Supporting Information

Sekhar et al. 10.1073/pnas.1303273110

SI Materials and Methods

Ubiquitin Sample Preparation. [U-^{15}N, ^{13}C] ~50% ^{2}H ubiquitin was overexpressed in *Escherichia coli* grown in 70% D_2O M9 minimal media with ^{13}C glucose and $^{15}NH_4Cl$ as the sole carbon and nitrogen sources and purified as discussed in the literature (1). Each 0.5-mL sample consisted of ~1.5 mM protein dissolved in 20 mM sodium acetate, 10 mM NaCl, 2 mM EDTA, 2 mM NaN_3, and 2 mM tryptophan, pH 5.0. Four samples were used in total with varying D_2O and glycerol concentrations (D_2O, %; glycerol, %): (10; 0) (10; 25), (35; 0), (35; 25).

Additional Details on Diffusion Measurements. The diffusion coefficients of the probe molecules in 200 mg/mL BSA, used to construct the molecular ruler of Fig. 4, were measured using pulsed field gradient (PFG) NMR methods (2) as outlined in *Materials and Methods*. Because the spectra for 10 out of the 11 probes were not isotope-edited (the exception being datasets of the ^{15}N-labeled FF domain), resonances arising from BSA can overlap with those of the probe molecules in the aliphatic (0.5–4.0 ppm) and aromatic/amide (6.5–10.5 ppm) regions of the spectrum. To minimize systematic errors in the measurement of probe peak intensities in BSA, we used a spin-lock-based T_2 filter (3) at the beginning of the water-suppressed longitudinal encode–decode pulse sequence (4). The spin-lock module was implemented using a weak ^{1}H radiofrequency field (2–4 kHz) for a duration of 40 ms immediately following the first excitation 90° pulse. The offset for the spin-lock field was chosen based on the particular probe molecule. BSA resonances were entirely dephased during this period. The spin-lock T_2 filter does lead to a loss in spectral sensitivity for each probe, the extent of which depends on the probe size. This loss is problematic only for the H13 DNA (400 μM) (5), and thus the spin-lock filter was not used in this case. In addition to the spin-lock module, resonances removed from BSA peaks were chosen where possible for diffusion analysis. For instance, the anomeric protons resonating at ~5.4 ppm were used to obtain diffusion constant estimates for sucrose, raffinose, stachyose, and H13 DNA.

Details on NMR Dynamics Measurements. Backbone ^{15}N R_1 and $R_{1\rho}$ rates, measured with enhanced sensitivity based pulse schemes (6), were used to estimate the overall rotational correlation time (τ_R) of ubiquitin. In this study the rotational diffusion tensor of ubiquitin was assumed to be isotropic (*i*) because detailed NMR relaxation studies have shown that it is only slightly anisotropic $D_{par}/D_{perp} = 1.17$ (7) and (*ii*) our goal is to quantify dynamics of Asn, Gln, Arg, and methyl-containing side chains where the appropriate bond vectors are dynamic and not well described by the orientations in a static structure. The τ_R values for ubiquitin, 25 °C, without and with 25% (vol/vol) glycerol are 5.58 ± 0.02 ns and 14.4 ± 0.02 ns, respectively.

Methyl ^{2}H R_1, R_2 rates of $^{13}CH_2D$ methyl groups in fractionally deuterated ubiquitin were measured using previously described experiments (8) and analyzed using the Lipari–Szabo (LS2) model (9), where local motion is described by two parameters: S^2_{axis}, which quantifies the amplitude of motion of the methyl threefold axis, and τ_f, the timescale of the dynamics. Details of the analysis are as published previously (8).

Side-chain dynamics of Asn, Gln ^{15}NHD, and Arg $^{15}N^{\varepsilon}$–H^{ε} moieties were studied by measuring ^{15}N spin relaxation rates, R_1 and $R_{1\rho}$, and ^{1}H–^{15}N NOEs using samples prepared with 35% D_2O/65% H_2O (^{15}NHD) and 10% D_2O/90% H_2O (Arg). Pulse schemes were identical to those used for the backbone amides (6). In the case of Asn/Gln the analysis included contributions from both ^{1}H–^{15}N and ^{2}H–^{15}N one-bond dipolar interactions. ^{15}N chemical shift anisotropy values and amide N–H or N–D bond vector lengths of $\Delta\sigma = -164.4$ ppm, $r = 1.02$ Å (10, 11) and -114.0 ppm, $r = 1.04$ Å (12) were used in the analysis of the Asn/Gln and Arg data, respectively. Relaxation data were analyzed with an extended Lipari–Szabo (LS3) (13) model in which a single-order parameter S_f^2 is used to describe the assumed infinitely fast local motion along with S_s^2 and τ_S for the slower (100 s of picoseconds to several nanoseconds) timescale motion. All analyses were performed using the program Dasha (14), supplemented with in-house written Matlab scripts. For the analysis of the side-chain relaxation data τ_R was set to the value obtained from the backbone amide R_1, $R_{1\rho}$ rates (15).

Order parameters and effective correlation times are plotted in Fig. S1 for ubiquitin in aqueous and 25% (vol/vol) glycerol buffers.

SI Text

Analysis of the I–N Transition as a Two-State Conformational Exchange Process. In the analysis described in the text and in a previous paper (16), establishing that the frictional forces originating from random solvent fluctuations play a dominant role in FF domain folding, we have assumed that the interconversion between I and N states (17) can be well approximated by a two-state transition. The exchange rates isolated from fits of NMR relaxation dispersion profiles assuming this simple model of exchange have subsequently been used in concert with the Kramers equation in the high-friction limit (Eqs. **1** and **2** in the main text) (18, 19) to calculate the effective hydrodynamic radius (EHR) of the I–N transition. In what follows we justify the use of a two-state exchange model to describe the I–N interconversion through a comparison between experimentally derived exchange parameters and those obtained on the basis of computations (see below) and then discuss how a more complex exchange process might modify the conclusions of this work.

