Biophysical Journal
Article

Biophysical Society

Molecular Effects of Concentrated Solutes on Protein Hydration, Dynamics, and Electrostatics

Luciano A. Abriata,[1,*] Enrico Spiga,[1] and Matteo Dal Peraro[1,*]
[1]Institute of Bioengineering, School of Life Sciences, École Polytechnique Fédérale de Lausanne and Swiss Institute of Bioinformatics, Lausanne, Switzerland

ABSTRACT Most studies of protein structure and function are performed in dilute conditions, but proteins typically experience high solute concentrations in their physiological scenarios and biotechnological applications. High solute concentrations have well-known effects on coarse protein traits like stability, diffusion, and shape, but likely also perturb other traits through finer effects pertinent at the residue and atomic levels. Here, NMR and molecular dynamics investigations on ubiquitin disclose variable interactions with concentrated solutes that lead to localized perturbations of the protein's surface, hydration, electrostatics, and dynamics, all dependent on solute size and chemical properties. Most strikingly, small polar uncharged molecules are sticky on the protein surface, whereas charged small molecules are not, but the latter still perturb the internal protein electrostatics as they diffuse nearby. Meanwhile, interactions with macromolecular crowders are favored mainly through hydrophobic, but not through polar, surface patches. All the tested small solutes strongly slow down water exchange at the protein surface, whereas macromolecular crowders do not exert such strong perturbation. Finally, molecular dynamics simulations predict that unspecific interactions slow down microsecond- to millisecond-timescale protein dynamics despite having only mild effects on pico- to nanosecond fluctuations as corroborated by NMR. We discuss our results in the light of recent advances in understanding proteins inside living cells, focusing on the physical chemistry of quinary structure and cellular organization, and we reinforce the idea that proteins should be studied in native-like media to achieve a faithful description of their function.

INTRODUCTION

Proteins typically experience complex media characterized by high solute concentrations in living organisms and biotechnological applications, but these conditions are largely ignored in most studies about structure, folding, dynamics, and function (1–3). In living systems, proteins are exposed to other proteins and macromolecules of varied kinds at total concentrations ranging from ~50 to ~300 g/L, depending on cell type, cellular and subcellular localization, and timing in the cell cycle (3–8). Small molecules from the central metabolic pathways contribute with an additional ~50–100 g/L of solutes (9); and many small-molecule osmolytes accumulate at even higher concentrations in vivo under stress conditions (10–13). Likewise, in biotechnological applications, proteins are stored in highly concentrated solutions of polyols/carbohydrates, refolded at high concentrations of small molecules like arginine after solubilization from inclusion bodies with chaotropic agents, separated through aqueous biphasic systems with high concentrations of polymers like polyethylene glycols (PEGs) or dextran, and even used as catalysts in media crowded with other proteins, as exemplified by many molecular biology enzymes which are used in high concentrations of bovine serum albumin (BSA) (14–18). The effects that make certain solutes useful in protein biotechnology at high concentrations, and the ultimate biological consequences of proteins working under high solute concentrations in vivo, are the subject of intensive research. The biological scenario is particularly challenging given the complex and heterogeneous composition of the cell interior, which is patent in the several contrasting reports of protein traits strongly affected both in vitro and in vivo (19–23), in those reports where kinetics are altered with minor effects on thermodynamics, those where kinetic and thermodynamic effects are only attenuated or even absent in vivo compared to in vitro, and those where effects are found to vary across cells and cell compartments (24–32). Clearly, dissecting the impact of concentrated solutes of various sizes and chemical natures on proteins is critical to better understand them under

Submitted March 11, 2016, and accepted for publication July 5, 2016.

*Correspondence: luciano.abriata@epfl.ch or matteo.dalperaro@epfl.ch

Enrico Spiga's present address is Laboratory for Computational Cell and Biology, The Francis Crick Institute, Mill Hill Laboratory, Midland Road, London NW1 1AT, United Kingdom.

Editor: Rohit Pappu.

http://dx.doi.org/10.1016/j.bpj.2016.07.011

relevant conditions. For this, bottom-up approaches based on experiments with single cosolutes and theoretical modeling are needed to complement and facilitate the understanding of in vivo observations (29).

Studies about the effects of high solute concentrations on proteins have historically focused on "coarse" protein traits like overall diffusion, compaction, and stability, and interpreted in the context of increased viscosity, preferential interaction or exclusion (11), and excluded volumes for macromolecular crowders (33). However, more recent investigations and theoretical studies have highlighted the role of direct attractive and repulsive interactions, to an extent that attractive interactions could even overcome the effects of excluded volumes (29,34–37). Also, recent evidence suggests that protein traits relevant at the residue and atomic levels, such as hydration, internal dynamics, folding pathways, and even catalytic activity, can be affected by concentrated conditions (22,38–43). One important problem within the bigger picture is clarifying to what extent the most abundant kinds of molecules of the cytoplasm and those used in protein biotechnology interact with proteins at high concentrations, and how the multiple effects of concentrated solutes work at the atomic level. Early thermodynamic studies showed that most kinds of small molecules compatible with proteins (i.e., amino acids, small polyols and sugars, and "inert" macromolecules like PEGs and polysaccharides, but not strong denaturants like urea) overall prefer to be excluded from interactions with proteins, either by favoring interactions with water or by steric exclusion (44–48). However, these works also pointed out that preferential exclusion does not preclude actual unspecific interactions at high concentrations, whose precise nature would depend on the physicochemical signatures of the solutes and protein surfaces.

Here, we have used NMR experiments complemented by atomistic molecular dynamics (MD) simulations to map at atomic resolution the interactions between solutes of different sizes and chemical natures on ubiquitin, used as a test protein, and to determine their primary effects on protein hydration, conformational flexibility, and electrostatics. By "primary effects," we mean those effects concerning traits applicable to the folded protein as a whole and that can be extrapolated at least to other soluble globular proteins like ubiquitin; we leave aside more complex observables like folding kinetics, overall stability, or binding affinity for ubiquitin's natural partners, which are expected to arise as consequences of the primary effects and might be more dependent on each particular protein.

MATERIALS AND METHODS

Protein production and handling

^{15}N-labeled ubiquitin was expressed from a plasmid kindly provided by Antonio Donaire (University of Murcia, Murcia, Spain) in BL21(DE3) cells induced with 1 mM isopropyl β-D-1-thiogalactopyranoside at OD_{600} = 0.8 for 4 h at 37°C. A cell lysate was produced by freezing, thawing, and sonication of the harvested cells in 1 mM phenylmethylsulfonyl fluoride and 10 mg/mL lysozyme. This lysate was cleared by precipitation with perchloric acid in ice under agitation, followed by 30 min centrifugation at 20,000 × g. The supernatant was then loaded into a cation exchange column equilibrated in ammonium acetate (20 mM, pH 4.5), and ubiquitin was eluted with a gradient to 1 M sodium chloride in the same buffer.

NMR experiments

Protein samples for NMR were 0.2 mM in concentration, prepared in 50 mM sodium phosphate at pH 7 with 10% D_2O. Titrations were performed from concentrated stock solutions of the solutes for low-concentration points, and then by direct addition of the solid solutes to 2- to 3-mL samples of protein. Slight adjustments to pH 7, made with a microelectrode, were required in some cases.

NMR spectra were acquired in a Bruker 600 MHz spectrometer equipped with a cryoprobe and an Avance III console. Spectra were acquired and processed with Bruker's TopSpin 3 program and analyzed with Sparky and custom Matlab scripts. NMR assignments were adjusted from those in Biological Magnetic Resonance Bank entry 5387 (49). Details about acquisition and processing parameters for all NMR spectra are given in the Supporting Material.

We note that the raw chemical shift changes observed along the titrations include a residue-specific effect of the solute at a given concentration mounted on top of a residue-independent, concentration-dependent offset. The residue-independent offset builds up as the properties of the solution change (which affects internal instrument calibration, as it is based on the water resonance) and includes many contributions that are hard to resolve, for example, from proton exchange between water and cosolute molecules and from changes in the dielectric properties of the solution. It is expected that at least part of this shift can be accounted for by calibrating the spectra to a signal from a reference compound that remains presumably unaffected by the global effects. Such an assumption is not necessarily valid for any reference compound, especially at the high solute concentrations used here, but aliphatic protons are in principle expected to be less affected than water-exchangeable protons. We hence recalibrated our spectra collected in buffer and in 300 g/L of each solute, assuming that the ^{1}H chemical shifts of the methyl protons of 2,2-dimethyl-2-silapentane-5-sulfonate (DSS, a commonly used water-soluble reference compound for NMR) are exactly at 0 ppm in both extreme conditions, and we indirectly recalibrated the ^{15}N axes accordingly. We have anyway opted to interpret the effects on crosspeak positions only within experiment at the last point of each titration, and we present our results considering both the water-referenced (i.e., raw) and the DSS-corrected chemical shifts. From the changes in ^{1}H and ^{15}N chemical shifts at 300 g/L relative to buffer conditions ($\Delta\delta_H$ and $\Delta\delta_N$) under each calibration scenario, we have computed the corresponding chemical shift perturbations (CSPs) (Figs. 1 and 2) as $CSP = \sqrt{\Delta\delta_H^2 + (\Delta\delta_N^2/25)}$.

MD simulations

For this work, we ran MD simulations on three systems that contained 300 g/L concentrations of either zwitterionic negative glutamate, zwitterionic positive arginine, or glucose, and three ubiquitin molecules inside a box of size ~100 Å × 100 Å × 100 Å. The systems in glucose and water also had Na^+ and Cl^- ions at 150 mM concentration each, whereas the systems with glutamate and arginine zwitterions contained Na^+ or Cl^- ions as required to achieve charge neutrality. Systems were built with the Packmol program (50) by randomly placing the solutes around the proteins at a minimal interatomic distance of 5 Å. Glucose parameters were obtained from the Glycam06 force field (51), and protein molecules were treated with the AMBER99SB-ILDN force

field. Parameterization of the charged zwitterionic glutamate and arginine was achieved with the AMBERTOOLS parameterization toolkit, with charges optimized according to the RESP methodology (52). The obtained parameters are similar to those reported and tested in an ad hoc work (53) and reproduce available interatomic radial distribution functions measured by neutron diffraction experiments in concentrated solutions (54–56). Each system was treated at a fully atomistic level including explicit TIP3P solvent (57), reaching 150,000–200,000 atoms. Simulations were run with NAnoscale MD (58) and extend to 600–800 ns. Comparisons with multiprotein systems in water without cosolutes correspond to data from our previous work (59). Further details about system equilibration and production and about trajectory analysis are given in the Supporting Material.

RESULTS

For our aims, NMR is advantageous as it takes place in solution state at room temperature, reports at the atomic/residue levels, and is based on ^{15}N resonances of the test protein which means that every other molecule is invisible and only the effects of the solutes on the test protein are observed. This is much like the methodology used in recent in vivo studies (26,27,60), but on a simpler matrix that improves spectral quality and facilitates interpretation of the results. MD in turn allows us to explain some of our NMR observations and propose other atomic-level effects that cannot be probed by NMR.

Effects of small and large polyols on ubiquitin

We started by testing glucose and glycerol as representatives of small polyols, which are abundant in biological tissues and are used in biotechnology to stabilize proteins against heat and cold denaturation. Stepwise addition of glucose up to 300 g/L (~1.7 M) to a 200 μM solution of ^{15}N-labeled ubiquitin followed by ^{1}H,^{15}N-heteronuclear single-quantum coherence (HSQC) NMR spectra revealed large CSPs for all N-H crosspeaks, both in the water-referenced and DSS-referenced data (Figs. 1, *A* and *B*, and S1). The measured CSPs begin to be significant (>0.01 ppm) above 14–35 g/L (~0.1–0.2 M) glucose and grow linearly with concentration. Although there is no clear pattern in the CSP plot against ubiquitin's amino acid sequence, the water-referenced and DSS-referenced CSPs display at the final titration point a weak positive correlation against the fraction of solvent exposure of the N-H moieties (Fig. 2). In other words, more exposed N-H groups are more perturbed than buried residues, consistent with the MD prediction of extensive interactions between glucose molecules and the protein surface (41,61). Surface mapping of the DSS-referenced CSPs further shows patches of preferential interaction (Fig. 3). Of importance for comparisons with other solutes later on, note that the CSPs at 300 g/L glucose are offset even for buried groups, and regardless of whether the raw or DSS-calibrated shifts are considered.

Stepwise additions of glycerol to a solution of ^{15}N ubiquitin revealed essentially the same effects on HSQC spectra as additions of glucose (Figs. 1, *A* and *B*, and 2). In contrast, the CSPs observed upon stepwise additions of 70 and 400 kDa polymers of glucose and fructose (Ficoll) were an order of magnitude smaller than in glucose or glycerol at similar mass/volume concentrations, remaining essentially at the noise level (Figs. 1 *C* and S1). This rules out strong effects from the small (~15%) drop in the dielectric constant (ε) of the solution in the titrations with the small molecules (as ε is similar for similar concentrations of glycerol, glucose, and Ficoll (Table S1)) and therefore suggests that direct interactions with the protein are the main cause of CSPs by the small polyols. In turn, the very low CSPs by the Ficoll polymers suggest they are much less capable of interacting with the protein than glycerol or glucose, despite their very similar chemical natures, and that any weak interactions with the polymers entail more random orientations that largely cancel out the net CSP, whereas the small polyols would have certain preferred binding locations and poses resulting in a net CSP >0.

Effects of small charged molecules on ubiquitin

We next tested as concentrated solutes three charged small molecules, glutamate, arginine, and glycine, which are interesting for many reasons. These three small zwitterions are known to stabilize proteins in solution, especially glycine which is a physiologically and technologically relevant protein-stabilizing osmolyte (14,17,62,63). Furthermore, glutamate is among the most abundant metabolites in *E. coli* cells (9), and arginine is a common additive for protein refolding and solubilization (14,64). More generally, they are representative of intracellular metabolites, which are typically of low molecular weight and high charge density (9).

