Protein Hydration Waters are Susceptible to Unfavorable Perturbations

Nicholas B. Rego,[†] Erte Xi,[‡] and Amish J. Patel[*,‡]

[†]*Biochemistry & Molecular Biophysics Graduate Group, University of Pennsylvania, Philadelphia, PA 19104*

[‡]*Department of Chemical & Biomolecular Engineering, University of Pennsylvania, Philadelphia, PA 19104*

E-mail: amish.patel@seas.upenn.edu

Abstract

The interactions of a protein, its phase behavior, and ultimately, its ability to function, are all influenced by the interactions between the protein and its hydration waters. Here we study proteins with a variety of sizes, shapes, chemistries, and biological functions, and characterize their interactions with their hydration waters using molecular simulations and enhanced sampling techniques. We find that akin to extended hydrophobic surfaces, proteins situate their hydration waters at the edge of a dewetting transition, making them susceptible to unfavorable perturbations. We also find that the strength of the unfavorable potential needed to trigger dewetting is roughly the same for all the proteins studied here, and depends primarily on the width of the hydration shell being perturbed. Our findings establish a framework for systematically classifying protein patches according to how favorably they interact with water.

Introduction

Biomolecular binding processes involve disrupting protein-water interactions and replacing them with direct interactions between the binding partners. Thus, protein-water interactions influence the thermodynamics and kinetics of protein interactions,[1–10] as well as the stability and phase behavior of protein solutions.[11–18] In this article, we characterize the overall interactions between proteins and their hydration waters by using specialized molecular simulations that employ an unfavorable potential to displace water molecules from the vicinity of a protein. Because displacing interfacial waters disrupts surface-water interactions, the less favorable those interactions (e.g., for hydrophobic surfaces), the easier it is to displace the interfacial waters.[19–21]

Indeed, both theory[22–25] and molecular simulations[26,27] have shown that the rare, low-density fluctuations, which are accessed when interfacial waters are displaced, are substantially more probable adjacent to a hydrophobic surface (than at a hydrophilic surface). Such enhanced low-density fluctuations situate water near extended hydrophobic surfaces at the edge of a collective dewetting transition, making them susceptible to unfavorable potentials.[22,26,27] The proximity of interfacial waters to a dewetting transition is also reflected in other collective interfacial properties, such as compressibility, transverse density correlations, and the distribution of water dipole orientations, among others.[19,22,26–35]

In contrast with simple hydrophobic or hydrophilic surfaces, proteins display nanoscopic chemical and topographical patterns, which influence their interactions with water in non-trivial ways.[20,36–48] By interrogating how pro-

tein hydration waters respond to an unfavorable potential, here we find that the hydration shells of diverse proteins are also situated at the edge of a dewetting transition. Such a resemblance of protein hydration shells to extended hydrophobic surfaces appears to arise from the fact that – even for protein surfaces that are enriched in polar and charged residues – roughly half the surface consists of hydrophobic atoms. Our findings, obtained by studying proteins across a broad range of sizes, chemistries, and functions, suggest that susceptibility to unfavorable perturbations is a common feature of soluble proteins with well-defined folded structures.

We also find that the strength of the unfavorable potential needed to trigger dewetting is inversely proportional to the width of the hydration shell that we choose to perturb, but is otherwise similar across all the proteins that we study. Our findings lay the groundwork for systematically disrupting protein-water interactions, and uncovering regions of proteins that have the weakest (hydrophobic) and the strongest (hydrophilic) interactions with water. A knowledge of the most hydrophobic protein regions could enable the prediction of the interfaces through which protein interact with one another.[49–52] Similarly, uncovering the most hydrophilic protein patches could result in the discovery of novel super-hydrophilic chemical patterns.[53]

Results and Discussion

Water near uniform, flat surfaces

To illustrate the molecular signatures of surface hydrophobicity, we first review the contrasting behavior of water near CH_3-terminated (hydrophobic) and OH-terminated (hydrophilic) self-assembled monolayer (SAM) surfaces, as was done in ref.[27] The SAM surfaces are not only flat and uniform, and thereby considerably simpler than proteins, but their hydrophobicity can also be defined unambiguously using a macroscopic measure, such as the water droplet contact angle. We focus on water molecules in a cylindrical observation volume, v, at the SAM-water interface, as shown in Figure 1a. We choose a radius, $R_v = 2$ nm, and a width, $w = 0.3$ nm, for the cylindrical v; with this choice, v at either SAM surface contains an average of roughly 120 waters. Following previous work,[27] we then perturb the interfacial waters in v by applying an unfavorable biasing potential, ϕN_v, where ϕ represents the strength of the potential, and N_v is the number of coarse-grained waters in v; a more precise definition of N_v is included in the Supporting Information.[56] The potential imposes an energetic penalty that increases linearly with N_v, so that as ϕ is increased, waters are displaced from v, resulting in a decrease in the average water numbers, $\langle N_v \rangle_\phi$, next to both SAM surfaces; see Figure 1b. The decrease in $\langle N_v \rangle_\phi$ with increasing ϕ is linear for the hydrophilic SAM. In comparison, the corresponding $\langle N_v \rangle_\phi$-values for the hydrophobic SAM are smaller for every ϕ, highlighting the relative ease of displacing waters. Moreover, the decrease in $\langle N_v \rangle_\phi$ with increasing ϕ is sensitive (or sigmoidal) near the hydrophobic surface rather than gradual (and linear) as it is near the hydrophilic surface. This contrast can be seen even more clearly in Figure 1c, which shows the susceptibility, $\chi_v \equiv -\partial \langle N_v \rangle_\phi / \partial(\beta\phi)$, as a function of $\beta\phi$; here, $\beta = 1/k_B T$, k_B is the Boltzmann constant, and T is the system temperature. The susceptibility is nearly constant for the hydrophilic surface. However, it shows a pronounced peak for the hydrophobic surface, suggesting that a collective dewetting of the interfacial waters can be triggered when a sufficiently strong unfavorable potential is applied.[27]