We first consider comparative bootstrap analyses (20) of experimental and simulated data, where the latter is explicitly two-state, to show that the distributions of exchange parameters (p_I, the fractional population of state I, and k_{ex}) obtained in fits of the experimentally derived dispersion profiles are consistent with a two-state exchange process. Simulated ideal two-state datasets ($R_{2,eff,sim}$ vs. ν_{CPMG}) have been constructed for each viscogen concentration (glycerol or BSA) as follows. Values of $R_{2,eff}$ at each concentration ($R_{2,eff,fit}$), obtained from global fits of experimental $R_{2,eff}$ rates ($R_{2,eff,expt}$) to a two-state model, have been used to construct simulated datasets by the addition of Gaussian noise scaled by the error in $R_{2,eff,expt}$ (σ_{R2}) as follows:

$$R_{2,eff,sim} = R_{2,eff,fit} + \sigma_{R2} * h(0,1),$$

where h(0,1) is a random number from a Gaussian distribution with a mean of 0 and a variance of 1. The $R_{2,eff,sim}(\nu_{CPMG})$ profiles so obtained are ideally two-state because they were generated by numerically propagating the Bloch–McConnell equations (21) for two-state exchange (using the in-house program CATIA, http://pound.med.utoronto.ca/software.html). Subsequently, we have used a bootstrap analysis of the resulting dispersion dataset, comprising, say, *m* profiles (one for each "residue") each with *n* ν_{CPMG} values [details are given in supporting information of Sekhar et al. (16)]. The "residues" chosen are those for which dispersion data are obtained at each viscogen concentration (see below). A series of 1,000 trial datasets are generated, whereby each of the n $R_{2,eff,sim}$ values in each of the m profiles is selected randomly an arbitrary number of times, such that the total number of $R_{2,eff}$ points in each dispersion profile is kept constant. Note that in a given trial dataset each of the dispersions is generated in a similar manner so that $R_{2,eff,sim}$ point j, $1 \le j \le n$, is either kept or replaced in each of the profiles. The 1,000 datasets so generated are globally fit to a two-state model of exchange to extract p_I and k_{ex}. Subsequently the procedure was repeated for the experimental data, using $R_{2,eff,expt}$ values rather than $R_{2,eff,sim}$ rates, and the resultant distributions of (p_I, k_{ex}) plotted in Fig. S2 for various glycerol and BSA concentrations. Note that the distributions obtained for experimental and simulated data are very similar and the mean k_{ex} values are identical within error, indicating that a third state, if present, does not significantly affect the experimental exchange rate constants as a function of viscogen concentration.

The simulated and experimental bootstrap datasets described above were also fit on a single residue basis to a two-state exchange model. Distributions of k_{IN} and k_{NI} for residues of the FF domain, both in the absence and in the presence of glycerol or BSA, are shown in Figs. S3–S5. The experimental and simulated distribution profiles for each residue are very similar, arguing again that the I–N interconversion can be well modeled as a two-site exchange process. A more complex exchange mechanism, involving additional states, would result in considerably larger differences in (p_I, k_{ex}) experimental distributions across the different residues because of residue-specific chemical shift differences in the additional states, as observed previously in studies of conformational exchange of the Fyn SH3 domain (17, 22). The relatively narrow distributions of single-residue k_{IN} and k_{NI} values in the experimental datasets both in the absence and in the presence of viscogen as well as the similarity of distributions in experimental and simulated datasets justifies the analysis of the relaxation dispersion data in the context of a two-site exchange model. It does not, however, prove that the exchange is in fact two-state (see below).

Support for the robustness of the two-state model of chemical exchange is obtained from the $\Delta\varpi_{IN}$ values generated from fits of experimental Carr–Purcell–Meiboom–Gill (CPMG) relaxation dispersion data recorded using different viscogen concentrations because chemical shifts are independent parameters not constrained during the fitting procedure. Values of $\Delta\varpi_{IN}$ from fits of profiles measured on samples with the highest concentrations of glycerol or BSA show very good correlations with values obtained in the absence of viscogen (Fig. S6), confirming that the dispersion measurements probe conformational fluctuations between the same set of states.

The above bootstrap and chemical shift analyses indicate that our CPMG dispersion data can be explained in the context of a two-state model. This does not prove that additional states are not present along or off the I–N pathway. Our measurements cannot discern the presence of states with fractional populations lower than ~0.2% (see below) and/or that are in fast exchange on the chemical shift timescale with I or N. In particular, if there are structurally distinct states in fast exchange with I, then the I state will be averaged. However, there are indications that the I state for the FF domain can be represented by a single structure. First, it has been possible to create a mimic of the intermediate state by truncating the C-terminal helix H4 of the FF domain, and this mimic has a very similar structure to that of I [backbone rmsd of ~2Å (23)]. Second, order parameters for I derived from chemical shifts are comparable to those of N for helices H1 and H2 and a part of helix H3 (24). This suggests that the intermediate is not a dynamic average of several structures, but may be faithfully represented by a single conformation.

Probing the Potential Contribution of a Third State. Fersht and coworkers (25) have shown from stopped-flow fluorescence experiments that the FF domain folds via an on-pathway intermediate, I, connecting native and unfolded (U) states. Under the conditions of the NMR experiments reported here and elsewhere (17, 24) the U state for the WT domain is present at a sufficiently low fractional population so as to not appreciably influence the dispersion profiles. Our data can therefore be well analyzed on the basis of a two-state exchange model (reduced $\chi^2 = 0.8$). A different scenario was observed for both the A17G and Q19G mutants of the FF domain (17). Here, the dispersion data could only be fit properly by including an additional state that was subsequently established to be U on the basis of the random coil chemical shifts obtained from analysis of relaxation dispersion profiles. Reduced χ^2 values decreased from 2.7 and 1.4 for A17G and Q19G, respectively, when a two-state analysis was performed, to less than 1 when a three-state model was used. The minimum population of the U state was determined to be 0.2% in these fits, with exchange rates, k_{ex}, between U and I ranging from 4,000 to 20,000 s^{-1}, depending on the temperature. Thus, at least in this case with the exchange parameters listed above the presence of a third state could be detected.

To get insight into whether the U state could influence the results of the present study we have considered a series of simulations, assuming a three-state exchange model, $U \leftrightarrow I \leftrightarrow N$. Hydrogen–deuterium exchange data recorded on the WT FF domain provide an upper estimate of 0.2% for the fractional population of U and we have used 0.1% in what follows. Values for k_{ex}(I–N) and p_I at each viscogen concentration (glycerol and BSA) have been taken from two-state fits of experimental dispersion data (16), with k_{ex}(U–I) assumed to be 10,000 s^{-1} and scaled with viscosity as for k_{ex}(I–N) such that the ratio k_{ex}(U–I)/k_{ex}(I–N) is constant. Finally, $\Delta\varpi_{IN}$ values used in simulations were those obtained from fits of CPMG dispersion curves to a two-state model in the absence of viscogen, whereas $\Delta\varpi_{UN}$ values were as reported for the Q19G mutant (17) because the corresponding shift differences for the WT FF domain have not been measured. Dispersion profiles were generated by numerically solving the Bloch–McConnell equations describing the exchange process (21) and Gaussian noise (SD corresponding to 2% of each $R_{2,eff}$ value) was added to each $R_{2,eff}$ rate. In this manner a set of dispersion profiles was created with the same number of "residues" as used in fits of the experimental data. A different dataset was generated for each input k_{ex}(I–N), k_{ex}(U–I), corresponding to different viscogen concentrations that were used experimentally. A second set of simulations was also performed in a manner similar to that described above but using a model of two-site exchange.