Stepwise additions of arginine hydrochloride, glycine, or sodium glutamate up to 300 g/L to a 200 μM solution of ^{15}N-labeled ubiquitin resulted in large CSPs across the HSQC spectra as evidenced by both the water-referenced and DSS-calibrated data (Fig. 1 *A*, and see an example of titration in Fig. S1). Although this seems similar to what we observed for glucose and glycerol, there are two important differences that could contain relevant information about the state of the protein in these solutions. First, the CSPs computed from DSS-referenced shifts follow a weak positive trend against the exposure of the N-H group, much as for glucose and glycerol but with the important difference that they are not offset. Second, the CSPs based on water-referenced shifts display negative, instead of positive, trends against N-H exposure, contrasting with the plots for glucose and glycerol where the positive trend against N-H exposure is clear in both the water-referenced and DSS-referenced data.

The positive trends for DSS-referenced CSPs against N-H exposure would reflect the interactions forced at the protein surface by all solutes (stronger for glucose and glycerol, as they seem to propagate to the protein

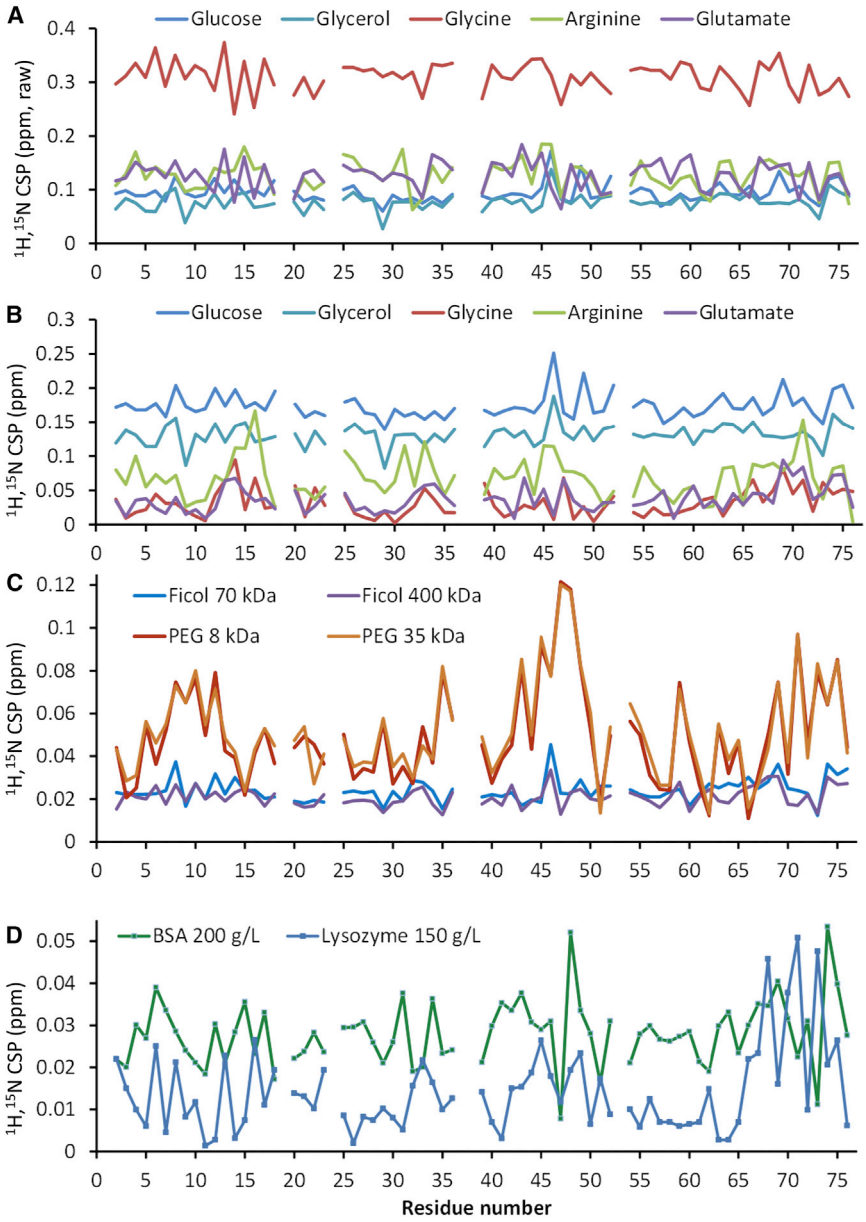

FIGURE 1 ^{1}H,^{15}N CSPs against ubiquitin's sequence at the final point of all titrations, which correspond to 300 g/L solute unless indicated otherwise. For the small molecules, CSPs computed from raw (*A*) and DSS-calibrated (*B*) shifts are shown; CSPs for the macromolecules (*C* and *D*) are from DSS-calibrated shifts (essentially identical to CSPs computed from raw shifts). Samples were studied in sodium phosphate, pH 7, at 300 K. To see this figure in color, go online.

interior). Meanwhile, the negative trends for water-referenced CSPs against N-H exposure observed for the amino acids but not for the polyols could reflect electrostatic effects, because concentrated solutions of amino acids have much (two- to threefold) higher dielectric constants than water, whereas concentrated solutions of small polyols have only slightly lowered dielectric constants (Table S1). We therefore explored this idea through MD simulations in 300 g/L sodium glutamate or arginine chloride in comparison to simulations in glucose and in pure water. These four simulations were 600–800 ns long, such that proteins and small molecules can amply diffuse in translational, rotational, and conformational spaces. The simulations reveal far fewer contacts between glutamate or arginine and the protein compared to glucose (Table S1). Moreover, they suggest a stronger preference for arginine and glutamate to get excluded from the protein surface, slightly increasing the density of the protein hydration layer (Fig. S2), which is consistent with trends observed in thermodynamic measurements (44–48). Because of their exclusion, the amino acid solutes diffuse quite freely and fast around the protein. Electrostatic calculations averaged throughout the simulated trajectories show that fast diffusion of the small zwitterions results in very ample fluctuations of the charges at the locations of the backbone protons, certainly much larger than the fluctuations observed in glucose which are similar to those in water (Fig. S3). The inverse dependence of electric field-induced chemical shifts on the effective dielectric constant at the nucleus (65,66) and the much lower dielectric constant inside the protein

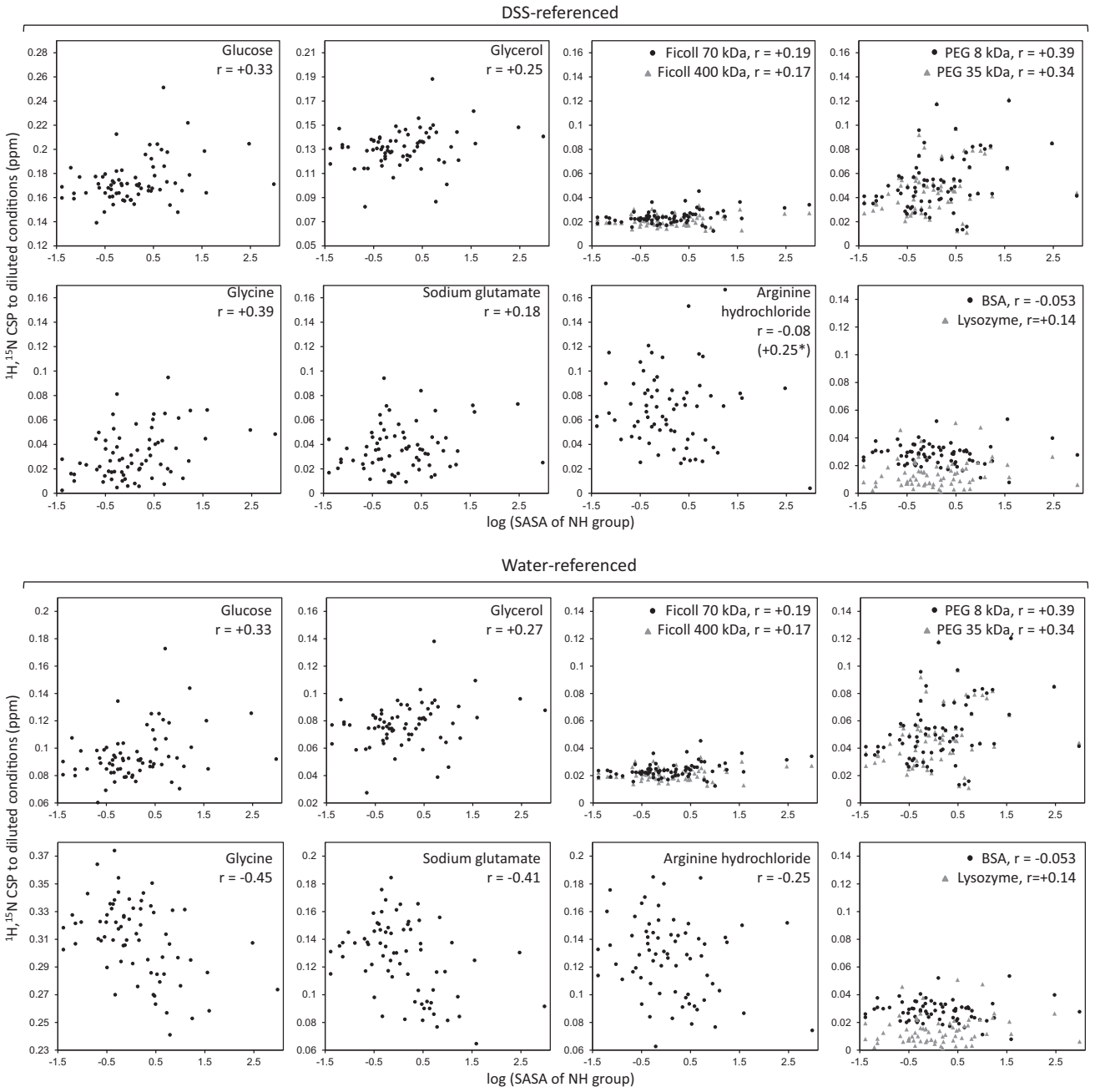

FIGURE 2 CSPs versus the logarithm of solvent-accessible surface areas for NH groups in ubiquitin, in 300 g/L solutions of all the studied solutes (unless indicated otherwise). CSPs computed from chemical shifts referenced to DSS (*top*) and from those referenced internally through the water resonance (*bottom*). All vertical axes span 0.17 ppm in DSS-referenced plots and 0.15 ppm in water-referenced plots, to facilitate comparison of the spread of values. In these plots, each dot corresponds to one ubiquitin residue for which NMR CSP data are available. All r values are Pearson correlation coefficients; they are significant at 0.01 and 0.05 levels when $|r| > 0.2775$ and 0.1983, respectively. In the top plot for arginine, the asterisk indicates the correlation coefficient when the point at the bottom right (2.98, 0.004) is excluded. Interactive versions of these plots with visible residue labels are available at http://lbm.epfl.ch/crowding2016.

than on its surface thus explain the trends observed for water-referenced CSPs with concentrated amino acids. Importantly, this finding implies that even if a small molecule does not interact strongly with the protein, if it is charged it can still induce long-range perturbations on the protein interior as it diffuses nearby. The potential consequences of this electrostatic coupling are especially interesting with regard to how metabolites flow in the cytoplasm, as we develop in the Discussion.

Effects of proteins and PEG on ubiquitin

Proteins are the major components by mass among intracellular solutes, and as such, they are especially prone to

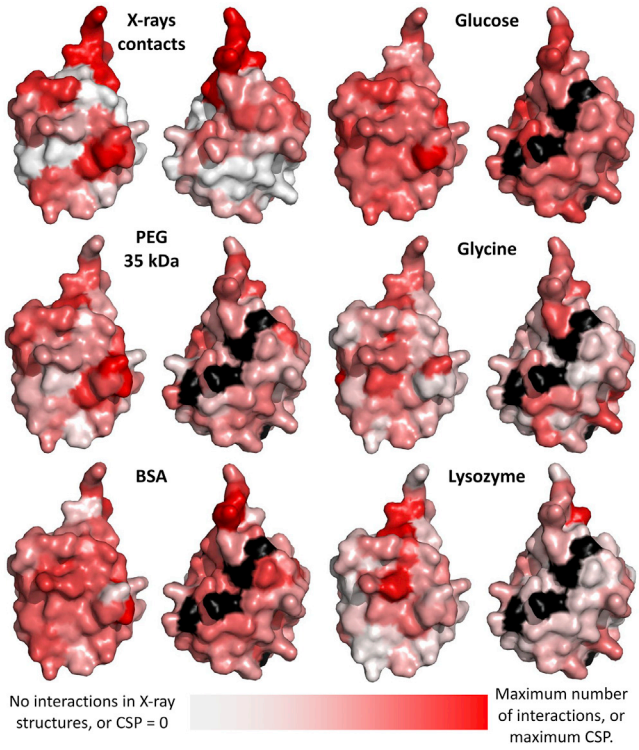

FIGURE 3 Ubiquitin's surface colored according to the number of interactions observed with other proteins in x-ray structures (*top left*) or according to the CSPs experienced under different solutes at high concentrations (all others). In each pair of structures, the righthand image is rotated by 180°. Surface patches colored black are residues for which no data are available. To see this figure in color, go online.

establish unspecific interactions in the cytoplasm, at least assuming fully random diffusion. Protein surfaces span a range of sizes, shapes, and exposed chemical signatures evolved for specific interactions with other cellular components. While these interactions are the focus of active investigations, other potential non-specific, presumably unintended interactions are rarely explored. We tested the effect on ^{15}N-ubiquitin of concentrated lysozyme and BSA, proteins that are in principle unrelated to ubiquitin and hence not expected to bind strongly to it.