Perturbing the protein hydration shell

Unlike the uniform SAM surfaces, proteins are heterogeneous, rugged, and amphiphilic. In ref.,[27] enhanced low-density fluctuations, which underpin susceptibility to dewetting, were observed adjacent to a hydrophobic patch on the BphC protein, whereas bulk-like (Gaussian) fluctuations were seen near a hydrophilic patch on the protein. However, in contrast with individual protein patches, protein surfaces on the

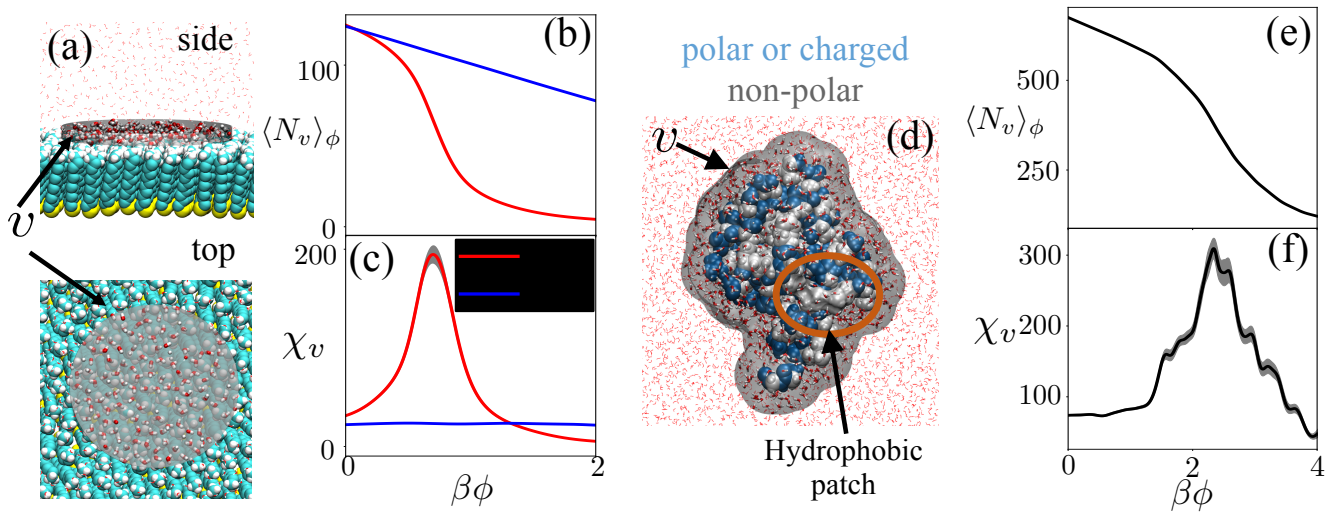

Figure 1: How interfacial waters respond to an unfavorable potential. (a) Simulation snapshots of the CH_3-terminated SAM-water interface. The SAM atoms are shown in space-fill representation; the N_v waters in the interfacial observation volume, v, are shown in the licorice representation, and the rest as lines. The cylindrical v is chosen to have a radius, $R_v = 2$ nm, and a width, $w = 0.3$ nm. (b) In response to an unfavorable biasing potential, ϕN_v, the average number of interfacial waters, $\langle N_v \rangle_\phi$, decreases at both the hydrophobic CH_3-terminated and the hydrophilic OH-terminated SAM surfaces. However, as the potential strength, ϕ is increased, $\langle N_v \rangle_\phi$ near the hydrophilic SAM decreases gradually; whereas, $\langle N_v \rangle_\phi$ near the hydrophobic SAM decreases sharply.[21,22,26,27] (c) These differences are also evident in the ϕ-dependence of the susceptibility, $\chi_v \equiv -\partial \langle N_v \rangle_\phi / \partial(\beta\phi)$, which is roughly constant for the hydrophilic SAM, but shows a marked peak for the hydrophobic SAM. Thus, interfacial waters at a hydrophobic surface are susceptible to unfavorable perturbations.[22,23,26,27] (d) Simulation snapshot of the ubiquitin protein, highlighting the observation volume, v, which contains waters in the first hydration shell of the protein. Protein atoms are shown in space-fill representation, and colored according to their atom types following ref.;[54] waters in v are shown as licorice, and the rest as lines. The well-characterized hydrophobic patch of ubiquitin, which mediates its interactions with other proteins, is also shown.[55] (e) The average number of protein hydration waters, $\langle N_v \rangle_\phi$, decreases in a sigmoidal manner as ϕ is increased. (f) The corresponding susceptibility, χ_v, displays a maximum, suggesting that protein hydration waters are also susceptible to unfavorable perturbations, akin to those near hydrophobic surfaces.

whole tend to have both hydrophilic residues to ensure solubility in water and hydrophobic residues to drive protein interactions. Given the overall amphiphilicity of proteins, how might their hydration waters respond to an unfavorable potential? Should we expect its hydration waters to be displaced gradually like the hydrophilic SAM surface? Or should we expect the protein hydration waters to undergo collective dewetting like the hydrophobic SAM surface?