Each of the datasets was subsequently fit using a two-site exchange model following an identical approach to that used in the analysis of the experimental data. Values of k_{ex} obtained from analysis of two- and three-state data for different viscogen concentrations are plotted in Fig. S7*A*. The correlation between the two values is high with minimal systematic offsets. Thus, for the exchange parameters considered here, only small errors in k_{ex} are introduced when fitting three-site exchange data to the simple two-site model. Next, viscosity-dependent k_{ex} values obtained by fitting the two- and three-state exchange data to a model of two-site chemical exchange were analyzed via the Kramers equation, $\tau_i = (\sigma + \eta)/K_i$, i ε $\{I,N\}$, $\tau_I = 1/k_{IN}$, $\tau_N = 1/k_{NI}$ to extract internal friction values (Fig. S7 *B–E*). The σ values so obtained from analysis of the two- and three-state datasets are very similar. Using K_i and σ values extracted from fits of the τ_i vs. η profiles from glycerol (two-state, Fig. S7*D*) we then evaluated the extent to which $\eta_{probe}*/\eta_{buff}$ values are affected by the presence of the small population of the U state. Values of $\eta_{probe}*/\eta_{buff}$ were obtained as described in the text with the exception that the simulated data of Fig. S7 *C* and *E* were used. Briefly, k_{IN} and k_{NI} that were generated from simulated two-state and three-state exchange data, 200 mg/mL BSA (Fig. S7 *C* and *E*), were used along with K_i and σ values (see above) to calculate $\eta_{probe}{*}/\eta_{buff} = 1.28 \pm 0.10$ (two-state exchange) and 1.22 ± 0.10 (three-state exchange). These small differences have little effect on calculated EHR values for the I–N transition (Fig. 4).

In summary, the two-state exchange model used to fit experimental CPMG dispersion profiles is the simplest that can account for our data. As with all fitting, it is not possible to prove that the chosen model is correct. However, the three-state interconversion considered here that includes a low-populated U state leads to very small changes in $\eta_{probe}{*}/\eta_{buff}$ that do not affect any of the conclusions of the present work.

1. Ecker DJ, et al. (1987) Gene synthesis, expression, structures, and functional activities of site-specific mutants of ubiquitin. *J Biol Chem* 262(29):14213–14221.
2. Johnson CS, Jr. (1999) Diffusion ordered nuclear magnetic resonance spectroscopy: Principles and applications. *Prog Nucl Magn Reson Spectrosc* 34(3):203–256.
3. Cavanagh J, Fairbrother WJ, Palmer AG, Rance M, Skelton NJ (2007) *Protein NMR Spectroscopy: Principles and Practice* (Academic, New York), 2nd Ed.
4. Altieri AS, Hinton DP, Byrd RA (1995) Association of biomolecular systems via pulsed field gradient NMR self-diffusion measurements. *J Am Chem Soc* 117(28):7566–7567.
5. Williamson JR, Boxer SG (1989) Multinuclear NMR studies of DNA hairpins. 1. Structure and dynamics of d(CGCGTTGTTCGCG). *Biochemistry* 28(7):2819–2831.
6. Farrow NA, et al. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. *Biochemistry* 33(19): 5984–6003.
7. Tjandra N, Feller SE, Pastor RW, Bax A (1995) Rotational diffusion anisotropy of human ubiquitin from 15N NMR relaxation. *J Am Chem Soc* 117(50):12562–12566.
8. Muhandiram D, Yamazaki T, Sykes BD, Kay LE (1995) Measurement of ^{2}H T$_1$ and T$_{1\rho}$ Relaxation times in uniformly ^{13}C-labeled and fractionally ^{2}H-labeled proteins in solution. *J Am Chem Soc* 117(46):11536–11544.
9. Lipari G, Szabo A (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J Am Chem Soc* 104(17):4546–4559.
10. Herzfeld J, Roberts JE, Griffin RG (1987) Sideband intensities in two-dimensional NMR spectra of rotating solids. *J Chem Phys* 86(2):597–602.
11. Buck M, et al. (1995) Structural determinants of protein dynamics: Analysis of 15N NMR relaxation measurements for main-chain and side-chain nuclei of hen egg white lysozyme. *Biochemistry* 34(12):4041–4055.
12. Trbovic N, et al. (2009) Protein side-chain dynamics and residual conformational entropy. *J Am Chem Soc* 131(2):615–622.
13. Clore GM, et al. (1990) Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. *J Am Chem Soc* 112(12):4989–4991.
14. Orekhov VY, Nolde D, Golovanov A, Korzhnev D, Arseniev A (1995) Processing of heteronuclear NMR relaxation data with the new software DASHA. *Appl Magn Reson* 9(4):581–588.
15. Lee LK, Rance M, Chazin WJ, Palmer AG, 3rd (1997) Rotational diffusion anisotropy of proteins from simultaneous analysis of 15N and 13C α nuclear spin relaxation. *J Biomol NMR* 9(3):287–298.
16. Sekhar A, Vallurupalli P, Kay LE (2012) Folding of the four-helix bundle FF domain from a compact on-pathway intermediate state is governed predominantly by water motion. *Proc Natl Acad Sci USA* 109(47):19268–19273.
17. Korzhnev DM, Religa TL, Lundström P, Fersht AR, Kay LE (2007) The folding pathway of an FF domain: Characterization of an on-pathway intermediate state under folding conditions by $^{(15)}$N, $^{(13)}$C$^{(\alpha)}$ and $^{(13)}$C-methyl relaxation dispersion and $^{(1)}$H/$^{(2)}$H-exchange NMR spectroscopy. *J Mol Biol* 372(2):497–512.
18. Kramers HA (1940) Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* 7(4):284–304.
19. Ansari A, Jones CM, Henry ER, Hofrichter J, Eaton WA (1992) The role of solvent viscosity in the dynamics of protein conformational changes. *Science* 256(5065): 1796–1798.
20. Efron B, Tibshirani R (1986) Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Stat Sci* 1(1):54–75.
21. McConnell HM (1958) Reaction rates by nuclear magnetic resonance. *J Chem Phys* 28(3):430–431.
22. Korzhnev DM, et al. (2004) Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR. *Nature* 430(6999):586–590.
23. Barette J, Velyvis A, Religa TL, Korzhnev DM, Kay LE (2012) Cross-validation of the structure of a transiently formed and low populated FF domain folding intermediate determined by relaxation dispersion NMR and CS-Rosetta. *J Phys Chem B* 116(23): 6637–6644.
24. Korzhnev DM, Religa TL, Banachewicz W, Fersht AR, Kay LE (2010) A transient and low-populated protein-folding intermediate at atomic resolution. *Science* 329(5997): 1312–1316.
25. Jemth P, et al. (2004) Demonstration of a low-energy on-pathway intermediate in a fast-folding protein by kinetics, protein engineering, and simulation. *Proc Natl Acad Sci USA* 101(17):6450–6455.