With lysozyme at 150 g/L (larger concentrations were difficult to achieve and produced extreme signal broadening), there is a basal CSP of ~0.01 ppm with some residues reaching 0.02–0.05 ppm, especially next to the C-terminal tail (Fig. 1 *C*). With BSA at 200 g/L (also here, higher concentrations led to extreme broadening), there is a basal CSP of ~0.02 ppm, and residues that reach values of 0.03–0.05 ppm (Fig. 1 *C*). Although these CSP values are barely above the noise level, they are interesting because they map to hydrophobic and charged residues in surface patches that partially overlap with the surface that ubiquitin employs to interact with many of its physiological targets as seen in crystallographic complexes (Fig. 3). In titrations with lysozyme, ubiquitin residues with peak CSPs are Lys6, Leu8, Ile13, Glu16, Lys33, Phe45, Gln49, His68, Val70, Leu71, and Leu73. With BSA, peak CSPs are at residues Lys6, Leu15, and Val17, which surround Glu16, Gln31 (followed by Glu32 and Lys33), Glu34, Leu43 (preceded by Arg42), Lys48, Glu64, His68, and Arg74. This listing suggests that weak binding is forced mainly through hydrophobic and electrostatic interactions. We last tested the hypothesis that hydrophobic forces are the main drivers of unspecific protein-protein interactions by titrating ^{15}N ubiquitin with PEG polymers of average weights 8 and 35 kDa, which are highly soluble yet display low polarity and other hydrophobic features (67–69). In these experiments, the highest CSPs map quite well to the hydrophobic surface patches that ubiquitin employs to interact with other proteins, reaching at 300 g/L values close to 0.1 ppm around Ile44 and Phe45 which are at the core of the hydrophobic surface patch (Fig. 1 *B* and Fig. 3).

Protein-solute interactions and effects of solutes on protein rotational diffusion from ^{15}N relaxation

The product of ^{15}N longitudinal and transverse relaxation times, R_1R_2, can be used to monitor weak, unspecific interactions between a test protein and a macromolecular crowder (70) thanks to its independence on viscosity for sufficiently large correlation times (>6–7 ns at 600 MHz ^{1}H frequency, as used here) (71). On the other hand, ^{15}N relaxation data is best known as a tool to measure protein rotational diffusion and internal dynamics through R_2/R_1 ratios.

Although the effects of concentrated solutes on protein diffusion and on diffusion-dependent events like protein association have been vastly studied both in vitro and in vivo (72–79), effects on internal dynamics have been only barely explored, with the prediction, based on MD simulations, that they will be dampened at high solute concentrations and upon interaction with other molecules (41,59,80). We have measured ^{15}N relaxation data of ubiquitin's backbone amides in buffer and in concentrated glucose, glycine, Ficoll 70 kDa and 400 kDa, and PEG 8 kDa and 35 kDa, as an additional proxy for unspecific interactions (besides the CSP mapping) and to determine the effects of these solutes on rotational diffusion and internal dynamics.

The sequence-averaged R_1R_2 values are all similar within their uncertainties, reaching a maximum difference of just 2–4 Hz for PEG and Ficoll with similar uncertainties (Table S1). This perfectly supports the independence of R_1R_2 products on viscosity, but also shows that they are very weak reporters on interactions compared to the CSPs, at least for our system and conditions. On a more residue-specific basis, R_1R_2 profiles against the protein sequence (Fig. S4) show deviations close to residues with pico- to nanosecond flexibility, like those of the C-terminal tail, and to residues that experience resonance broadening, like Thr9, Glu24, Gly53, and the C-terminal tail, in all our concentrated solutions. This is not unexpected, because

the R_1R_2 product depends on the fourth power of the N-H order parameter and on locally high N-H exchange contributions to R_2 (71). This is especially clear for the PEG polymers, where strong mapping of the CSPs to specific regions of the protein sequence and surface indicates that R_1R_2's sensitivity to internal dynamics exceeds, and indeed blurs, its potential to detect interactions: First, CSP profiles with PEG show very clear patterns that are not evident in the R_1R_2 profiles. Second, regions close to residues whose NH groups are known to undergo microsecond- to millisecond-timescale exchange in buffer (Thr9, Glu24, and Gly53) display increased R_1R_2 and R_2/R_1 values in concentrated conditions that most likely arise from their intrinsic exchange processes rather than from weak binding. Last, residues of the C-terminal tail, very flexible in the picosecond to nanosecond timescale, show low R_1R_2 values that most likely reflect their fast ample motions also in concentrated media. Thus, it seems like most information about interactions in the R_1R_2 product is blurred by the stronger effects of dynamics, at least for ubiquitin with this set of concentrated solutes. Only for the smaller PEG polymer are there significant positive deviations of the R_1R_2 product (and R_2/R_1 ratios) matching regions where CSP mapping indicates binding.

Effects of solutes on ubiquitin diffusion and internal dynamics from ^{15}N relaxation experiments and atomistic MD simulations

Median R_2/R_1 ratios and the derived global correlation times in the different media (Table S1) are consistent with previous reports of translational and rotational diffusion becoming highly nonlinear on viscosity at high solute concentrations (72–79). Briefly, the correlation time increases linearly with viscosity for small molecules, whereas there is a strong negative deviation for Ficoll and PEG polymers, indicating that the microviscosity experienced by the protein is far lower than the macroviscosities of the polymer solutions (Fig. S5). For concentrated proteins, despite the fact that their viscosities are just above those of concentrated small molecules, the very high R_2 values suggest very slow tumbling due to stronger binding (72), but of a more random nature, i.e., less specific, given the very low CSPs.

^{15}N R_2/R_1 profiles (Fig. 4) show that ubiquitin dynamics are quite, albeit not totally, conserved across all conditions, since lower-than-average and higher-than-average ratios (which report on subnanosecond flexibility and micro-to-millisecond exchange processes, respectively) map to similar sequence segments. Profiles of root mean-square fluctuations per residue measured in MD simulations show almost no impact from high concentrations of arginine or glutamate, and only a mild dampening of flexibility in concentrated glucose (Fig. 4). However, inspection of slower, collective motions in the MD trajectories (mapped as projections on principal components derived from crystallographic data of ubiquitin bound to different proteins (61) (Figs. S6 and S7) suggests much slower exploration of its conformational space in glutamate and arginine relative to water, and extremely restricted conformational space exploration in glucose. Likewise, previous MD studies have shown only mild effects of protein-protein interactions, even unspecific ones, on backbone flexibility, despite large restrictions on conformational-space exploration (59,80). These predictions from MD are hard to probe experimentally and their magnitudes may be affected by inaccuracies in the simulations, but they strongly suggest that slow collective motions are much more affected by concentrated conditions than are fast local dynamics.

Perturbations of hydration structure and dynamics at the protein surface

High solute concentrations are known to slow down water dynamics and affect its structure, with hydrophobic and hydrophilic surfaces causing different effects (10,81–84) and with water structure and dynamics being perturbed up to 30–40 Å away from the surfaces of macromolecular crowders (85–87). These effects tune the solvent contribution to dielectric properties, molecular diffusion, and wetting capability. On the other hand, thermodynamic measurements between the tested solutes and various proteins indicate a general tendency for them to favor preferential hydration (44–48). In turn, our previous MD work predicted that large glucose concentrations would slow down the exchange of water molecules between the surface of a protein and the bulk (41), which new simulations presented here also suggest for arginine and glutamate despite their lower tendency to interact (Table S2).

In our NMR experiments, we observed, for all the small molecules tested, that the crosspeaks for Thr9 and Gly75, which are very weak in our diluted conditions, gained intensity upon titration instead of broadening out by the slower rotational tumbling as observed for the other crosspeaks (Figs. S1 and S8). This did not occur with Ficoll, PEG, and protein crowders, where all crosspeaks broadened. Specific resonance sharpening for Thr9 and Gly75 crosspeaks indicates that their intrinsic N-H exchange rates decrease faster than the decrease in global tumbling rate originated by the increased viscosity. This slowdown of the N-H exchange rate parallels what was observed for ubiquitin encapsulated inside reverse micelles, where the much slower water exchange rates even facilitated the measurement of crosspeaks to the water resonance in nuclear Overhauser spectroscopy (NOESY) spectra, otherwise invisible in diluted conditions at pH 7 (88). This constituted indeed our second proof of slower water exchange at the protein surface in the presence of small molecules at high concentrations. Also in such solutions, and without any micelle encapsulation, we could detect multiple crosspeaks to the water resonance in ^{15}N-resolved NOESY spectra (Fig. S9). The

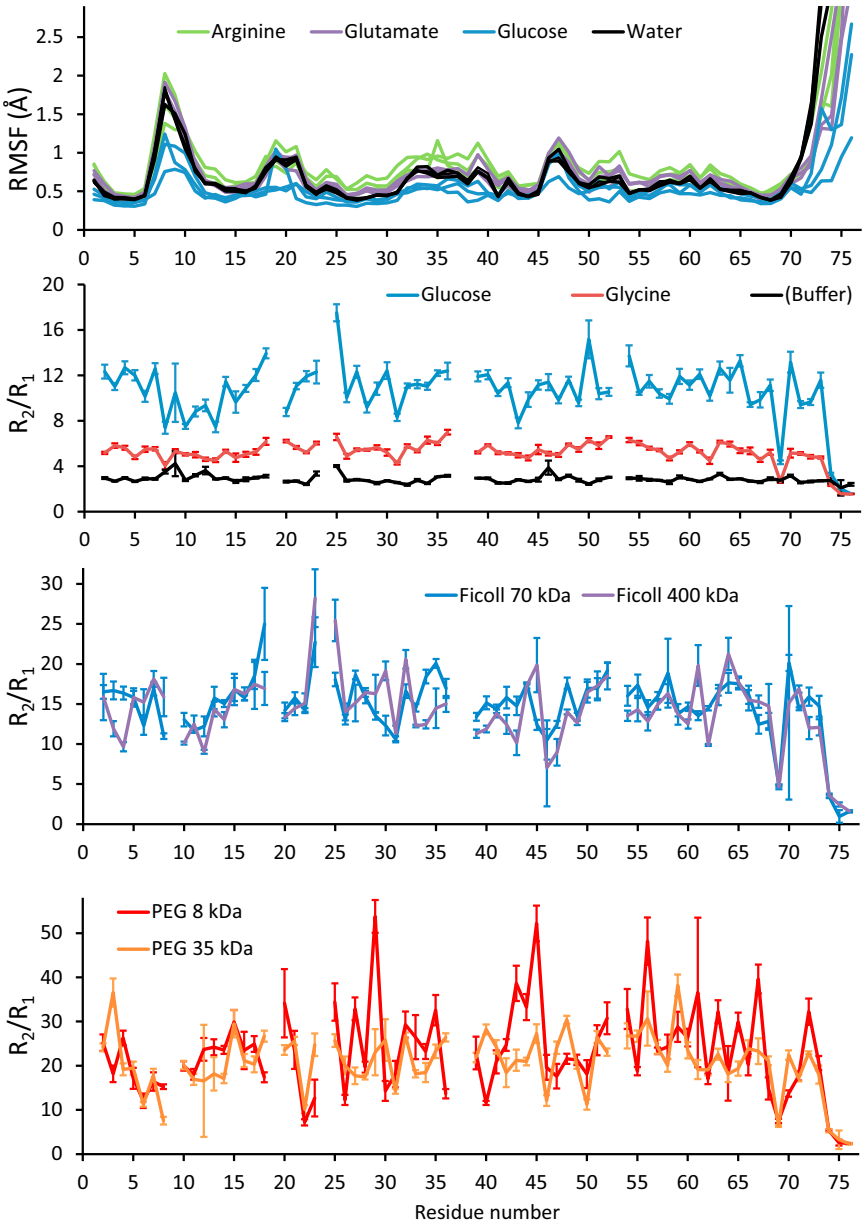

FIGURE 4 Fast ubiquitin dynamics as observed in simulations (*top*; root mean-square fluctuation of Cα atoms) and by ^{15}N relaxation (R_2/R_1 ratios). All systems contain 300 g/L of the indicated solute. In each R_2/R_1 profile, a positive deviation from the average indicates slow (microsecond- to millisecond-timescale) dynamics, whereas a ratio smaller than the average indicates fast (picosecond to nanosecond) fluctuations. Gaps correspond to weak or absent crosspeaks and three prolines in ubiquitin's sequence; error bars correspond to one standard error propagated from T_1 and T_2 uncertainties. The median value for each R_2/R_1 profile reflects the global tumbling rate in those conditions (provided in Table S1 as the global correlation time, which is larger for slower diffusional rotation), which combines direct effects from viscosity and from dynamic association with other molecules in the solution, i.e., weak interactions. In the root mean-square fluctuation plots from MD simulations, higher values indicate sequence segments of high flexibility in the sampled timescale (picosecond to submicrosecond). To see this figure in color, go online.

crosspeaks are not as many as observed in the reverse micelles, suggesting that the effects of high solute concentrations on water dynamics are milder than those produced by encapsulation. Yet, making our point, the more exposed N-H groups are those that gain more intensity in HSQC spectra and give stronger NOESY crosspeaks to the water resonance at high solute concentrations. With the Ficoll polymers, crosspeaks to the water resonance were also observed, but they were weaker and for fewer residues; with the PEG polymers, only a few such crosspeaks were clear.

MD simulations of ubiquitin in glutamate, arginine, and glucose also reveal a strong decrease in the exchange rates of surface water with the bulk, which we quantified at ~10-fold (Table S2). This magnitude is surely affected by force-field inaccuracies, but the trend is strong; moreover, the conclusion that water exchanges more slowly from the protein surface is consistent with previous measurements of H/D exchange rates for proteins in concentrated osmolyte solutions (89).

DISCUSSION

Solute size and chemistry jointly tune the extent of unspecific interactions

Studies from the 1970s to the early 2000s measured, through thermodynamic methods, the tendency of several solutes of various sizes and chemical features to preferentially bind to a series of proteins. Their main conclusion was that molecules that do not behave as strong protein denaturants,

i.e., amino acids, small polyols, small carbohydrates, and "inert" macromolecules like PEGs and polysaccharides, are preferentially excluded from interactions with proteins when present at high concentrations, favoring their hydration (44–48). For small molecules, such exclusion arises from their preference for interactions with solvent molecules over interactions with proteins. For inert macromolecular crowders, exclusion stems chiefly from steric repulsion, although actual interactions play a role as well in many cases, leading to a balance between entropic contributions from excluded volumes and enthalpic contributions from interactions, even for presumably inert polymers (35,46). Those studies recognized, though, that preferential exclusion by either mechanism does not preclude actual unspecific interactions between the solutes and the proteins. Estimations based on the chemical potentials measured in those works suggested small average numbers of molecules interacting with protein surfaces at any time, even if they were preferentially excluded overall. Moreover, those studies anticipated that the precise affinity of the interactions would vary throughout chemically diverse patches of protein surfaces and would depend on the physicochemical properties of the solutes, as recently investigated by measuring interactions of protein-group mimics with glycerol and PEG (90). Now, in this work, sensitive NMR techniques allow us to probe these interactions at the residue level by looking at proteins labeled with ^{15}N without interference from the concentrated solutes. Interpretation of the NMR results is not always straightforward, but MD simulations help propose explanations for many observations, besides providing a tool to probe questions not answerable by NMR.