To address these questions, we first study ubiquitin, a highly-conserved protein involved in numerous signaling pathways, including protein degradation.[57] Although it is a relatively small protein (76 amino acid residues), ubiquitin displays many of the characteristic features of a soluble globular protein, including a stable folded structure, a chemically and topographically heterogeneous surface, and interactions with diverse molecules that are crucial to its function (Figure 1d). Many of these interactions are mediated by a well-documented hydrophobic patch,[55] which is also shown in Figure 1d.

To characterize the overall strength of protein-water interactions, once again, we apply a biasing potential, ϕN_v, where N_v is now

the number of coarse-grained waters in the entire protein hydration shell, v. The hydration shell, v, is defined as the union of spherical sub-volumes centered on all the protein heavy atoms, with each sub-volume chosen to have the same radius, R_v. Such a definition allows v to capture the ruggedness of the underlying protein surface, with the width of the hydration shell determined by R_v. Here we choose $R_v = 0.6$ nm so that only waters in the first hydration shell of the protein are included in v (Figure 1d). The decrease in the average number of ubiquitin hydration waters, $\langle N_v \rangle_\phi$, in response to the unfavorable potential, ϕN_v, is shown in Figure 1e. Interestingly, $\langle N_v \rangle_\phi$ displays a sigmoidal dependence on ϕ, akin to that for the hydrophobic CH_3-terminated SAM surface (Figure 1b). Correspondingly, a clear peak is also observed in the susceptibility around $\phi^* \approx 2\ k_B T$ (Figure 1f). Thus, the hydration shell of the inherently amphiphilic and incredibly complex surface of the ubiquitin protein dewets collectively in response to an unfavorable perturbation.

How the hydration shells of diverse proteins respond to unfavorable potentials

Is ubiquitin unique? If not, how general is the susceptibility of protein hydration waters to an unfavorable potential? To address this question, we studied six additional proteins, spanning a range of sizes, chemical patterns, and functional roles; see Figure 2 as well as Figure S3 in the Supporting Information. Given the importance of hydrophobic surface moieties in situating the interfacial waters at the edge of a dewetting transition, we initially considered other proteins with well-defined hydrophobic patches. First, we consider the fungal protein, Hydrophobin II, which is highly surface-active, and is known to self-assemble at water-vapor interfaces.[60] Although Hydrophobin II is charge neutral overall, the protein surface displays 10 charged residues. In Figure 2a, we show both the hydrophobic face of Hydrophobin II, which is enriched in hydrophobic residues, as well as the remainder of the protein, which has been shown to be super-hydrophilic.[20] As with ubiquitin, the susceptibility, χ_v, for Hydrophobin II also displays a marked peak. Next, we consider the human hepatitis B viral capsid protein, which has a net charge of -6, but displays an even larger hydrophobic patch than the one on Hydrophobin II (Figure 2b). That patch drives the binding of two capsid proteins to form a dimer which further assembles into a 240-protein capsid shell.[61] The viral capsid protein also displays a clear peak in χ_v. Does the collective dewetting seen in the above proteins stem from the presence of extended hydrophobic patches on their surfaces? Although many proteins possess such patches, not all of them do; instead, most protein surfaces display chemical patterns that are amphiphilic, and feature only smaller hydrophobic regions. The signaling protein MDM2 contains such a modest hydrophobic groove, which is nevertheless important from a functional standpoint; it enables MDM2 to exercise control over cellular senescence by binding to the transactivation domain of the tumor-suppresor protein p53.[62] As shown in Figure 2c, MDM2 also displays a peak in susceptibility, χ_v.

Might the large or small, but well-defined hydrophobic patches on ubiquitin, Hydrophobin II, the capsid sub-unit, and MDM2 be responsible for rendering their hydration shell waters susceptible to unfavorable perturbation? To address this question, we study proteins that are known for being anomalously hydrophilic or charged. Malate dehydrogenase is a large hydrophilic protein with 61 charged surface residues. The protein dimerizes into a metabolic enzyme, and plays an important role in the citric acid cycle. The interface through which the protein monomers bind is fairly hydrophilic, featuring 5 charged residues and several other polar residues. In fact, regions of the binding interface are so hydrophilic that they hold on their waters even in the bound state. In other words, the binding interface features structured waters that bridge the two interacting proteins; such bridging waters, which are observed in the crystal structure of the malate dehydrogenase dimer, are shown in Fig-