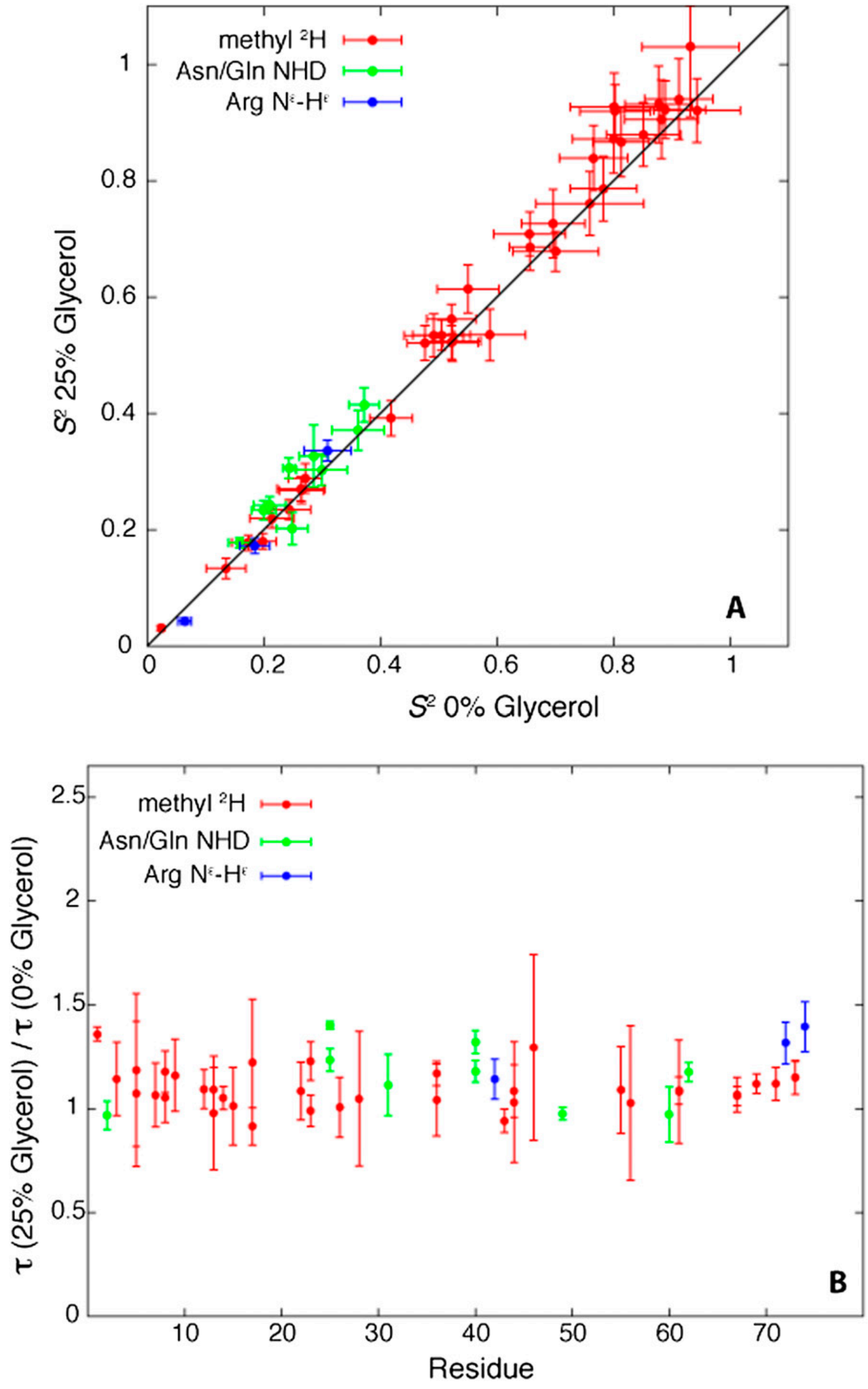

Fig. S1. Solvent viscosity has a very small effect on protein side-chain picosecond–nanosecond (ps–ns) timescale dynamics. (*A*) Comparison of ubiquitin side-chain S^2 values obtained from analysis of $^{13}CH_2D$ ^{2}H (methyl ^{2}H, red) and ^{15}N (NHD, Asn/Gln, green; Nᵋ–Hᵋ Arg, blue) relaxation data recorded in aqueous buffer (0% glycerol) and buffer that includes 25% glycerol (increase in macroscopic viscosity of ∼2.5). Values of S^2_{axis} and $S^2 = S^2_s \times S^2_f$ are plotted for methyl groups and side-chain NHs, respectively. (*B*) Ratio of τ_f (methyl) and τ_S(side-chain ^{15}N) values obtained from analysis of data recorded with and without glycerol. Values of τ increase on average by a factor of ∼1.3 despite the 2.5-fold increase in solvent viscosity indicating that side-chain motions are only slightly influenced by the increased solvent viscosity.