Although the exact thermodynamic preferential interactions reported through the years for small solutes depended on each protein, the general trend is that amino acids tend to be more excluded than small sugars and polyols from protein surfaces (44–48). Consistent with that trend, our results from MD simulations (Fig. S2 and Table S2) and NMR (Fig. 2) show that glucose and glycerol (neutral small molecules rich in hydrogen-bond acceptors and donors) have a higher tendency to interact with the protein than the tested amino acids (small highly charged molecules). On the other hand, the CSPs induced by the macromolecular crowders suggest that very large polar surfaces rich in hydrogen-bond donors and acceptors (*Ficoll*, Fig. 1, *B* and *C*) are not much forced into unspecific interactions at high concentrations, but instead, hydrophobic surfaces are (*PEG*, Fig. 1 *B*). Similarly, fullerene molecules that are hydroxylated just enough to confer solubility, but still hydrophobic, also undergo interactions of a mainly hydrophobic nature with ubiquitin (91), and a recent study showed that α-synuclein undergoes unspecific interactions through aromatic hydrophobic residues in relevant cell types (27). That hydrophobic interactions with large molecules are particularly favored is consistent with the independent finding that protein-like aromatic molecules tend to interact with PEG (90); indeed, one of the strongest PEG-interacting regions observed here for ubiquitin maps around Phe45, whose aromatic ring is exposed. Finally, from the data using proteins as crowders it appears that both hydrophobic and electrostatic interactions play roles in establishing unspecific interactions. For other, more charged proteins (ubiquitin is neutral at pH 7 as used here), the contribution of electrostatic interactions might be more pronounced.

Joint consideration of the CSPs and relaxation data suggests that small solutes would undergo "weak docking" (thus producing significant CSPs) and simply a viscosity-dependent drag of rotational tumbling (Fig. S5), whereas protein interactions with macromolecules would involve more random poses and orientations, resulting in smaller CSPs (because different contributions simply cancel out), but with large effects on rotational tumbling because of their large molecular weights, augmented in the case of protein crowders by the stronger nature of the attractive interactions (Table S2; Fig. S5). Overall, it therefore appears that the extent and specificity of weak interactions between proteins and solutes depends on their sizes and surface properties: polar contacts are favored for small solutes, and contacts through hydrophobic regions for macromolecular crowders, although the latter are likely tuned by local electrostatics and hydrogen bonding.

Modulation of protein and water dynamics under concentrated conditions

In addition to the linear and nonlinear slowdowns in rotational and translational diffusion already dissected in the literature (72–79), our MD results indicate that concentrated media also slow down the collective protein dynamics that take place on timescales slower than tens of nanoseconds, which can be regarded as diffusion in a conformational space, although they do not affect much the fluctuations that occur on nanosecond and faster timescales. In line with these complex effects of concentrated solutes on protein dynamics, an NMR study of side-chain dynamics in calmodulin showed that microsecond-to-millisecond, but not picosecond-to-nanosecond, motions are affected in cell lysates compared to reference conditions in buffer (22). We note here that the few works reporting no or only mild effects on protein dynamics in vivo deal with intrinsically disordered proteins, which undergo extensive fast motions but no collective, conformational dynamics, which we identified as being most affected.

Effects on conformational dynamics are important because they could in principle modulate enzyme functionality, as observed upon mutation in some systems (92), thus potentially altering intracellular biochemistry or the properties of an enzyme in vitro. Our simulations predict that stronger-interacting molecules cause larger slowdown of the collective fluctuations (compare, for example, glutamate or arginine against glucose in Fig. S6), from which we extrapolate that the very

large polar and inert structures of the cells will not lead to important drags on internal dynamics, as they do not tend to interact, whereas unspecific protein-protein interactions would have more profound effects. In the cytoplasm, the net effect on the internal dynamics of a given protein will therefore depend on the extension and kind of unspecific interactions it experiences.

Regarding the slowdown of hydration dynamics at the protein surface, our results show a strong effect from small solutes but not from macromolecular crowders, although the latter do alter global water properties. Therefore, we can expect an overall mild net effect inside cells, because small molecules reach concentrations of 50–100 g/L out of the total solute concentration. Recent NMR experiments tailored to measure N-H exchange rates on two proteins in *E. coli* cells indeed showed little effect of the cytoplasm on the rates for most residues (26).

We finally note that effects on conformational and hydration dynamics could translate into effects on kinetics, but not necessarily on thermodynamics, as indeed reported in many works (26,28–32,41).

Causes and consequences of quinary protein structure and cytoplasmic organization

The growing picture about the intracellular architecture proposes that the aqueous phase of the cytoplasm is crowded and dynamic but ordered (4,93), with some degree of structural organization at the molecular level that has even been proposed to be reminiscent of the very emergence of life on Earth (94). Moreover, molecular crowding and the regulation of solute concentrations inside cells are closely related to various cellular processes, with variable, still-debated effects on kinetics and thermodynamics (95,96). Based on our findings, and in the context of findings by others, we hypothesize about the practical implications of high solute concentrations inside cells regarding cytoplasm structure and dynamics. First, we expect that very large, approximately immobile structures with hydrophilic surfaces, for example, from cytoskeleton fibers or granules of macromolecular carbohydrates, will behave similarly to Ficoll and will therefore keep soluble globular proteins largely excluded, not forcing interactions. Such large intracellular polar surfaces would simply lie beyond the upper limit for molecular crowding effects (97). On the other hand, hydrophobic patches on these large polar structures could act as points of contact for hydrophobic regions of other proteins, thus scaffolding quinary structures. Likewise, since even soluble globular proteins usually have a few exposed hydrophobic patches, unspecific protein-protein interactions are more likely to proceed through these patches, resulting in non-random, weakly ordered arrangements possibly tuned by electrostatic complementarity and hydrogen bonding. The distribution of a small number of surface hydrophobic patches limits the possible arrangements of soluble components, whose surfaces are mostly polar, leading to some degree of organization that could be relevant for ordering weak complexes, metabolons, and the local cytoplasmic environment. Supporting the idea that hydrophobic interactions mediate formation of weak assemblies, a recent experiment-based 3D model of the metabolon formed by enzymes of the Krebs cycle showed that most residues involved in the weak-contact surfaces are hydrophobic (98). Also, involvement of surface hydrophobic patches in unspecific interactions has been observed inside cells for several proteins (99). And reaching a size scale on the order of micrometers, liquid-liquid demixing mediated by hydrophobic effects at high protein concentrations is involved in the assembly of weak macromolecular complexes (100).

Interactions between small solutes and proteins could in principle strongly alter the diffusion of the small molecules in a molecule-specific fashion, despite the effective viscosity of the cytoplasm being only around two to three times higher than that of water (25,101). Since most intracellular metabolites are small charged molecules, we expect that they will behave like the tested amino acids rather than like the tested polyols. Therefore, we expect they will tend to interact very weakly with proteins and instead diffuse quite freely, possibly through diluted channels in the cytoplasmic matrix, as put forward in a description of the cytoplasm as a two-phase system of supercrowded gel and subcrowded cytosol in the form of flow-conducting circuits (4). It is interesting, though, that even without extensive interactions, small highly charged molecules are able to exert long-range perturbations into buried residues as they diffuse near the protein. The ultimate effect that this coupling could have on protein properties is to be explored, but interestingly, the 3D structure of the Krebs-cycle metabolon assembly mentioned above (98) suggests that channels with charged surfaces could direct the flow of substrates and products through the different active sites of the associated enzymes. Coupling of the strong electric fields of proteins to the charged metabolites could assist in directing their flow through the metabolon.

High solute concentrations in the context of protein variability and evolution

We finally note that ubiquitin's region most affected by the high solute concentrations maps to a surface patch around the exposed hydrophobic residues Ile44 and Phe45. This region has exquisitely evolved for binding different partners, being highly sensitive to mutations (102). In contrast, amino acid variation in soluble proteins is largely constrained by the requirement of polar amino acids at surfaces (103,104). We speculate that the tight functional constraints acting around ubiquitin's partners-binding surface might have prevented its evolution toward diminishing unspecific binding to it.

CONCLUSION

Studies of proteins in concentrated media that consider several types of molecules besides inert macromolecular crowders have gained momentum in the last decade, contributing to the discovery of new effects beyond those originated by excluded volumes and viscosity. We provided herein a systematic physical chemical investigation with the aim of disentangling the multiple facets of the complex interacting networks that proteins experience in native and other relevant environments. Our results, while helping to put into context recent findings on the biological effects of intracellular crowding, highlight once more the importance of studying proteins in their relevant conditions.

SUPPORTING MATERIAL

Supporting Materials and Methods, nine figures, and two tables are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(16)30542-2.

AUTHOR CONTRIBUTIONS

L.A.A. and M.D.P. designed the project; L.A.A. and E.S. performed and analyzed molecular dynamics simulations; L.A.A. produced protein samples and performed and analyzed NMR spectra; all three authors contributed to manuscript writing.

ACKNOWLEDGMENTS

We thank Antonio Donaire (University of Murcia, Murcia, Spain) for providing us with the plasmid to express ubiquitin; Sylvia Ho, Gisou van der Goot, and David Hacker (EPFL, Switzerland) for help with protein preparations; and Andrés Binolfi (Instituto de Investigaciones para el Descubrimiento de Fármacos de Rosario of the National Scientific and Technical Research Council of Argentina) for discussions.

This work was supported with funds from the École Polytechnique Fédérale de Lausanne (EPFL) and the Swiss National Science Foundation, and through access to high-performance computing resources provided by the Partnership for Advanced Computing in Europe and Centro Svizzero di Calcolo Scientifico (CSCS). L.A.A. acknowledges the European Molecular Biology Organization and the Marie Curie Actions for a Long-Term Postdoctoral Fellowship.

REFERENCES

1. Wirth, A. J., and M. Gruebele. 2013. Quinary protein structure and the consequences of crowding in living cells: leaving the test-tube behind. *BioEssays.* 35:984–993.
2. Ellis, R. J., and A. P. Minton. 2003. Cell biology: join the crowd. *Nature.* 425:27–28.
3. Zimmerman, S. B., and S. O. Trach. 1991. Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of *Escherichia coli*. *J. Mol. Biol.* 222:599–620.
4. Spitzer, J., and B. Poolman. 2013. How crowded is the prokaryotic cytoplasm? *FEBS Lett.* 587:2094–2098.
5. Conlon, I., and M. Raff. 2003. Differences in the way a mammalian cell and yeast cells coordinate cell growth and cell-cycle progression. *J. Biol.* 2:7.
6. Zeskind, B. J., C. D. Jordan, …, P. Matsudaira. 2007. Nucleic acid and protein mass mapping by live-cell deep-ultraviolet microscopy. *Nat. Methods.* 4:567–569.
7. Cheung, M. C., R. LaCroix, …, D. J. Ehrlich. 2013. Intracellular protein and nucleic acid measured in eight cell types using deep-ultraviolet mass mapping. *Cytometry A.* 83:540–551.
8. Wiśniewski, J. R., M. Y. Hein, …, M. Mann. 2014. A "proteomic ruler" for protein copy number and concentration estimation without spike-in standards. *Mol. Cell. Proteomics.* 13:3497–3506.
9. Bennett, B. D., E. H. Kimball, …, J. D. Rabinowitz. 2009. Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat. Chem. Biol.* 5:593–599.
10. Yancey, P. H., M. E. Clark, …, G. N. Somero. 1982. Living with water stress: evolution of osmolyte systems. *Science.* 217:1214–1222.
11. Timasheff, S. N. 1993. The control of protein stability and association by weak interactions with water: how do solvents affect these processes? *Annu. Rev. Biophys. Biomol. Struct.* 22:67–97.
12. Santoro, M. M., Y. Liu, …, D. W. Bolen. 1992. Increased thermal stability of proteins in the presence of naturally occurring osmolytes. *Biochemistry.* 31:5278–5283.
13. Wang, A., and D. W. Bolen. 1997. A naturally occurring protective system in urea-rich cells: mechanism of osmolyte protection of proteins against urea denaturation. *Biochemistry.* 36:9101–9108.
14. Leibly, D. J., T. N. Nguyen, …, W. C. Van Voorhis. 2012. Stabilizing additives added during cell lysis aid in the solubilization of recombinant proteins. *PLoS One.* 7:e52482.
15. Gil, D., and A. G. Schrum. 2013. Strategies to stabilize compact folding and minimize aggregation of antibody-based fragments. *Adv. Biosci. Biotechnol.* 4 (4a):73–84.
16. Lee, J., E.-W. Lin, …, H. D. Maynard. 2013. Trehalose glycopolymers as excipients for protein stabilization. *Biomacromolecules.* 14:2561–2569.
17. Arakawa, T., K. Tsumoto, …, D. Ejima. 2007. Biotechnology applications of amino acids in protein purification and formulations. *Amino Acids.* 33:587–605.
18. Hatti-Kaul, R. 2001. Aqueous two-phase systems. A general overview. *Mol. Biotechnol.* 19:269–277.
19. Monteith, W. B., R. D. Cohen, …, G. J. Pielak. 2015. Quinary structure modulates protein stability in cells. *Proc. Natl. Acad. Sci. USA.* 112:1739–1742.
20. Dedmon, M. M., C. N. Patel, …, G. J. Pielak. 2002. FlgM gains structure in living cells. *Proc. Natl. Acad. Sci. USA.* 99:12681–12684.
21. Barbieri, L., E. Luchinat, and L. Banci. 2015. Protein interaction patterns in different cellular environments are revealed by in-cell NMR. *Sci. Rep.* 5:14456.
22. Latham, M. P., and L. E. Kay. 2012. Is buffer a good proxy for a crowded cell-like environment? A comparative NMR study of calmodulin side-chain dynamics in buffer and *E. coli* lysate. *PLoS One.* 7:e48226.
23. Shi, X., Y. H. Foo, …, T. Wohland. 2009. Determination of dissociation constants in living zebrafish embryos with single wavelength fluorescence cross-correlation spectroscopy. *Biophys. J.* 97:678–686.
24. Gnutt, D., M. Gao, …, S. Ebbinghaus. 2015. Excluded-volume effects in living cells. *Angew. Chem. Int. Ed. Engl.* 54:2548–2551.
25. Ye, Y., X. Liu, …, C. Li. 2013. ^{19}F NMR spectroscopy as a probe of cytoplasmic viscosity and weak protein interactions in living cells. *Chemistry.* 19:12705–12710.
26. Smith, A. E., L. Z. Zhou, and G. J. Pielak. 2015. Hydrogen exchange of disordered proteins in *Escherichia coli*. *Protein Sci.* 24:706–713.
27. Theillet, F.-X., A. Binolfi, …, P. Selenko. 2016. Structural disorder of monomeric α-synuclein persists in mammalian cells. *Nature.* 530:45–50.
28. Phillip, Y., V. Kiss, and G. Schreiber. 2012. Protein-binding dynamics imaged in a living cell. *Proc. Natl. Acad. Sci. USA.* 109:1461–1466.