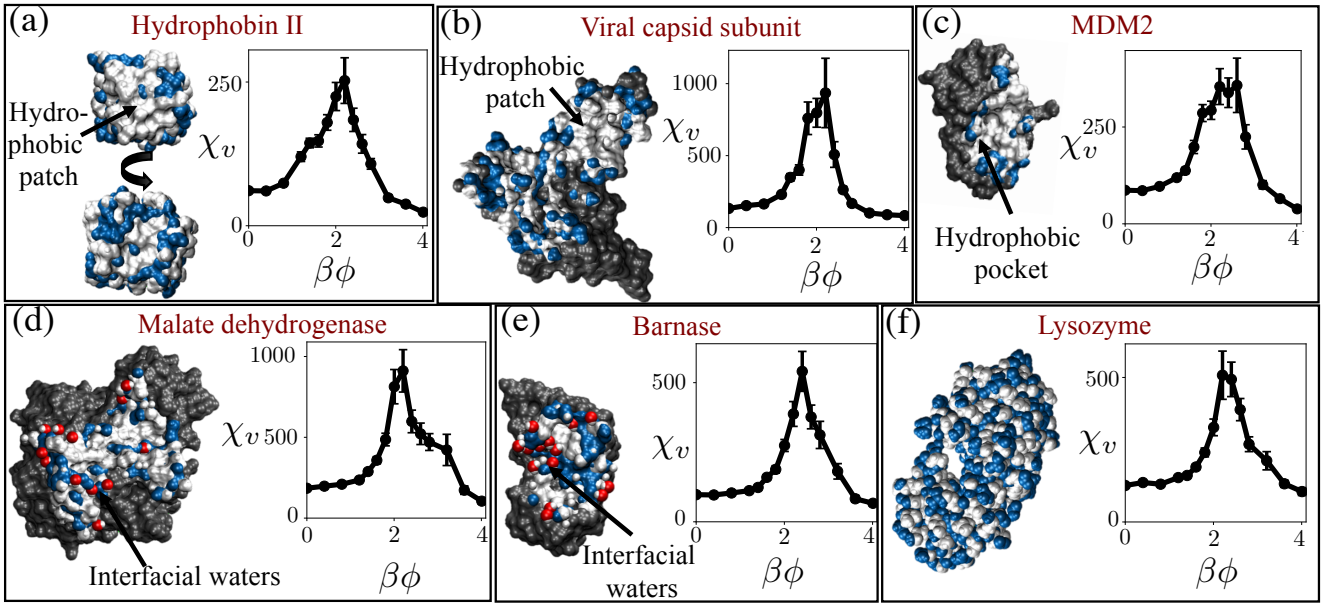

Figure 2: The hydration shells of proteins with diverse sizes, shapes, chemical patterns, and functional roles are susceptible to an unfavorable potential. For each protein, surface atoms of interest are shown in white (non-polar) or blue (polar or charged); the rest are shown in gray. Susceptibilities of the protein hydration waters, $\chi_v \equiv -\partial\langle N_v\rangle_\phi/\partial(\beta\phi)$, to the biasing potential strength, ϕ, are also shown. (a) Hydrophobin II is a small fungal protein that adsorbs to water-vapor interfaces via a large hydrophobic patch (top); the rest of the protein is amphiphilic (bottom).[58,59] (b) The protein that makes up a structural sub-unit of the hepatitis B viral capsid is shown, along with the amphiphilic interface through which two protein sub-units bind; the hydrophobic patch on the binding interface is also highlighted. (c) The signaling protein, MDM2, interacts with its binding partner, p53, through a smaller interaction interface, resembling a crevice that is lined with hydrophobic residues. (d) The enzyme malate dehydrogenase forms a homo-dimer through a binding interface, which is sufficiently hydrophilic that it retains waters in the bound state; those waters are shown in red/white. (e) The bacterial RNase barnase employs five charged residues to bind with its inhibitor barstar in one of the most stable complexes known; structural waters at the barnase-barstar binding interface are also shown. (f) Lysozyme is a soluble protein with a highly amphiphilic surface that resembles a checker pattern. For each of these diverse proteins, the susceptibility of waters in the entire protein hydration shell to the unfavorable potential, χ_v displays a marked peak. Moreover, the peak in χ_v occurs at roughly $\phi^* \approx 2\ k_{\mathrm{B}}T$ for each protein.

ure 2d.[63,64] In contrast, the binding interfaces of the proteins discussed previously are entirely dry. Interestingly, even for the largely hydrophilic malate dehydrogenase protein, we observe a peak in susceptibility, χ_v, in response to an unfavorable perturbation (Figure 2d).

Another fairly hydrophilic protein that features a charged interaction interface is barnase, a bacterial RNase that interacts with its inhibitor, barstar, in one of the strongest known protein-protein interactions.[65] The high-affinity sub-picomolar binding between barnase and barstar is facilitated by the formation of electrostatic contacts between five positively charged residues on barnase and five negatively charged residues on barstar.[65] Remarkably, a clear peak in susceptibility is also observed for barnase (Figure 2e). Finally, we study T4 lysozyme, a bacteriophage protein, which catalyzes the hydrolysis of the peptidoglycan layer of bacterial cell walls,[66] and has 45 charged surface residues with an overall charge of +9. Lysozyme does not appear to participate in interactions with proteins other than its substrates, or to possess a clear hydrophobic patch; rather, it displays a checkered pattern of hydrophobic and

hydrophilic atoms (Figure 2f). As with all of the proteins studied here, the hydration shell of lysozyme also displays a marked peak in susceptibility. Our results, obtained across proteins with a diversity of sizes, biological functions, and surface chemistries, thus suggest that susceptibility to an unfavorable potential is a general property of protein hydration waters.

Characterizing protein surfaces: residues vs atoms

Although collective dewetting in response to an unfavorable potential may be expected for proteins that are fairly hydrophobic, it is somewhat surprising that even the more hydrophilic and charged proteins display such behavior. To better characterize the similarities and differences in the surface chemistries of the seven proteins discussed above, we plot the fraction of their surface residues that are charged and hydrophobic in Figures 3a and 3b, respectively. As expected, the more hydrophobic proteins have a smaller fraction of charged residues and a larger fraction of hydrophobic residues, with the charge fraction ranging from 0.15 to 0.35, and the hydrophobic fraction varying from 0.5 to 0.2. To interrogate whether the surface chemistries of these seven proteins are representative of the larger class of folded, globular proteins, we additionally estimated these quantities for an expanded set containing a total of 20 proteins. The results are included in Figure S4 of the Supporting Information, and highlight that the characteristics of proteins studied here are indeed representative of typical proteins. How then do we understand the sensitivity of such diverse protein hydration shells to unfavorable potentials, and their resemblance to extended hydrophobic surfaces?