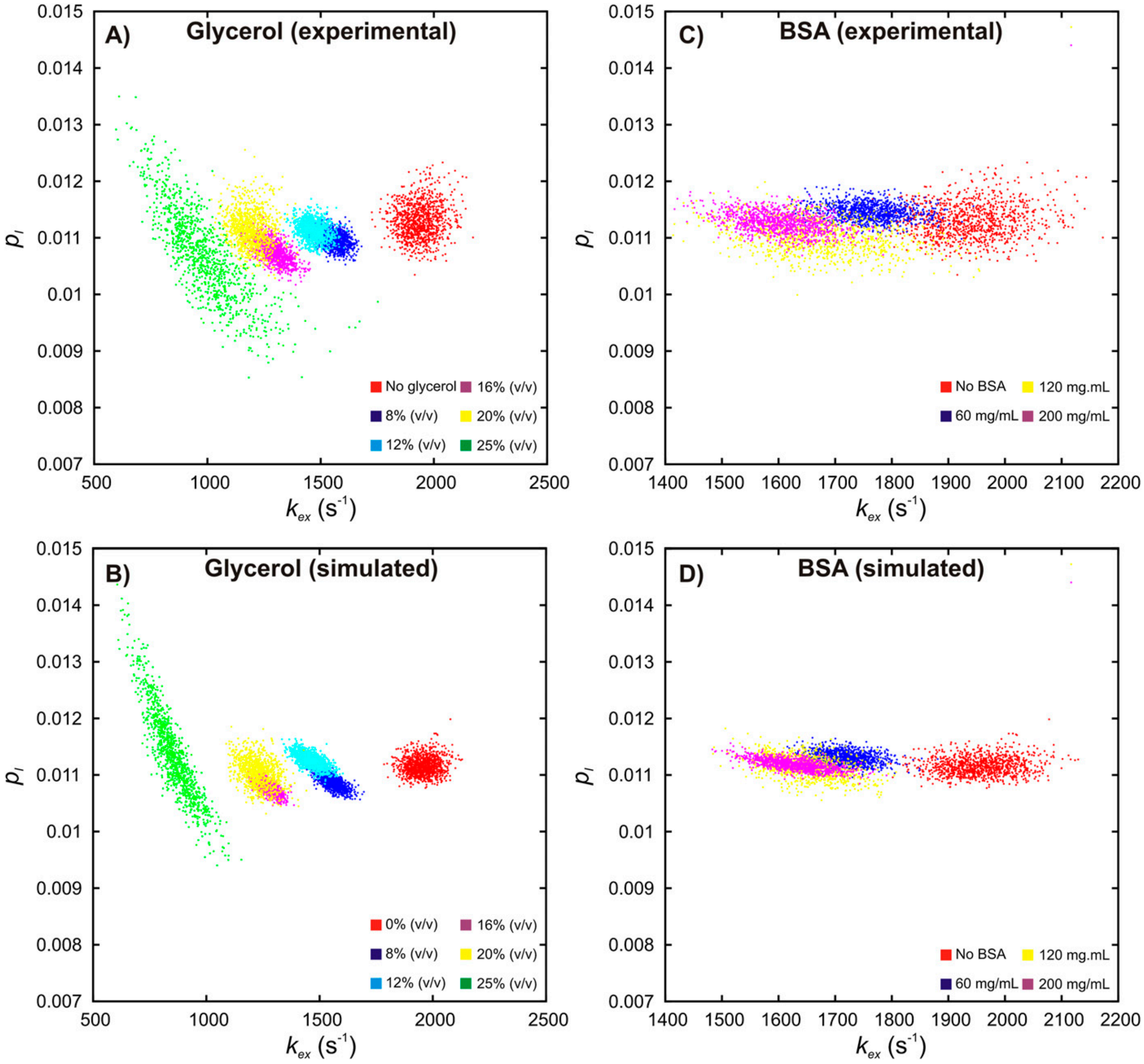

Fig. S2. Distributions of (p_I, k_{ex}) for various (*A* and *B*) glycerol and (*C* and *D*) BSA concentrations. Distributions were obtained by a bootstrapping analysis involving (*A* and *C*) experimental or (*B* and *D*) simulated data and fitting the resulting datasets globally to a model of two-state conformational exchange. Ideal two-state simulated data (*B* and *D*) were obtained as described in *SI Text*.

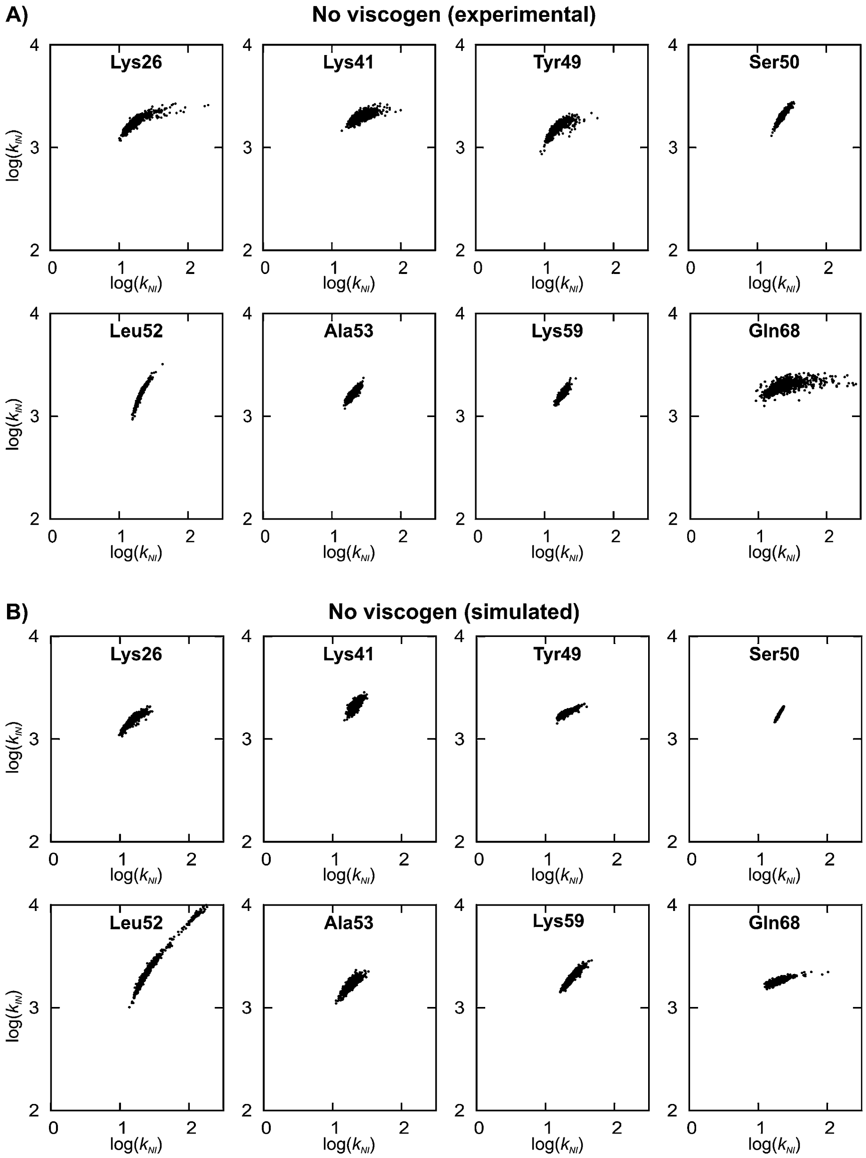

Fig. S3. Distributions of k_{IN} [= $(1 - p_I)*k_{ex}$] and k_{NI} (= p_I*k_{ex}) obtained in the absence of viscogen. Distributions were obtained by a bootstrapping analysis of (*A*) experimental or (*B*) simulated data and fitting the bootstrapped datasets on a per-residue basis to a model of two-state conformational exchange. Ideal two-state simulated data for *B* was obtained as described in *SI Text*.