29. Phillip, Y., and G. Schreiber. 2013. Formation of protein complexes in crowded environments—from in vitro to in vivo. *FEBS Lett.* 587:1046–1052.

30. König, I., A. Zarrine-Afsar, …, B. Schuler. 2015. Single-molecule spectroscopy of protein conformational dynamics in live eukaryotic cells. *Nat. Methods.* 12:773–779.

31. Ebbinghaus, S., A. Dhar, …, M. Gruebele. 2010. Protein folding stability and dynamics imaged in a living cell. *Nat. Methods.* 7:319–323.

32. Dhar, A., K. Girdhar, …, M. Gruebele. 2011. Protein stability and folding kinetics in the nucleus and endoplasmic reticulum of eucaryotic cells. *Biophys. J.* 101:421–430.

33. Minton, A. P. 1981. Excluded volume as a determinant of macromolecular structure and reactivity. *Biopolymers.* 20:2093–2120.

34. Miklos, A. C., C. Li, …, G. J. Pielak. 2010. Volume exclusion and soft interaction effects on protein stability under crowded conditions. *Biochemistry.* 49:6984–6991.

35. Minton, A. P. 2013. Quantitative assessment of the relative contributions of steric repulsion and chemical interactions to macromolecular crowding. *Biopolymers.* 99:239–244.

36. Breydo, L., K. D. Reddy, …, V. N. Uversky. 2014. The crowd you're in with: effects of different types of crowding agents on protein aggregation. *Biochim. Biophys. Acta.* 1844:346–357.

37. Kuznetsova, I. M., B. Y. Zaslavsky, …, V. N. Uversky. 2015. Beyond the excluded volume effects: mechanistic complexity of the crowded milieu. *Molecules.* 20:1377–1409.

38. Chen, E., A. Christiansen, …, P. Wittung-Stafshede. 2012. Effects of macromolecular crowding on burst phase kinetics of cytochrome c folding. *Biochemistry.* 51:9836–9845.

39. Homouz, D., L. Stagg, …, M. S. Cheung. 2009. Macromolecular crowding modulates folding mechanism of alpha/beta protein apoflavodoxin. *Biophys. J.* 96:671–680.

40. Pozdnyakova, I., and P. Wittung-Stafshede. 2010. Non-linear effects of macromolecular crowding on enzymatic activity of multi-copper oxidase. *Biochim. Biophys. Acta.* 1804:740–744.

41. Spiga, E., L. A. Abriata, …, M. Dal Peraro. 2014. Dissecting the effects of concentrated carbohydrate solutions on protein diffusion, hydration, and internal dynamics. *J. Phys. Chem. B.* 118:5310–5321.

42. Breydo, L., A. E. Sales, …, V. N. Uversky. 2015. Effects of polymer hydrophobicity on protein structure and aggregation kinetics in crowded milieu. *Biochemistry.* 54:2957–2966.

43. Kuznetsova, I. M., K. K. Turoverov, and V. N. Uversky. 2014. What macromolecular crowding can do to a protein. *Int. J. Mol. Sci.* 15:23090–23140.

44. Arakawa, T., D. Ejima, …, S. N. Timasheff. 2007. Suppression of protein interactions by arginine: a proposed mechanism of the arginine effects. *Biophys. Chem.* 127:1–8.

45. Gekko, K., and S. N. Timasheff. 1981. Mechanism of protein stabilization by glycerol: preferential hydration in glycerol-water mixtures. *Biochemistry.* 20:4667–4676.

46. Bhat, R., and S. N. Timasheff. 1992. Steric exclusion is the principal source of the preferential hydration of proteins in the presence of polyethylene glycols. *Protein Sci.* 1:1133–1143.

47. Kita, Y., T. Arakawa, …, S. N. Timasheff. 1994. Contribution of the surface free energy perturbation to protein-solvent interactions. *Biochemistry.* 33:15178–15189.

48. Courtenay, E. S., M. W. Capp, …, M. T. Record, Jr. 2000. Vapor pressure osmometry studies of osmolyte-protein interactions: implications for the action of osmoprotectants in vivo and for the interpretation of "osmotic stress" experiments in vitro. *Biochemistry.* 39:4455–4471.

49. Babu, C. R., P. F. Flynn, and A. J. Wand. 2001. Validation of protein structure from preparations of encapsulated proteins dissolved in low viscosity fluids. *J. Am. Chem. Soc.* 123:2691–2692.

50. Martínez, L., R. Andrade, …, J. M. Martínez. 2009. PACKMOL: a package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.* 30:2157–2164.

51. Kirschner, K. N., A. B. Yongye, …, R. J. Woods. 2008. GLYCAM06: a generalizable biomolecular force field. Carbohydrates. *J. Comput. Chem.* 29:622–655.

52. Fox, T., and P. A. Kollman. 1998. Application of the RESP methodology in the parametrization of organic solvents. *J. Phys. Chem. B.* 102:8070–8079.

53. Horn, A. H. C. 2014. A consistent force field parameter set for zwitterionic amino acid residues. *J. Mol. Model.* 20:2478.

54. Daub, C. D., K. Leung, and A. Luzar. 2009. Structure of aqueous solutions of monosodium glutamate. *J. Phys. Chem. B.* 113:7687–7700.

55. McLain, S. E., A. K. Soper, and A. Watts. 2006. Structural studies on the hydration of L-glutamic acid in solution. *J. Phys. Chem. B.* 110:21251–21258.

56. Shukla, D., and B. L. Trout. 2010. Interaction of arginine with proteins and the mechanism by which it inhibits aggregation. *J. Phys. Chem. B.* 114:13426–13438.

57. Jorgensen, W. L., J. Chandrasekhar, …, M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.

58. Phillips, J. C., R. Braun, …, K. Schulten. 2005. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26:1781–1802.

59. Abriata, L. A., and M. Dal Peraro. 2015. Assessing the potential of atomistic molecular dynamics simulations to probe reversible protein-protein recognition and binding. *Sci. Rep.* 5:10549.

60. Binolfi, A., A. Limatola, …, P. Selenko. 2016. Intracellular repair of oxidation-damaged α-synuclein fails to target C-terminal modification sites. *Nat. Commun.* 7:10251.

61. Abriata, L. A., E. Spiga, and M. Dal Peraro. 2013. All-atom simulations of crowding effects on ubiquitin dynamics. *Phys. Biol.* 10:045006.

62. Han, Y., B.-S. Jin, …, J.-H. Lee. 2007. Effects of sugar additives on protein stability of recombinant human serum albumin during lyophilization and storage. *Arch. Pharm. Res.* 30:1124–1131.

63. Arakawa, T., and S. N. Timasheff. 1985. The stabilization of proteins by osmolytes. *Biophys. J.* 47:411–414.

64. Arakawa, T., K. Tsumoto, …, D. Ejima. 2007. The effects of arginine on protein binding and elution in hydrophobic interaction and ion-exchange chromatography. *Protein Expr. Purif.* 54:110–116.

65. Augspurger, J. D., J. G. Pearson, and D. Schwartz. 1992. Chemical-shift ranges in proteins. *J. Magn. Reson.* 100:342–357.

66. Hass, M. A. S., M. R. Jensen, and J. J. Led. 2008. Probing electric fields in proteins in solution by NMR spectroscopy. *Proteins.* 72:333–343.

67. Chen, J., S. K. Spear, …, R. D. Rogers. 2005. Polyethylene glycol and solutions of polyethylene glycol as green reaction media. *Green Chem.* 7:64–82.

68. Harris, D. P., A. T. Andrews, …, J. A. Asenjo. 1997. The application of aqueous two-phase systems to the purification of pharmaceutical proteins from transgenic sheep milk. *Bioseparation.* 7:31–37.

69. Hammes, G. G., and P. R. Schimmel. 1967. An investigation of water-urea and water-urea-polyethylene glycol interactions. *J. Am. Chem. Soc.* 89:442–446.

70. Li, C., and G. J. Pielak. 2009. Using NMR to distinguish viscosity effects from nonspecific protein binding under crowded conditions. *J. Am. Chem. Soc.* 131:1368–1369.

71. Kneller, J. M., M. Lu, and C. Bracken. 2002. An effective method for the discrimination of motional anisotropy and chemical exchange. *J. Am. Chem. Soc.* 124:1852–1853.

72. Wang, Y., C. Li, and G. J. Pielak. 2010. Effects of proteins on protein diffusion. *J. Am. Chem. Soc.* 132:9392–9397.

73. Kozer, N., Y. Y. Kuttner, …, G. Schreiber. 2007. Protein-protein association in polymer solutions: from dilute to semidilute to concentrated. *Biophys. J.* 92:2139–2149.

74. Kuttner, Y. Y., N. Kozer, …, G. Haran. 2005. Separating the contribution of translational and rotational diffusion to protein association. *J. Am. Chem. Soc.* 127:15138–15144.

75. Banks, D. S., and C. Fradin. 2005. Anomalous diffusion of proteins due to molecular crowding. *Biophys. J.* 89:2960–2971.

76. Sanabria, H., Y. Kubota, and M. N. Waxham. 2007. Multiple diffusion mechanisms due to nanostructuring in crowded environments. *Biophys. J.* 92:313–322.

77. Dix, J. A., and A. S. Verkman. 2008. Crowding effects on diffusion in solutions and cells. *Annu. Rev. Biophys.* 37:247–263.

78. Konopka, M. C., I. A. Shkel, …, J. C. Weisshaar. 2006. Crowding and confinement effects on protein diffusion in vivo. *J. Bacteriol.* 188:6115–6123.

79. Wang, Y., L. A. Benton, …, G. J. Pielak. 2012. Disordered protein diffusion under crowded conditions. *J. Phys. Chem. Lett.* 3:2703–2706.

80. Peters, J. H., and B. L. de Groot. 2012. Ubiquitin dynamics in complexes reveal molecular recognition mechanisms beyond induced fit and conformational selection. *PLOS Comput. Biol.* 8:e1002704.

81. Ball, P. 2008. Water as an active constituent in cell biology. *Chem. Rev.* 108:74–108.

82. Liu, P., X. Huang, …, B. J. Berne. 2005. Observation of a dewetting transition in the collapse of the melittin tetramer. *Nature.* 437:159–162.

83. Lum, K., D. Chandler, and J. D. Weeks. 1999. Hydrophobicity at small and large length scales. *J. Phys. Chem. B.* 103:4570–4577.

84. Major, R. C., J. E. Houston, …, X.-Y. Zhu. 2006. Viscous water meniscus under nanoconfinement. *Phys. Rev. Lett.* 96:177803.

85. King, J. T., E. J. Arthur, …, K. J. Kubarych. 2014. Crowding induced collective hydration of biological macromolecules over extended distances. *J. Am. Chem. Soc.* 136:188–194.

86. Sterpone, F., G. Stirnemann, and D. Laage. 2012. Magnitude and molecular origin of water slowdown next to a protein. *J. Am. Chem. Soc.* 134:4116–4119.

87. Harada, R., Y. Sugita, and M. Feig. 2012. Protein crowding affects hydration structure and dynamics. *J. Am. Chem. Soc.* 134:4842–4849.

88. Nucci, N. V., M. S. Pometun, and A. J. Wand. 2011. Site-resolved measurement of water-protein interactions by solution NMR. *Nat. Struct. Mol. Biol.* 18:245–249.

89. Foord, R. L., and R. J. Leatherbarrow. 1998. Effect of osmolytes on the exchange rates of backbone amide protons in proteins. *Biochemistry.* 37:2969–2978.

90. Knowles, D. B., I. A. Shkel, …, M. T. Record. 2015. Chemical interactions of polyethylene glycols (PEGs) and glycerol with protein functional groups: applications to effects of PEG and glycerol on protein processes. *Biochemistry.* 54:3528–3542.

91. Zanzoni, S., A. Ceccon, …, M. D'Onofrio. 2015. Polyhydroxylated [60]fullerene binds specifically to functional recognition sites on a monomeric and a dimeric ubiquitin. *Nanoscale.* 7:7197–7205.

92. González, M. M., L. A. Abriata, …, A. J. Vila. 2016. Optimization of conformational dynamics in an epistatic evolutionary trajectory. *Mol. Biol. Evol.* 33:1768–1776.

93. Luby-Phelps, K. 2013. The physical chemistry of cytoplasm and its influence on cell function: an update. *Mol. Biol. Cell.* 24:2593–2596.

94. Spitzer, J., G. J. Pielak, and B. Poolman. 2015. Emergence of life: physical chemistry changes the paradigm. *Biol. Direct.* 10:33.

