interfaces. *Phys Rev Lett* 102(15). doi:10.1103/PhysRevLett.102.156101.
3. Jamadagni SN, Godawat R, Garde S (2011) Hydrophobicity of Proteins and Interfaces: Insights from Density Fluctuations. *Annu Rev Chem Biomol Eng* 2(1):147–171.
4. Levine ZA, et al. (2016) Surface force measurements and simulations of mussel-derived peptide adhesives on wet organic surfaces. *Proc Natl Acad Sci*:201603065.
5. Mondello M, Grest GS, Webb EB, Peczak P (1998) Dynamics of n-alkanes: Comparison to Rouse model. *J Chem Phys* 109(2):798–805.
6. Cornell WD, et al. (1995) A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J Am Chem Soc* 117(19):5179–5197.
7. Swope WC, Andersen HC, Berens PH, Wilson KR (1982) A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *J Chem Phys* 76(1):637–649.
8. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) LINCS: A Linear Constraint Solver for molecular simulations. *J Comput Chem* 18(12):1463–1472.
9. Essmann U, et al. (1995) A smooth particle mesh Ewald method. *J Chem Phys* 103(19):8577–8593.
10. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. *J Chem Phys* 81(8):3684.
11. Nosé S (1984) A molecular dynamics method for simulations in the canonical ensemble. *Mol Phys* 52(2):255–268.
12. Hoover WG (1985) Canonical dynamics: Equilibrium phase-space distributions. *Phys Rev A* 31(3):1695–1697.
13. Martyna GJ, Klein ML, Tuckerman M (1992) Nosé-Hoover chains: The canonical ensemble via continuous dynamics. *J Chem Phys* 97(4):2635–2643.
14. Romero-Vargas Castrillón S, Giovambattista N, Aksay IA, Debenedetti PG (2009) Effect of Surface Polarity on the Structure and Dynamics of Water in Nanoscale Confinement. *J Phys Chem B* 113(5):1438–1446.
15. Argyris D, Cole DR, Striolo A (2009) Dynamic Behavior of Interfacial Water at the Silica Surface. *J Phys Chem C* 113(45):19591–19600.
16. Widom B (1963) Some Topics in the Theory of Fluids. *J Chem Phys* 39(11):2808–2812.
17. Willard AP, Chandler D (2010) Instantaneous liquid interfaces. *J Phys Chem B* 114(5):1954–1958.
18. Lewiner T, Lopes H, Vieira AW, Tavares G (2003) Efficient Implementation of Marching Cubes' Cases with Topological Guarantees. *J Graph Tools* 8(2):1–15.
19. van der Walt S, et al. (2014) scikit-image: image processing in Python. *PeerJ* 2:e453.