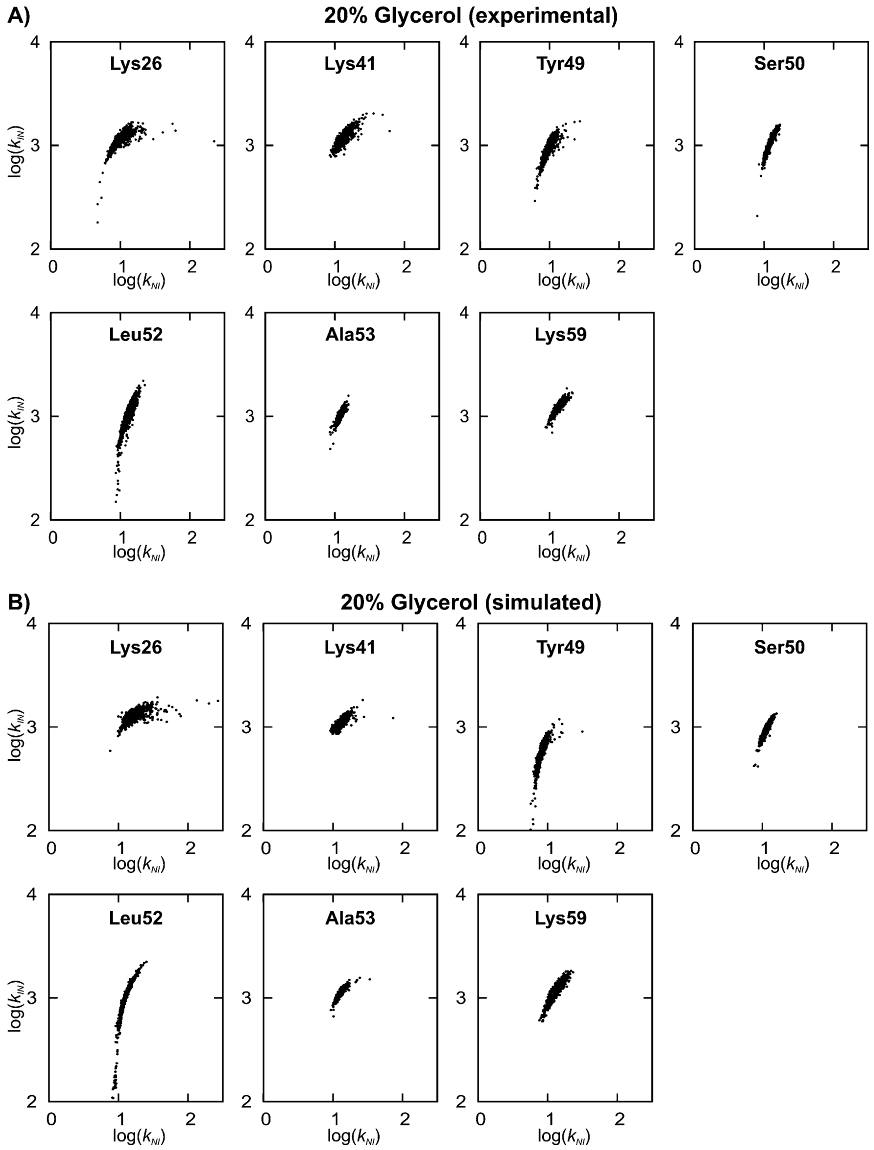

Fig. S4. Distributions of k_{IN} [= (1 − p_I)*k_{ex}] and k_{NI} (= p_I*k_{ex}) obtained in the presence of 20% (vol/vol) glycerol. Distributions were obtained by a bootstrapping analysis of (*A*) experimental or (*B*) simulated data and fitting the bootstrapped datasets on a per-residue basis to a model of two-state conformational exchange. Ideal two-state simulated data for *B* was obtained as described in *SI Text*. The seven residues chosen are those for which experimental dispersion data are available over all glycerol concentrations examined.