95. Mourão, M. A., J. B. Hakim, and S. Schnell. 2014. Connecting the dots: the effects of macromolecular crowding on cell physiology. *Biophys. J.* 107:2761–2766.

96. Hoffman, L., X. Wang, …, M. N. Waxham. 2015. Relative cosolute size influences the kinetics of protein-protein interactions. *Biophys. J.* 109:510–520.

97. Miklos, A. C., C. Li, …, G. J. Pielak. 2011. An upper limit for macromolecular crowding effects. *BMC Biophys.* 4:13.

98. Wu, F., and S. Minteer. 2015. Krebs cycle metabolon: structural evidence of substrate channeling revealed by cross-linking and mass spectrometry. *Angew. Chem. Int. Ed. Engl.* 54:1851–1854.

99. Freedberg, D. I., and P. Selenko. 2014. Live cell NMR. *Annu. Rev. Biophys.* 43:171–192.

100. Li, P., S. Banjade, …, M. K. Rosen. 2012. Phase transitions in the assembly of multivalent signalling proteins. *Nature.* 483:336–340.

101. Mastro, A. M., M. A. Babich, …, A. D. Keith. 1984. Diffusion of a small molecule in the cytoplasm of mammalian cells. *Proc. Natl. Acad. Sci. USA.* 81:3414–3418.

102. Roscoe, B. P., K. M. Thayer, …, D. N. A. Bolon. 2013. Analyses of the effects of all ubiquitin point mutants on yeast growth rate. *J. Mol. Biol.* 425:1363–1377.

103. Abriata, L. A., T. Palzkill, and M. Dal Peraro. 2015. How structural and physicochemical determinants shape sequence constraints in a functional enzyme. *PLoS One.* 10:e0118684.

104. Abriata, L. A., C. Bovigny, and M. Dal Peraro. 2016. Detection and sequence/structure mapping of biophysical constraints to protein variation in saturated mutational libraries and protein sequence alignments with a dedicated server. *BMC Bioinformatics.* 17:242.

Biophysical Journal, Volume *111*

Supplemental Information

Molecular Effects of Concentrated Solutes on Protein Hydration, Dynamics, and Electrostatics

Luciano A. Abriata, Enrico Spiga, and Matteo Dal Peraro

Supporting Material for

Molecular Effects of Concentrated Solutes on Protein Hydration, Dynamics and Electrostatics

L. A. Abriata,[*] E. Spiga and M. Dal Peraro[*]

Additional Methods for NMR Experiments

Titrations were monitored through 2D phase-sensitive, ^{1}H-detected ^{1}H-^{15}N HSQC spectra using water flip-back pulses and gradient selection with decoupling during acquisition (pulse program hsqcetf3gpsi from the Bruker library, the closest literature description is that by Schleucher, Schwendinger, Sattler, Schmidt, Schedletzky, Glaser, Sorensen and Griesinger *J. Biomol. NMR* 1994). The resolution at acquisition was 2k points in the direct ^{1}H dimension and 128 points in the indirect ^{15}N dimension, on sweep widths of 12 and 32 ppm, respectively, and centered at 4.699 and 115 ppm, respectively. Spectra were processed with 2k points in ^{1}H and 256 points in ^{15}N, using linear prediction through 32 coefficients and a sine bell function in the ^{15}N dimension. A baseline of polynomial degree 5 was applied. At each point of the titration, we optimized the lock phase, the shim currents and the hard pulse length for ^{1}H. For all solutes but arginine and glutamate, the pulse length remained close to its value in buffer, *i.e.* around 11 μs. For arginine and glutamate, the 90º pulse for ^{1}H increased from ~11 μs to ~17 μs (glutamate) or ~18 μs (arginine) at the highest concentrations. We further point out that these titrations are reproducible and that the differences observed for CSPs among residues within spectra are well beyond the uncertainties in the chemical shift measurements (except, as indicated in the text, for the Ficoll crowders and for several residues in the titrations with proteins).

^{15}N relaxation data (T_1 and T_2) were acquired through pseudo-3D (direct ^{1}H dimension and indirect ^{15}N dimension, plus an extra dimension for time delays) phase-sensitive spectra with water flip-back pulses, gradient selection and decoupling during acquisition (Bruker pulse programs hsqct1etf3gpsi3d which uses inversion recovery to measure T_1, and hsqct2etf3gpsitc3d which uses interleaved temperature-compensated Carr–Purcell–Meiboom–Gill blocks to measure T_2). The relaxation delays for T_1 measurement were 8, 64, 136, 232, 336, 472, 664, 800 and 1600 ms for the determination of longitudinal relaxation times; and those for T_2 measurement were 33.9, 67.8, 101.8, 135.7, 169.6, 203.5, 237.4 and 271.4 ms. These values were chosen based on previous work (Tjandra, Feller, Pastor and Bax *J. Am. Chem. Soc.* 1995) with the addition of the 1600 ms delay for T_1 measurement which was especially important at high solute concentrations of small molecules. A recycle delay of 5 seconds was employed in all experiments given the long T_1 values operative under some of the high-concentration conditions. Acquisition and processing parameters of the ^{1}H-^{15}N HSQC planes in these pseudo-3D spectra was as described above for HSQC spectra. We further notice that we only measured ^{15}N relaxation on selected systems containing high-solute concentrations but with low ionic strength, where hard pulses remain as low as in buffer.

^{15}N-resolved NOESY (or NOESY-HSQC) spectra were acquired with a 3D (direct ^{1}H dimension, indirect ^{15}N dimension and indirect ^{1}H NOE dimension) ^{1}H-detected phase-sensitive spectrum with gradient selection and decoupling during acquisition (pulse program noesyhsqcetgp3d of the Bruker library, built up from Marion, Kay, Sparks, Torchia and Bax, *J. Am. Chem. Soc.* 1989). The sweep widths were of 14, 32 and 14 ppm centered at 4.699, 115 and 4.699 ppm, with 2k, 48 or 64 detected points and 2k, 64 and 128 points for processing, all respectively for the direct ^{1}H dimension, the indirect ^{15}N dimension and the indirect ^{1}H dimension. The mixing time (during which both NOEs and exchange processes can build crosspeak intensity) was 40 ms as used in similar works testing ubiquitin hydration dynamics (Nucci, Pometun and Wand *Nat. Struct. Mol. Biol.* 2011).

Additional Methods for Molecular Dynamics Simulations

The equilibration phase of each simulation involved 5 steps. First, conjugate gradient energy minimization with constrained Cα atoms. Second, warm-up from 10 K to 300 K in 30 K increments along 10000 steps followed by 20000 further steps at 300 K, with a timestep of 1 fs and constraints on the Cα atoms to their initial positions. Third, equilibration at 300K and 1 atm with constrained Cα atoms with a time step of 1 fs for 2 ns. Constraints on Cα atoms were then removed and the systems warmed up again from 10 to 300 K in 30 K increments along 10000 steps, followed by 20000 steps additional 1 fs steps at 300 K. This was then held at 300 K for 20,000 additional 2 fs steps, continuing into the production runs. The temperature was controlled using a Langevin thermostat with a damping coefficient of 5/ps. The pressure was kept at 1 atm using the Berendsen barostat with a relaxation time constant of 1 ps (warm-ups) or with a Langevin-Nosé-Hoover barostat with barostat oscillation time of 200 fs and damping time of 100 fs (constant temperature equilibrations and production runs). Non-bonded interactions (van der Waals and Coulombic) were treated with a cut-off of 12 Å with the construction of a neighbor list having pair list distance of 13.5 Å. Outside the cut-off the Coulombic interactions are calculated using the Particle Mesh Ewald method as implemented in NAMD.

Root mean squared fluctuations (RMSF) were computed for each Cα atom as:

$$RMSF_i = \sqrt{\frac{\sum_j \left(r_{i,j} - \bar{r}_i\right)^2}{N}}$$

where r_i denotes the position of Cα atom *i*, and *j* runs through the analyzed frames of the simulation. Calculations were done after alignment to the starting structure.

Translational diffusion coefficients were computed from the mean square displacements of each protein's center of mass during the simulation:

$$MSD(t')_{protein} = \frac{1}{t_{\max}} \sum_{t_0=1}^{t_{\max}} (\bar{r}(t_0 + t') - \bar{r}(t_0))^2$$

(where t_{max} is the simulation time in frames and *r* denotes positions of the protein's center of mass) by taking the slope of this time series and dividing by 6.

Interactions between water or solute molecules and the protein were investigated by computing survival probabilities for water-protein and sugar-protein contacts. Two atoms were considered to be in contact when the distance between them was lower than 1.1 times the sum of their Van der Waals radii as optimized in recent works. The survival probability was computed as reported by Marchi's group based on Impey's original formulation for ion hydration:

$$N(t) = \frac{1}{N_t} \sum_{n=1}^{Nt} \sum_j P_j(t_n, t)$$

where $Pj(t_n,t)$ takes the values of 1 if the j^{th} water/solute molecule is in contact with the protein between time t_n and t_n+t, and zero otherwise, and N_t is the number of frames. The decay in the survival probability for water was then fitted to a stretched exponential (where the stretching parameter γ varies from 0 for non-ideal diffusion to 1 for ideal diffusion) combined with two or more simple exponentials, as described by Marchi:

$$N(t) = n_s \exp(-(t/\tau_s)^{\gamma_S}) + \sum_{i=2}^{4} n_i \exp(-(t/\tau_i))$$

where n_s and n_i are the number of water/solute molecules with residence times τ_s and τ_i on the surface of the protein on each timescale (time scales reported in Table S2).

The radial distribution functions of water and solute molecules around the protein surface were evaluated using VMD's (Humphrey et al *J. Mol. Graphics* 1996) Radial Distribution Plugin.

Table S1. Properties of the tested solutes, and main results from NMR experiments and MD simulations.

Solute at 300 g/L	Structure, molar mass and polarity	Molar conc.	Dynamic viscosity η, 300K (mPa s)	Dielectric constant ε at 300K in concentrate aqueous solution	NMR					MD	
					CSP at 300 g/L						
					Range (ppm)	Trend vs. SASA(N)	τ_c (ns)	NOESY to water	R_1R2 median ± median deviation	Average number of molecules in contact with protein	Timescale for full exploration of conform-ational space
(none)	--		0.798	78			4.2	None	13 ± 1		Few tens of ns
Glucose	Polyol, 180 g/mol	1.67	2.21	62	0.060 – 0.173	+	10.2	Strong	14 ± 1	75	> us
Glycerol	Polyol 92 g/mol	3.26	1.87	68	0.027 – 0.138	+	(N.D.)	Strong	(N.D.)		
Sodium glutamate	Negative Zwitterion 169 g/mol	1.78	N.A.	N.A., but >>78 according to MD	0.065 – 0.184	-	(N.D.)	Strong	(N.D.)	34	Hundreds of ns
Arginine hydro-chloride	Positive Zwitterion 210.7 g/mol	1.42	1.3		0.063 – 0.185	-	(N.D.)	Strong	(N.D.)	33	Hundreds of ns
Glycine	Neutral Zwitterion 75.1	3.99	1.77	166	0.241 – 0.374	-	6.6	Strong	14 ± 1		
Ficoll 70kDa	Sucrose polymer, avg. 70 kDa	0.0043	20-30	N.A.	0.012 – 0.045	+	12.3	Strong, for few residues	17 ± 2		
Ficoll 400kDa	Sucrose polymer, avg. 400 kDa	$7.5\ 10^{-4}$ (avg.)	50-60	57	0.013 – 0.034	+	11.9	Strong, for few residues	17 ± 3		
PEG 8kDa	Polymer of etylene oxide, avg. 8 kDa	0.038 (avg.)	20-50	~50	0.011 – 0.121	+	15.1	Weak, for few residues	15 ± 4		
PEG 35kDa	Polymer of etylene oxide, avg. 35 kDa	0.0086 (avg.)	N.A. but >> 50	N.A.	0.013 – 0.121	+	14.5	Weak, for few residues	15 ± 2		
Lysozyme (200 g/L)	14.3 kDa	0.011 (avg.)	1.6	N.A. but << 78	0.001 – 0.051	?	N.D. but likely high	Weak, for few residues	(N.D.)		
Bovine serum albumin	66.5 kDa	0.0045 (avg.)	4.8	N.A. but ~90 at 70 g/L	0.008 – 0.053	?	N.D. but likely high	Weak, for few residues	(N.D.)		

N.D. = Not determined; N.A. = Not Available

Table S2. Analysis of solute-protein and water-protein contacts in MD simulations.