To answer this question, we draw inspiration from work by Kapcha and Rossky, who highlighted that amino acid residues are not monolithic, but are instead heterogeneous, and are composed of both hydrophobic and hydrophilic atoms.[54] Kapcha and Rossky thus advocate adopting an atom-centric, rather than a residue-centric view of the protein surface. They further suggest that an atom be classified as hydrophobic only if the magnitude of its partial charge is less than 0.25 in the OPLS force field, and hydrophilic otherwise.[54] Following these authors, we plot the fraction of surface atoms (not residues) that are hydrophobic in Figure 3c. Interestingly, we find that the fraction of surface atoms that are hydrophobic is not only larger than the corresponding fraction of surface residues, but that roughly half the protein surface (or more) consists of hydrophobic atoms. Importantly, the fraction of surface atoms that are hydrophobic is uniformly high for all proteins studied here.

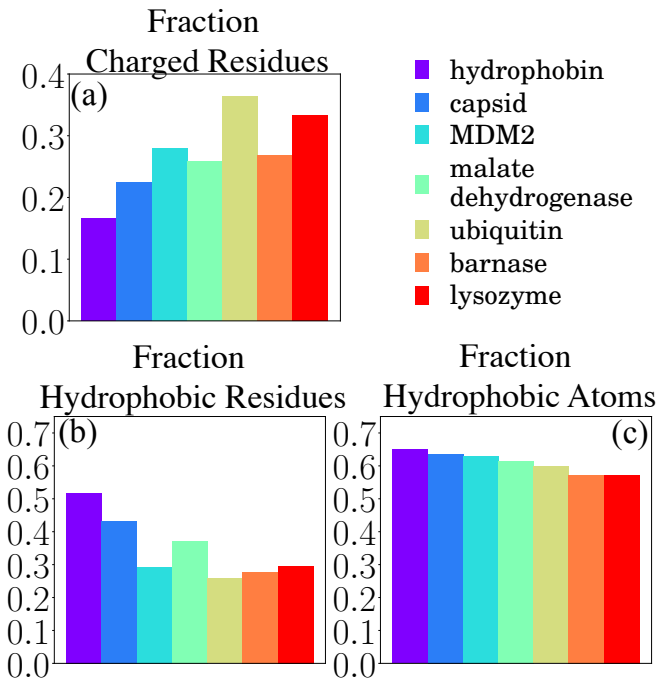

Figure 3: An atom-centric view of diverse protein surfaces reveals that they are more hydrophobic than anticipated by the corresponding residue-centric view. (a) Fraction of protein surface residues that are charged. (b) Fraction of surface residues that are hydrophobic. (c) Fraction of protein surface atoms that are hydrophobic according to the Kapcha-Rossky classification.[54]

To better understand these results, we analyzed the atomic composition of the surface residues; as shown in Figure S7 of the Supporting Information, roughly 80% of atoms belonging to hydrophobic residues are hydrophobic, but nearly 50% of atoms belonging to hydrophilic (polar or charged) residues are

also hydrophobic. Thus, although polar and charged surface residues do not contribute as heavily to hydrophobic surface atoms, they nevertheless have substantive contributions. We note that following Kapcha and Rossky, we classify not just the protein heavy atoms, but also hydrogen atoms as being either hydrophobic or hydrophilic.[54] If the protein hydrogen atoms are excluded from the analysis, the fraction of hydrophobic surface atoms is somewhat lower, but remains close to 0.5 as shown in Figure S5 of the Supporting Information. We also vary the Kapcha-Rossky[54] partial charge threshold of 0.25 systematically from 0.15 to 0.35, and find that the fraction of hydrophobic surface atoms is qualitatively similar to that in Figure 3c; see Figure S6 of the Supporting Information. Our results thus suggest that for a wide variety of proteins, roughly half the surface consists of hydrophobic atoms; these atoms situate the protein hydration waters at the edge of a dewetting transition, making them particularly susceptible to unfavorable potentials.

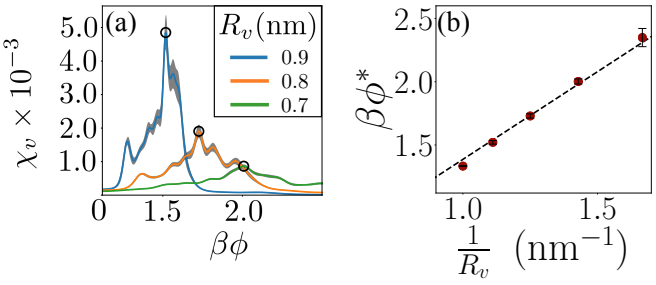

Figure 4: How the location of the peak in susceptibility depends on the width of the protein hydration shell, v. (a) Susceptibility of the ubiquitin hydration waters to the biasing potential strength, ϕ, is shown for hydration shells with different widths; as R_v is increased, the peak in χ_v shifts to lower ϕ-values. (b) The peak location, $\beta\phi^*$ as a function of R_v^{-1} (symbols), and a linear fit to the data (line) are shown.