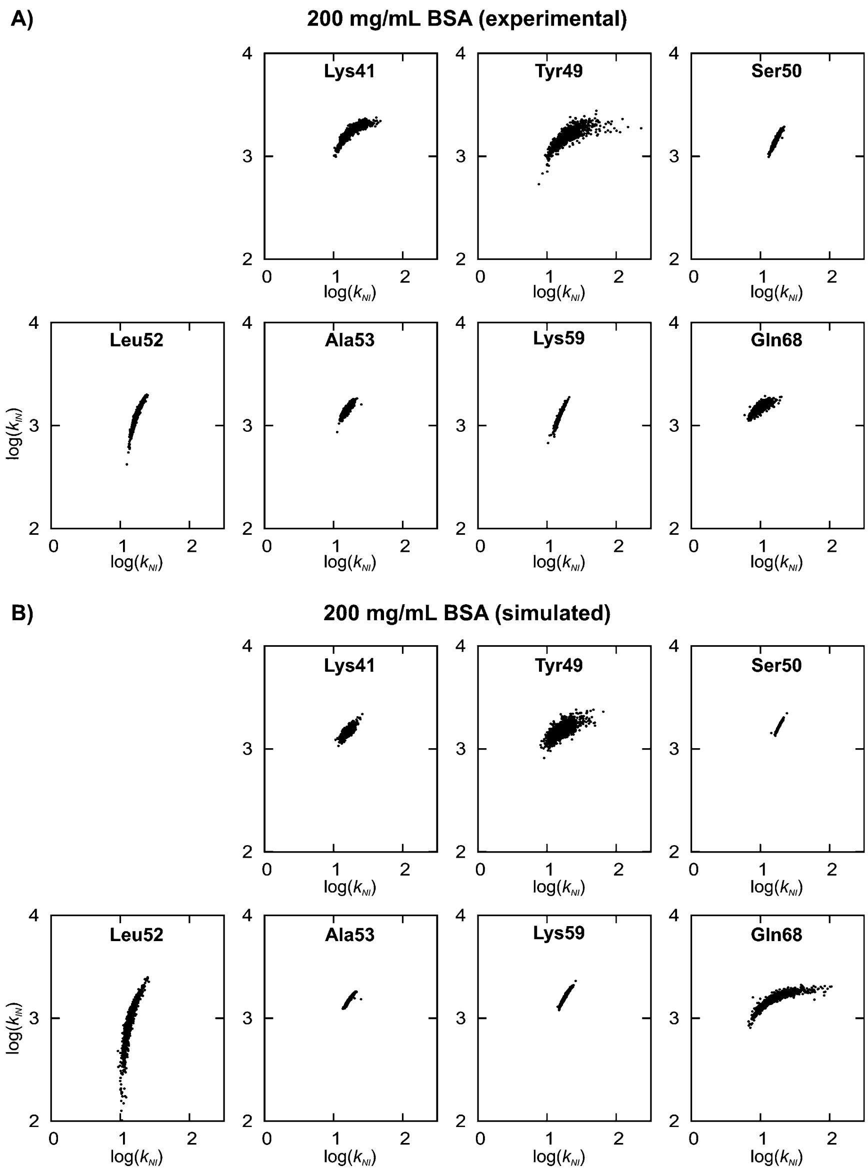

Fig. S5. Distributions of k_{IN} [= $(1 - p_I)*k_{ex}$] and k_{NI} (= p_I*k_{ex}) obtained in the presence of 200 mg/mL BSA. Distributions were obtained by a bootstrapping analysis of (*A*) experimental or (*B*) simulated data and fitting the bootstrapped datasets on a per-residue basis to a model of two-state conformational exchange. Ideal two-state simulated data for *B* was obtained as described in *SI Text*. The seven residues chosen are those for which experimental dispersion data are available over all BSA concentrations examined.

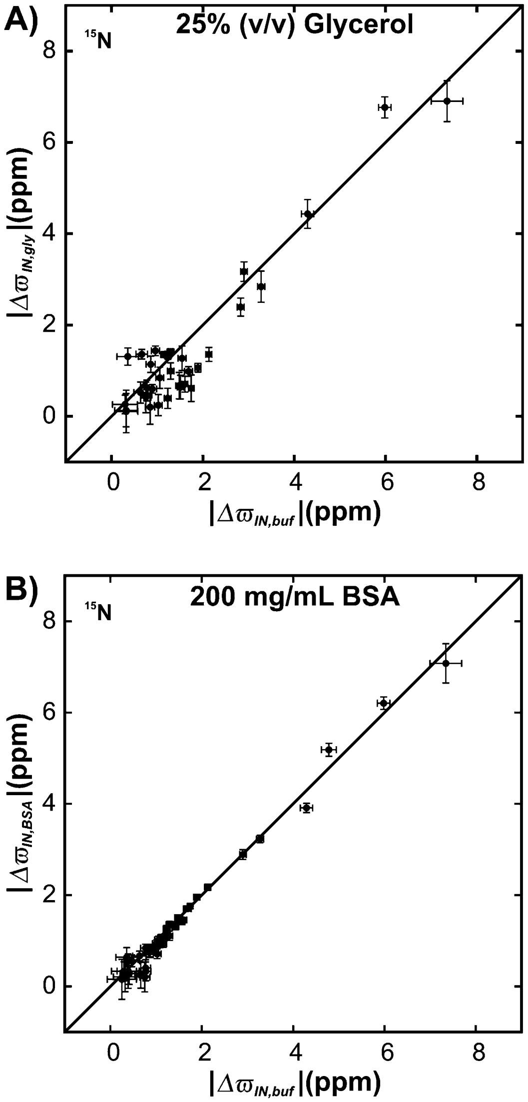

Fig. S6. Correlation of ^{15}N $\Delta\varpi_{IN}$ values obtained from global fits of relaxation dispersion curves acquired at 11.7 and 18.8 T and measured using samples dissolved in buffer (buf) and in (*A*) 25% (vol/vol) glycerol (gly) or (*B*) 200 mg/mL BSA. A two-state model of conformational exchange was used in all analyses. The solid line is $y = x$. Data for this figure is from Sekhar et al. (16).