	Water	Glucose	Arginine	Glutamate
Water exchange				
Fastest timescale (ps)	30	170	200	200
Second fastest (ns)	0.14	0.7	1	2
Third fastest (ns)	0.92	7	10	17
Slowest timescale (ns)	13	100	100	180
Average number of solutes interacting with protein surface	--	75	33	34
Dielectric constant	77.6	69.5	>> 78	>> 78
Range of rotational diffusion times	1.8 ns	> 1000 ns	23 – 31 ns	129 – 664 ns
	From Spiga et al. *J. Phys. Chem. B* 2014	(Data for solutes from simulations in this work)		

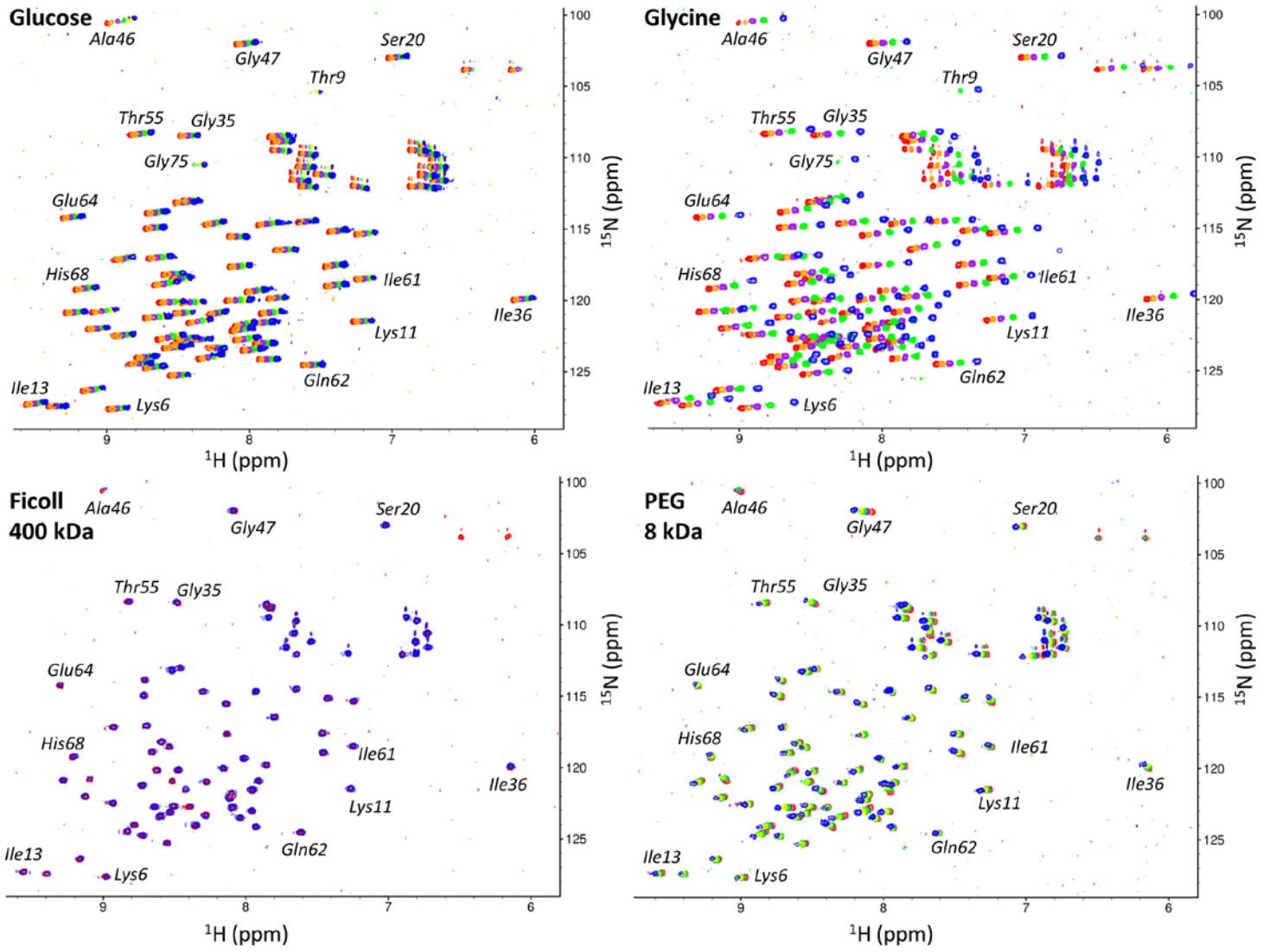

Figure S1. Effects on ^{1}H,^{15}N HSQC spectra (internally referenced to the water resonance) of stepwise additions of glucose, glycine, Ficoll 400 kDa or PEG 8 kDa to a 200 µM solution of ^{15}N-labeled ubiquitin, reaching in all cases 300 g/L final solute concentrations. In the four cases, the spectrum colored red is in pure buffer; then, color codes are:

- Glucose: orange = 50 g/L, violet = 150 g/L, green = 200 g/L, yellow = 250 g/L, blue = 300 g/L.
- Glycine: orange=45 g/L, violet=91 g/L, green=166 g/L, blue=300 G/L.
- Ficoll 400 kDa: blue is in 300 g/L.
- PEG 8 kDa: green=100 g/L, yellow=200g/L, blue=300 g/L

In the four spectra, the crosspeak for Alanine 46 is folded in the ^{15}N dimension.

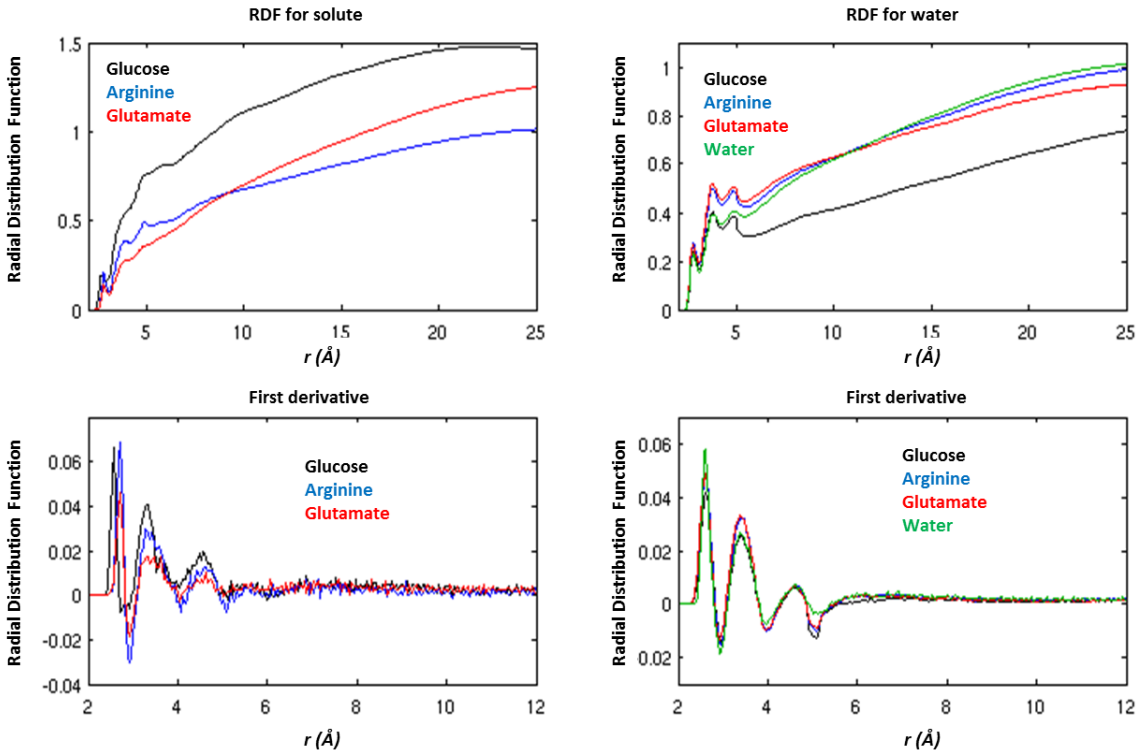

Figure S2. Radial distribution functions (top) for solutes (left) and water (right) relative to the protein surfaces, and first derivative with respect to the distance (bottom).

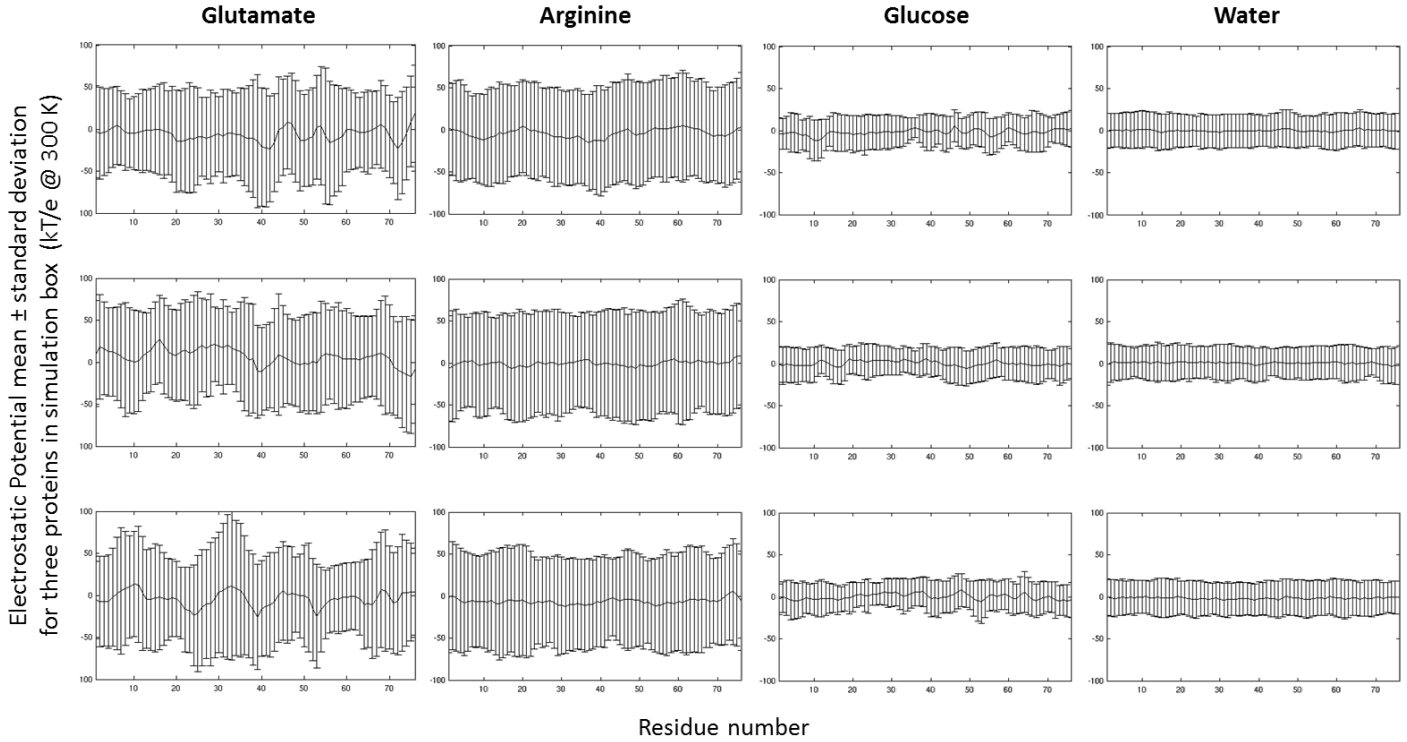

Figure S3. Trajectory average and standard deviation of the electrostatic potentials at each N atom of ubiquitin, for three proteins in the simulation boxes in water, glucose, glutamate and arginine.

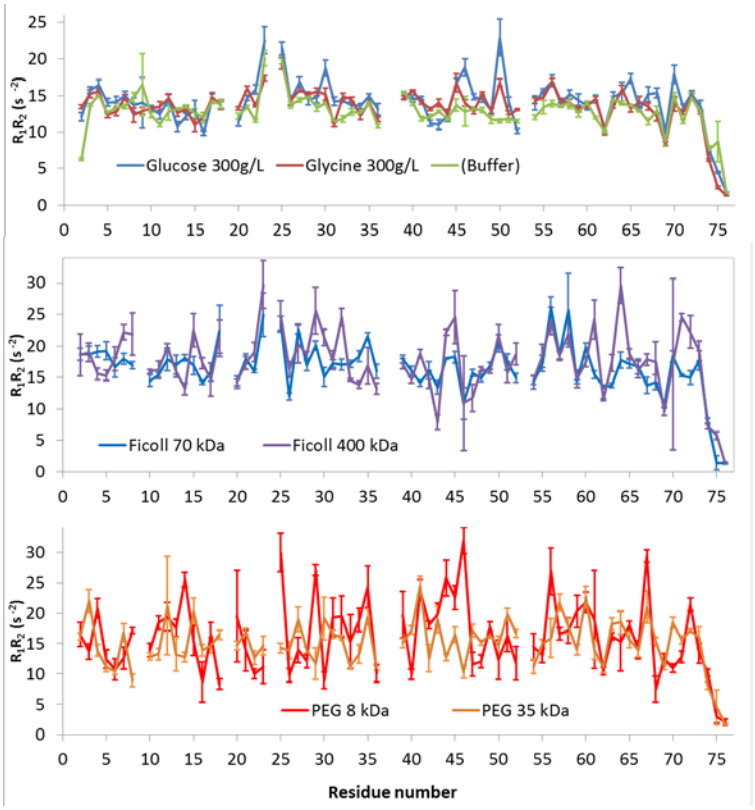

Figure S4. R_1R_2 products against ubiquitin's sequence, in buffer, concentrated glucose, glycine, Ficoll polymers and PEG polymers. Gaps correspond to weak or absent cross peaks and three prolines in ubiquitin's sequence. Error bars correspond to one standard error propagated from T1 and T2 uncertainties.

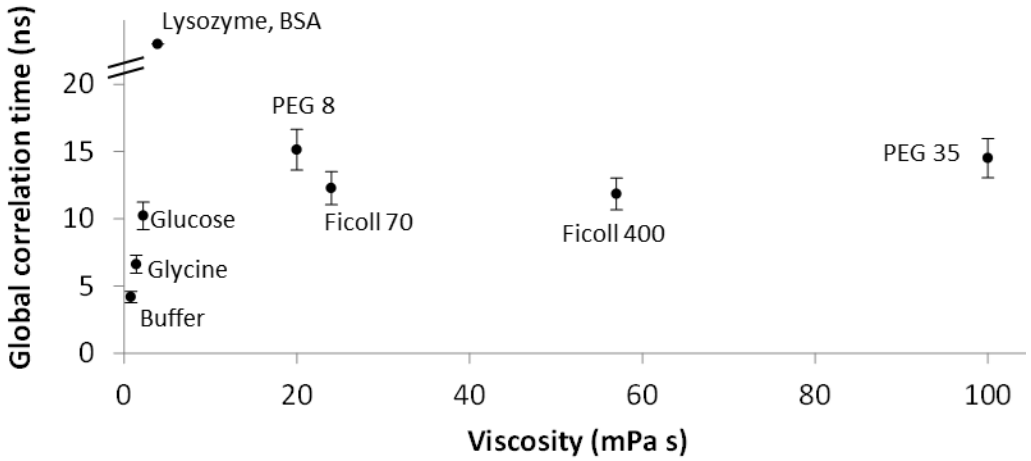

Figure S5. Global correlation time against viscosity for ubiquitin solutions studied by ^{15}N relaxation at 300 g/L of the added solutes. In lysozyme and BSA we do not report actual values of the global correlation times, we simply indicate that it will be larger than 20 ns given the spectral broadening. All viscosities are as reported in Table S1.