Strength of unfavorable potential needed to trigger dewetting

Interestingly, not only do all the proteins studied here show a peak in χ_v, the location of this characteristic peak in susceptibility is roughly the same, $\phi^* \approx 2\ k_BT$, in all cases (Figure 2 and Figure S2a of Supporting Information). We also note that although there is substantial variation in χ_v at ϕ^* across different proteins (Figure 2 and Figure S2b of Supporting Information), much of the variation can be explained by differences in protein sizes, and vanishes upon normalizing χ_v by the unbiased average number of waters, $\langle N_v \rangle_0$ (Figure S2d of Supporting Information). To understand these observation, we first recognize that for $\phi > 0$, the biasing potential, ϕN_v, favors configurations with lower N_v-values and thereby lower densities. As a result, the biasing potential effectively lowers the pressure in the protein hydration shell to: $P_{\text{eff}} \approx P - \phi\rho_w$, where P is the system pressure and ρ_w is the molar density of liquid water. For a sufficiently large ϕ, the tension (negative pressure) exerted on the protein hydration waters prompts the nucleation of vapor in certain regions of v. The biasing potential thus stabilizes a water-vapor interface; the pressure drop across this interface is related to the corresponding interfacial tension, γ, and the mean interfacial curvature, $\bar{\kappa}$, according to the Young-Laplace equation, $P - P_{\text{eff}} = \gamma\bar{\kappa}$. Moreover, because the biasing potential is only experienced by waters in v, the interfacial curvature is determined by the radius, R_v, of the spherical sub-volumes defining v, such that $\bar{\kappa} \propto 1/R_v$. Thus, the biasing potential strength needed to trigger vapor nucleation ought to be: $\phi^* \approx (P - P_{\text{eff}})/\rho_w \propto (\gamma/\rho_w)(1/R_v)$. By systematically varying R_v, and repeating our calculations for ubiquitin, we find that as R_v is increased, the peak in susceptibility indeed shifts to lower values (Figure 4a). Furthermore, as shown in Figure 4b, ϕ^* also varies linearly with $1/R_v$ as predicted; the best fit line through the origin has a slope of 1.38 nm, which is comparable to $\beta\gamma/\rho_w = 0.45$ nm, estimated using $\rho_w = 33$ nm^{-3} and $\gamma = 60.2$ mJ/nm^2 for SPC/E water.

The above discussion also suggests a way to experimentally perturb protein hydration waters in a manner that closely mimics the potential employed in our simulations – subjecting a protein solution to tension. However, the potential in our simulations is only felt by waters in v, whereas the tension in the experiments would be experienced by all waters; thus, care would have to be exercised to prevent the cavitation of bulk water. Because both the unfavorable potential in our simulations and the tension in the experiments proposed above alter the chemical potential of the vapor relative to the liquid, an alternative way to dewet the protein hydration shell could be through the introduction a small hydrophobic solute into the protein solution (e.g., methane). Systematically varying the solute chemical potential then ought to result in the nucleation of solute regions in the protein hydration shell.

Conclusions and Outlook

Proteins employ intricate topographical and chemical patterns, which have evolved to facilitate their many biological functions. Although the space of such patterns is immense, it is likely constrained by common characteristics that all proteins must possess in order to function properly. For example, all proteins must have favorable interactions with water to be soluble, which requires the presence of hydrophilic groups on their surfaces. Conversely, protein surfaces must also feature hydrophobic regions that interact poorly with water, and provide a driving force for proteins to interact with other molecules. In this article, we shed light on how proteins accomplish these competing goals by balancing their overall interactions with their hydration waters. We show that roughly half the atoms on the protein surface are hydrophobic – a fact that can be obfuscated by focusing on surface residues rather than surface atoms. We also find that the hydration shells of diverse proteins – even those with highly charged surfaces and amphiphilic interaction interfaces – are highly susceptible to an unfavorable potential. Our results thus suggest that hydrophobic atoms on the protein surface situate its hydration waters at the edge of a dewetting transition, which can be triggered by an unfavorable perturbation.

The susceptibility of protein hydration waters to an unfavorable biasing potential, seen here across diverse proteins, arises from fat low-density tails in the underlying distribution of protein hydration waters, i.e., from enhanced low-density fluctuations.[27] A small change to the protein surface can modulate the fatness of the tails[46] and thereby the strength of the potential, ϕ^*, required to trigger dewetting in subtle and non-trivial ways; however, such small changes can nevertheless lead to substantial changes in the free energy required to dewet the protein and form interfacial cavities. Protein hydrophobicity, which is quantified by the ease of interfacial cavity formation,[19,20] can thus vary substantially with only small changes in the chemical or topographical patterns presented by the protein surface. Importantly, such a sensitive dependence of hydrophobicity on the protein context is underpinned by a susceptibility to dewetting; given our finding that the hydration shells of diverse proteins undergo collective dewetting, we expect context-dependent hydrophobicity to also be a general characteristic of proteins.[46]

Indeed, manifestations of such context-dependent hydrophobicity have been observed across a number of systems and interfacial properties;[20,36–48] our findings are consistent with these observations, and suggest that they stem from the proximity of hydration waters to collective dewetting transitions. Signatures of collective transitions have also been observed in studies of partially hydrated proteins. For example, Cui *et al.* found that partially hydrated proteins undergo a percolation transition at a critical value of protein hydration.[67] Similarly, in studying the uptake of water from the vapor phase by proteins, Debenedetti and co-workers found protein hydration to display hysteresis – a hallmark of collective transitions – between the adsorption and desorption branches of the isotherm.[16,68] These authors also found that polar and charged residues contributed to the collective wetting of a dry protein (and the as-

sociated hysteresis);[69] correspondingly, here we find that non-polar regions of the protein give rise to the collective dewetting of a hydrated protein.

We also find that the biasing potential strength needed to trigger dewetting is inversely proportional to the width of the hydration shell, but does not depend substantially on the particular protein being perturbed. Thus, our results not only suggest that susceptibility to an unfavorable perturbation is a general feature of the hydration shells of proteins, but also highlight that the strength of the perturbation needed to trigger dewetting is rather similar across different proteins. This finding suggests a near-universal calibration of the perturbation strength across diverse proteins, i.e., by considering ϕ relative to ϕ^*. It also establishes a framework for systematically classifying how favorable the interactions between water and different parts of the protein surface are. In particular, we expect that locations on the protein surface that dewet at low ϕ-values (relative to ϕ^*) will correspond to the most hydrophobic regions on the protein surface, whereas regions that retain their waters even at high ϕ-values will be the most hydrophilic. Given the importance of hydrophobicity in driving protein interactions, regions of the protein surface that dewet most readily may correspond closely with patches on the protein that participate in interactions.[6,49,52,70] We are investigating whether such hydrophobic protein regions could serve as predictors of protein interaction sites, and plan to report our findings in a future study. Similarly, identification of the most hydrophilic regions of the protein could facilitate the discovery of novel super-hydrophilic chemical patterns; an understanding of what enables such patterns to have strong interactions with water could also serve as the basis for the rational design of protein non-fouling surfaces or surfaces that display super-oleophobicity underwater.[53,71]

Acknowledgement A.J.P. gratefully acknowledges financial support from the National Science Foundation (CBET 1652646, CHE 1665339, and UPENN MRSEC DMR 1720530), and a fellowship from the Alfred P. Sloan Research Foundation (FG-2017-9406). N.B.R. was supported by the National Science Foundation grant CBET 1652646. The authors thank Shekhar Garde and Baron Peters for insightful discussions.

Supporting Information Available

In the Supporting Information, we include details of our simulations, information on enhanced sampling techniques that we use, plots supporting Figure 2, and additional analysis pertaining to the diversity of the proteins studied here.

References

(1) Ashbaugh, H. S.; Hatch, H. W. *J. Am. Chem. Soc.* **2008**, *130*, 9536–9542.

(2) Krone, M. G.; Hua, L.; Soto, P.; Zhou, R.; Berne, B. J.; Shea, J.-E. *J. Am. Chem. Soc.* **2008**, *130*, 11066–11072.

(3) Levy, Y.; Onuchic, J. N. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 389–415.

(4) Berne, B. J.; Weeks, J. D.; Zhou, R. *Annu. Rev. Phys. Chem.* **2009**, *60*, 85–103.

(5) Jamadagni, S. N.; Godawat, R.; Garde, S. *Annu. Rev. Chem. Biomol. Eng.* **2011**, *2*, 147–171.

(6) Vashisth, H.; Abrams, C. F. *Proteins: Struct., Funct., Bioinf.* **2013**, *81*, 1017–1030.

(7) Baron, R.; McCammon, J. A. *Annu. Rev. Phys. Chem.* **2013**, *64*, 151–75.

(8) Tiwary, P.; Limongelli, V.; Salvalaglio, M.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, E386–E391.

(9) Tiwary, P.; Mondal, J.; Morrone, J. A.; Berne, B. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 12015–12019.

(10) Bellissent-Funel, M.-C.; Hassanali, A.; Havenith, M.; Henchman, R.; Pohl, P.; Sterpone, F.; van der Spoel, D.; Xu, Y.; Garcia, A. E. *Chem. Rev.* **2016**, *116*, 7673–7697.

(11) Jaenicke, R. *J. Biotechnol.* **2000**, *79*, 193–203.

(12) Liu, S. J. S.; Shahrokh, Z.; Jun, *J. Pharm. Sci.* **2004**, *93*, 1390–1402.

(13) Shen, V. K.; Cheung, J. K.; Errington, J. R.; Truskett, T. M. *Biophys. J.* **2006**, *90*, 1949–60.

(14) Trevino, S. R.; Scholtz, J. M.; Pace, C. N. *J. Pharm. Sci.* **2008**,

(15) Thirumalai, D.; Reddy, G.; Straub, J. E. *Acc. Chem. Res.* **2012**, *45*, 83–92.

(16) Debenedetti, J. C. P.; Pablo, *J. Phys. Chem. Lett.* **2012**, *3*, 2713–2718.

(17) Chong, S.-H.; Ham, S. *Angew. Chem., Int. Ed.* **2014**, *53*, 3751–3751.

(18) Remsing, R. C.; Xi, E.; Patel, A. J. *J. Phys. Chem. B* **2018**, *122*, 3635–3646.

(19) Patel, A. J.; Varilly, P.; Jamadagni, S. N.; Acharya, H.; Garde, S.; Chandler, D. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 17678–17683.

(20) Patel, A. J.; Garde, S. *J. Phys. Chem. B* **2014**, *118*, 1564–1573.

(21) Xi, E.; Remsing, R. C.; Patel, A. J. *J. Chem. Theory Comput.* **2016**, *12*, 706–713.

(22) Lum, K.; Chandler, D.; Weeks, J. D. *J. Phys. Chem. B* **1999**, *103*, 4570–4577.

(23) Varilly, P.; Patel, A. J.; Chandler, D. *J. Chem. Phys.* **2011**, *134*, 074109.

(24) Vaikuntanathan, S.; Rotskoff, G.; Hudson, A.; Geissler, P. L. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, E2224–E2230.

(25) Xi, E.; Patel, A. J. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, 4549–4551.

(26) Patel, A. J.; Chandler, D. *J. Phys. Chem. B* **2010**, *114*, 1632–1637.

(27) Patel, A. J.; Varilly, P.; Jamadagni, S. N.; Hagan, M. F.; Chandler, D.; Garde, S. *J. Phys. Chem. B* **2012**, *116*, 2498–2503.

(28) Rossky, C. Y. L.; McCammon, J. A.; J., P. *J. Chem. Phys.* **1984**, *80*, 4448–4455.

(29) Mittal, J.; Hummer, G. *Proc. Nat. Acad. Sci.* **2008**, *105*, 20130–20135.