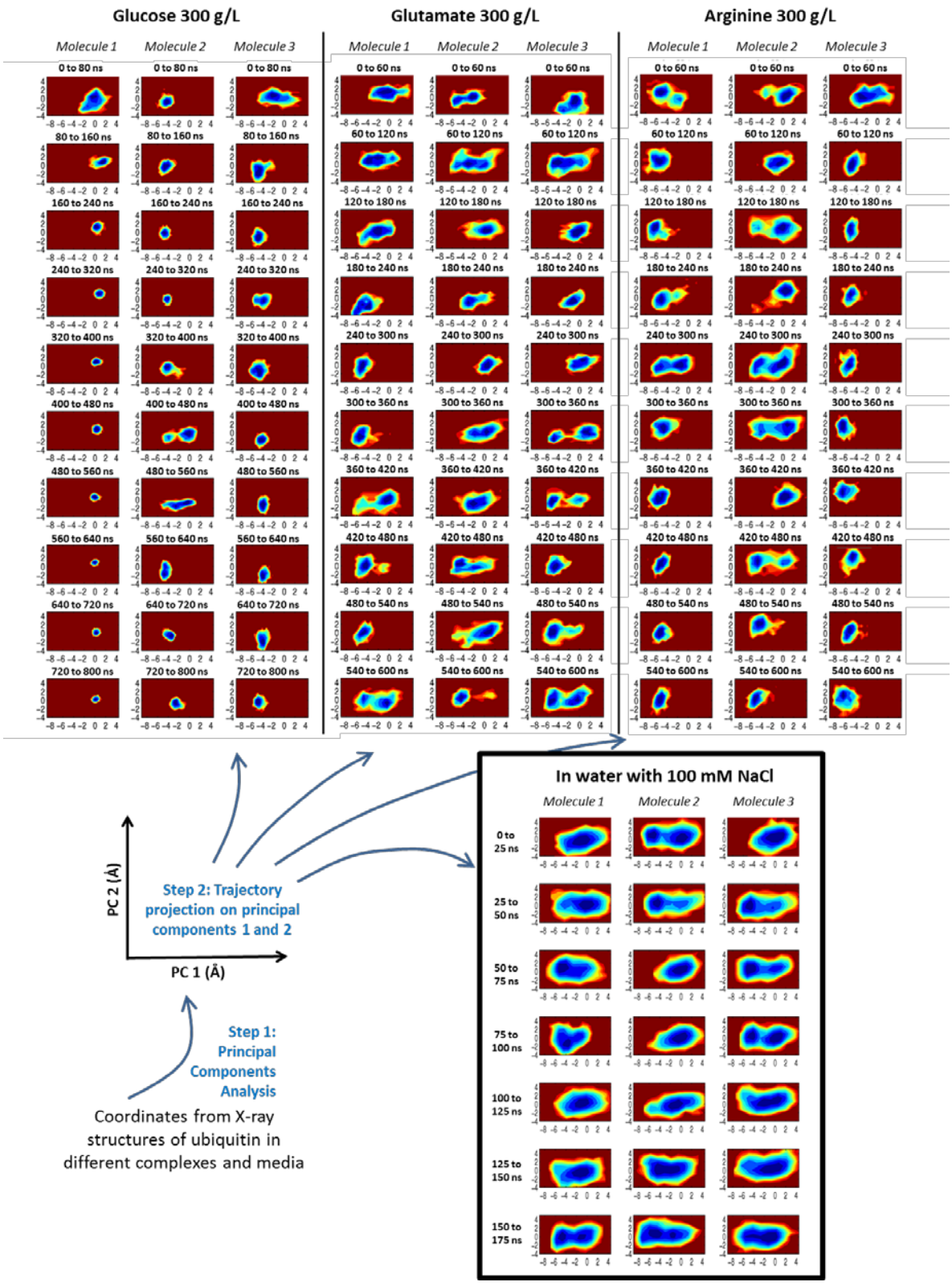

Figure S6. Projection of the conformations of each protein simulated in each system, on a plane of principal components that captures the conformational variability of ubiquitin when bound to different proteins as seen in X-ray structures. The x and y axes are principal components 1 and 2, respectively accounting for 42.6 and 12.5 % of the variability in the X-ray data, and both having units of Ångström. The diagram on the bottom left graphically explains how it was built (details can be found in Abriata et al *Phys. Biol.* 2013 or in Spiga et al *J. Phys. Chem. B* 2014). The inset shows how three ubiquitin molecules in a single box in water explore the conformational space much faster than in the presence of concentrated solutes (notice 25 ns time windows) (From Abriata and Dal Peraro *Scientific Reports* 2015, before protein aggregation).

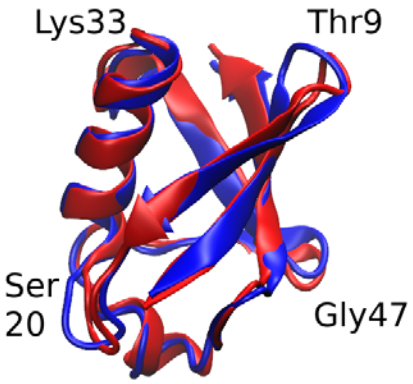

Figure S7. Structures of ubiquitin conformations at the deepest points of the basins observed in the simulations in water. The structure in red corresponds to the left-most basin, the one in blue to the right-most basin. Important amino acids from each loop are indicated. The flexible carboxi-terminal tail (residues 72-76) is not shown.

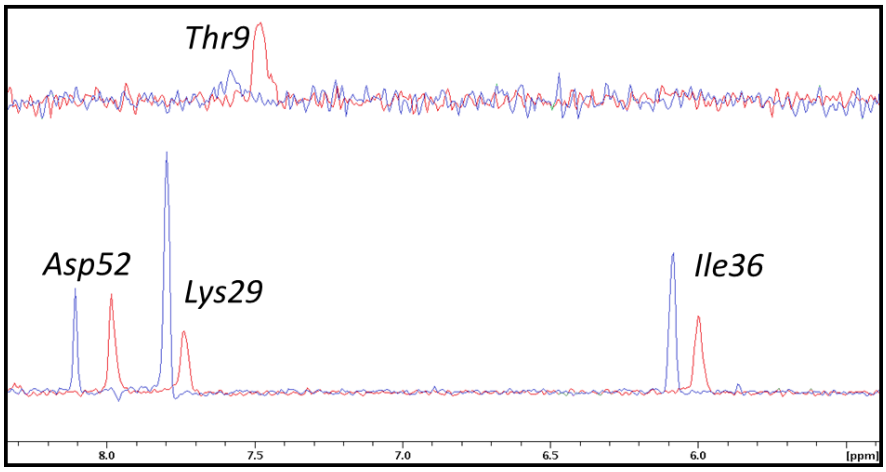

Figure S8. ^{1}H slices of ^{1}H,^{15}N HSQC spectra showing how the signal-to-noise ratio increases for Thr9's N,H cross peak in 300 g/L glucose (red) relative to buffer (blue) while the other resonances broaden (exemplified with Asp52, Lys29 and Ile36). The two spectra were acquired with the same protein concentration (200 μM) in the same buffer at controlled pH 7, with the same receiver gain, number of scans and resolution in the ^{15}N dimension, each with its own optimized tuning, matching and pulses.

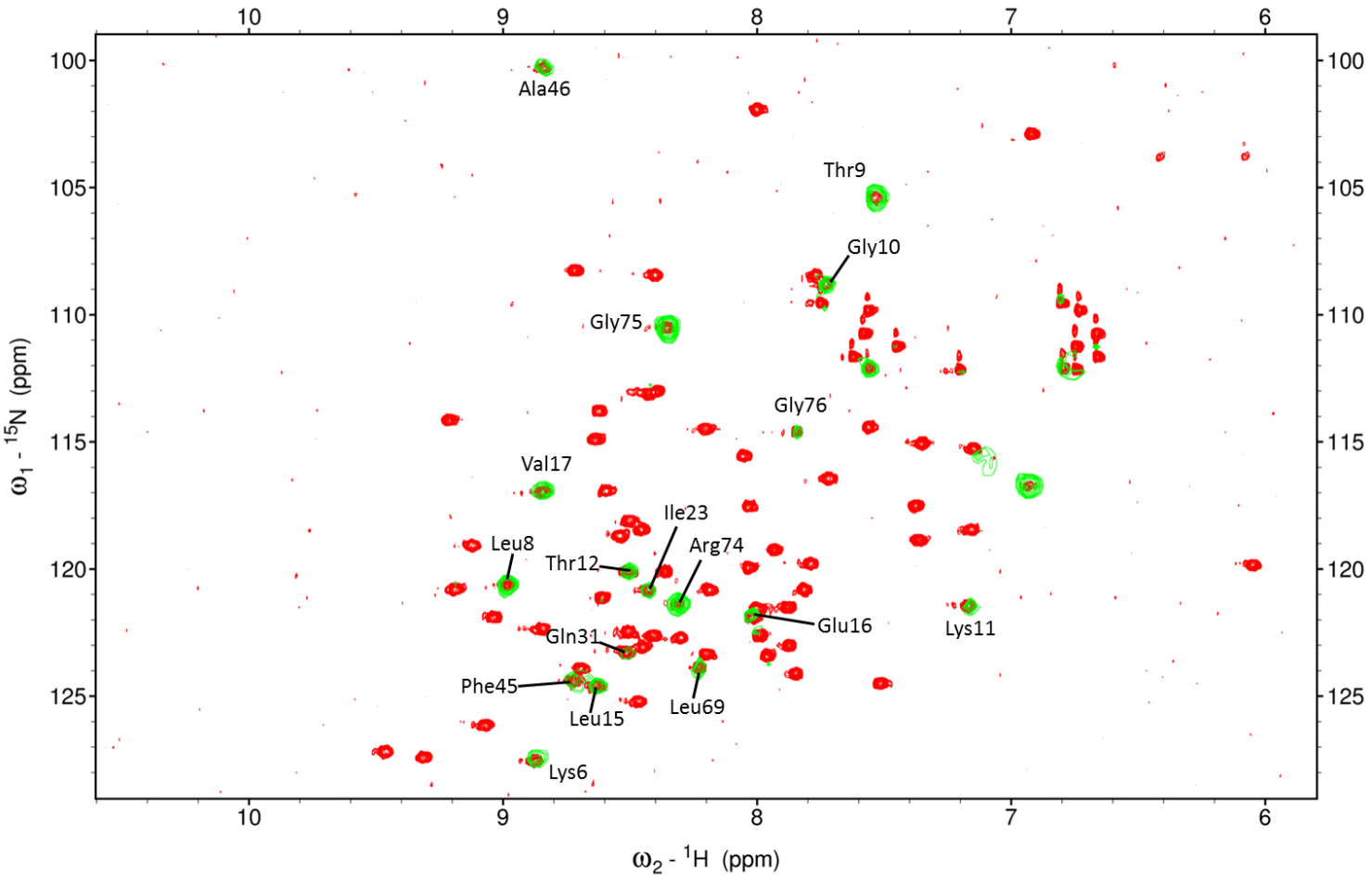

Figure S9. Plane at 4.7 ppm in the NOESY ^{1}H dimension of a 3D NOESY-HSQC spectrum of ubiquitin in 300 g/L glucose (green) overlaid on top of an HSQC spectrum of the same sample in in identical conditions. Crosspeaks from backbone resonances with enhanced NOESY intensities are labeled.

References for viscosity and static dielectric constant data in Table S1

Telis et al. Viscosity of Aqueous Carbohydrate Solutions at Different Temperatures and Concentrations, *International Journal of Food Properties* (2007) 10 (1):185-195

Mason et al. The Viscosities of Aqueous Solutions of Amino Acids at 25 and 35º, *Journal of the American Chemical Society* (1952) 74 (5):1287-1290

Saini et al. Volumetric analysis of L-arginine hydrochloride in aqueous and aqueous tetrahydrofuran solution at 303.15 K *Archives of Applied Science Research* (2012) 4 (5):2068-2076

Wyman et al. The Dielectric Constant of Solutions of Amino Acids and Peptides, *Journal of the American Chemical Society* (1933) 55:908-914

Ciu et al. The Effect of Molecular Crowding on the Stability of Human c-MYC Promoter Sequence I-Motif at Neutral pH, *Molecules* (2013) 18, 12751-12767; doi:10.3390/molecules181012751

Capuano et al. Electrostatic and Excluded Volume Effects on the Transport of Electrolytes in Poly(ethylene glycol)—Water "Mixed Solvents". *J. Phys. Chem. B* 2003, 107, 12363–12369

Czerwinski et al. The structural basis for the perturbed pKa of the catalytic base in 4-oxalocrotonate tautomerase: kinetic and structural effects of mutations of Phe-50. *Biochemistry* (2001) 40, 1984–1995

Mali et al. Dielectric relaxation of poly ethylene glycol-water mixtures using time domain technique. *Indian Journal of Pure & Applied Physics* (2007) 45, 476-481

Zaslavsky et al. Dielectric Properties of Water in the Coexisting Phases of Aqueous Polymeric Two-phase Systems. *J. Chem. Soc. Faraday Trans.* (1989) 85(9), 2857-2865

Stenger et al. Molecular Weight Dependence of the Depletion Attraction and its Effects on the Competitive Adsorption of Lung Surfactant. *Biochim. Biophys. Acta* (2008) 10:2032-2040

Cametti et al. Dielectric Relaxation Spectroscopy of Lysozyme Aqueous Solutions: Analysis of the δ-Dispersion and the Contribution of the Hydration Water. *J. Phys. Chem. B* (2011) 115, 7144-7153

Floros et al. Detailed study of the dielectric function of a lysozyme solution studied with molecular dynamics simulations. *Eur. Biophys. J.* (2015) DOI 10.1007/s00249-015-1052-7

Grant et al. The Dielectric Behavior of Aqueous Solutions of Bovine Serum Albumin from Radiowave to Microwave Frequencies. *J. Phys. Chem.* (1968) 72(13), 4373-4380