(30) Sarupria, S.; Garde, S. *Phys. Rev. Lett.* **2009**, *103*, 037803.

(31) Godawat, R.; Jamadagni, S. N.; Garde, S. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 15119–15124.

(32) Willard, A. P.; Chandler, D. *J. Phys. Chem. B* **2010**, *114*, 1954–1958.

(33) Heyden, M.; Tobias, D. J. *Phys. Rev. Lett.* **2013**, *111*, 218101.

(34) Sosso, G. C.; Caravati, S.; Rotskoff, G.; Vaikuntanathan, S.; Hassanali, A. *J. Phys. Chem. A* **2016**, *121*, 370–380.

(35) Shin, S.; Willard, A. P. *J. Phys. Chem. B* **2018**,

(36) Zhou, R.; Huang, X.; Margulis, C. J.; Berne, B. J. *Science* **2004**, *305*, 1605–1609.

(37) Liu, P.; Huang, X.; Zhou, R.; Berne, B. J. *Nature* **2005**, *437*, 159–162.

(38) Giovambattista, N.; Lopez, C. F.; Rossky, P. J.; Debenedetti, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 2274–2279.

(39) Giovambattista, N.; Debenedetti, P. G.; Rossky, P. J. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 15181–15185.

(40) Acharya, H.; Vembanur, S.; Jamadagni, S. N.; Garde, S. *Faraday Discuss.* **2010**, *146*, 353–365.

(41) Daub, C. D.; Wang, J.; Kudesia, S.; Bratko, D.; Luzar, A. *Faraday Discuss.* **2010**, *146*, 67–77.

(42) Mittal, J.; Hummer, G. *Faraday Discuss.* **2010**, *146*, 341–352.

(43) Wang, J.; Bratko, D.; Luzar, A. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 6374–6379.

(44) Fogarty, A. C.; Laage, D. *J. Phys. Chem. B* **2014**, *118*, 7715–7729.

(45) Harris, R. C.; Pettitt, B. M. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 14681–14686.

(46) Xi, E.; Venkateshwaran, V.; Li, L.; Rego, N.; Patel, A. J.; Garde, S. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, 201700092.

(47) Monroe, J. I.; Shell, M. S. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, 8093–8098.

(48) Heyden, M. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* e1390.

(49) Bogan, A. A.; Thorn, K. S. *J. Mol. Biol.* **1998**, *280*, 1–9.

(50) DeLano, W. L. *Curr. Opin. Struct. Biol.* **2002**, *12*, 14–20.

(51) Nooren, I. M.; Thornton, J. M. *EMBO J.* **2003**, *22*, 3486–3492.

(52) White, A. W.; Westwell, A. D.; Brahemi, G. *Expert Rev. Mol. Med.* **2008**, *10*, e8.

(53) Chen, S.; Cao, Z.; Jiang, S. *Biomaterials* **2009**, *30*, 5892–5896.

(54) Kapcha, L. H.; Rossky, P. J. *J. Mol. Biol.* **2014**, *426*, 484–498.

(55) Winget, J. M.; Mayor, T. *Mol. Cell* **2010**, *38*, 627–635.

(56) Patel, A. J.; Varilly, P.; Chandler, D.; Garde, S. *J. Stat. Phys.* **2011**, *145*, 265–275.

(57) Hershko, A.; Ciechanover, A. *Annu. Rev. Biochem.* **1982**, *51*, 335–364.

(58) Hakanpää, J.; Linder, M.; Popov, A.; Schmidt, A.; Rouvinen, J. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2006**, *62*, 356–367.

(59) Wösten, H. A. B.; de Vocht, M. L. *Biochim. Biophys. Acta, Rev. Biomembr.* **2000**, *1469*, 79–86.

(60) Hakanpaa, J.; Linder, M.; Popov, A.; Schmidt, A.; Rouvinen, J. *Acta Crystallogr., Sect D: Biol. Crystallogr.* **2006**, *62*, 356–367.

(61) Wynne, S. A.; Crowther, R. A.; Leslie, A. G. *Mol. Cell* **1999**, *3*, 771–780.

(62) Kussie, P. H.; Gorina, S.; Marechal, V.; Elenbaas, B.; Moreau, J.; Levine, A. J.; Pavletich, N. P. *Science* **1996**, *274*, 948–953.

(63) Rodier, F.; Bahadur, R. P.; Chakrabarti, P.; Janin, J. *Proteins: Struct., Funct., Bioinf.* **2005**, *60*, 36–45.

(64) Zaitseva, J.; Meneely, K. M.; Lamb, A. L. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2009**, *65*, 866–869.

(65) Schreiber, G.; Fersht, A. R. *J. Mol. Biol.* **1995**, *248*, 478–486.

(66) Shoichet, B. K.; Baase, W. A.; Kuroki, R.; Matthews, B. W. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 452–456.

(67) Cui, D.; Ou, S.; Patel, S. *Proteins: Struct., Funct., Bioinf.* **2014**, *82*, 3312–3326.

(68) Kim, S. B.; Palmer, J. C.; Debenedetti, P. G. *J. Phys. Chem. B* **2015**, *119*, 1847–1856.

(69) Kim, S. B.; Sparano, E. M.; Singh, R. S.; Debenedetti, P. G. *J. Phys. Chem. Lett.* **2017**, *8*, 1185–1190.

(70) Keskin, O.; Ma, B.; Nussinov, R. *J. Mol. Biol.* **2005**, *345*, 1281–1294.

(71) Si, Y.; Dong, Z.; Jiang, L. *ACS Cent. Sci.* **2018**,

Graphical TOC Entry