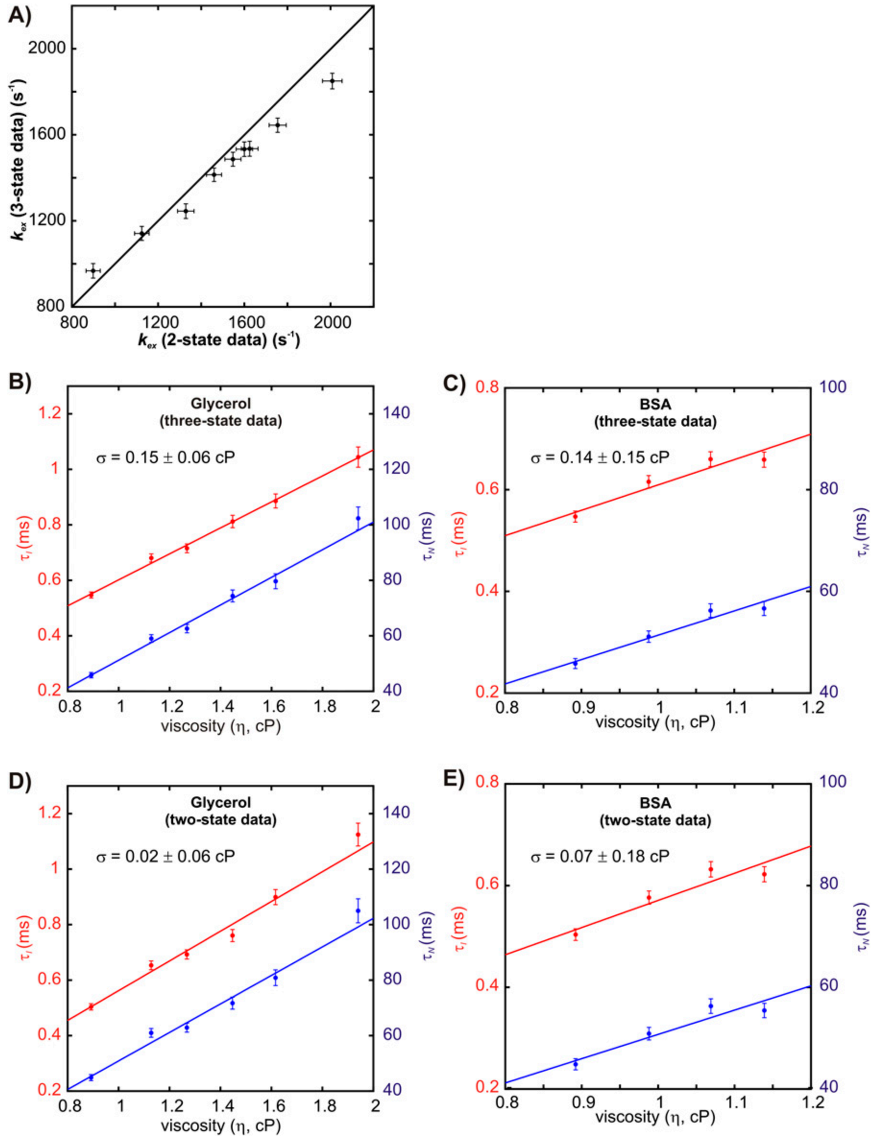

Fig. S7. (*A*) Correlation plot of k_{ex} values obtained from fitting simulated two-state and three-state data to a two-state model of conformational exchange. Simulated data were constructed as described in *SI Text*. The solid line is $y = x$. Viscosity-dependent lifetimes of the intermediate ($\tau_I = 1/k_{IN}$) and native states ($\tau_N = 1/k_{NI}$) in (*B* and *D*) glycerol and (*C* and *E*) BSA obtained by fitting simulated three-state (*B* and *C*) and two-state (*D* and *E*) relaxation dispersion data to a two-state model of chemical exchange. Solid lines are fits of the data to a Kramers equation of the form $\tau_i = (\sigma + \eta)/K_i$, i ε {I,N}.

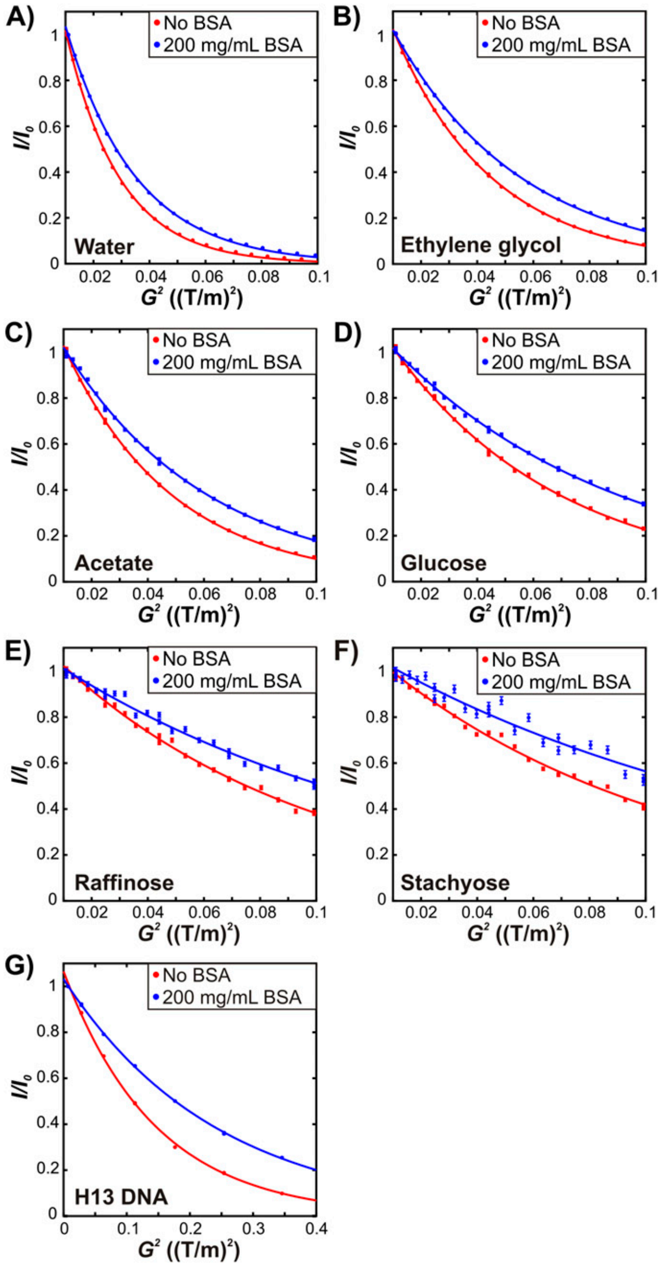

Fig. S8. Intensity decays observed in PFG NMR diffusion spectra as a function of the square of the gradient strength in the absence (red) and presence (blue) of 200 mg/mL BSA for (*A*) water, (*B*) ethylene glycol, (*C*) acetate, (*D*) glucose, (*E*) raffinose, (*F*) stachyose, and (*G*) H13 DNA (5). The intensities are normalized to the value at the first gradient strength (I_0). Errors in the data (vertical bars) were estimated by acquiring duplicate points at selected gradient strengths and, if not visible, are smaller than the circles denoting the data. Solid lines are fits of the data points to a single exponential decay curve (Eq. **3**). A wider range of gradient strengths was used for H13 DNA because of its larger size and hence slower diffusion than that of the other probe molecules.

Table S1. R_H **values calculated for each probe from PFG measurements (Eq. 5)**

Probe	R_H,[†] Å	No. of resonances used[‡]
Water	1.12 ± 0.01	1
Acetate	2.24 ± 0.01	1
Ethylene glycol	2.05 ± 0.01	1
Dioxane	2.12 ± 0.02	1
Glycerol	2.58 ± 0.01	1
Glucose	3.5 ± 0.2	2
Sucrose	4.50 ± 0.08	4
Raffinose	5.3 ± 0.3	5
Stachyose	5.9 ± 0.1	4
H13 DNA	8.9 ± 0.5	4
FF domain	13.0 ± 0.3	10

[†] R_H values match well with measurements made in literature where available (1–5).

[‡] Diffusion constants used in the calculation of R_H were based on averages over the number of peaks indicated.

1. Williamson JR, Boxer SG (1989) Multinuclear NMR studies of DNA hairpins. 1. Structure and dynamics of d(CGCGTTGTTCGCG). *Biochemistry* 28(7):2819–2831.
2. Chatterjee C, Martinez D, Gerig JT (2007) Interactions of trifluoroethanol with [val5]angiotensin II. *J Phys Chem B* 111(31):9355–9362.
3. Hong SU, Malaisamy R, Bruening ML (2007) Separation of fluoride from other monovalent anions using multilayer polyelectrolyte nanofiltration membranes. *Langmuir* 23(4): 1716–1722.
4. Ribeiro AC, et al. (2006) Binary mutual diffusion coefficients of aqueous solutions of sucrose, lactose, glucose, and fructose in the temperature range from (298.15 to 328.15) K. *J Chem Eng Data* 51(5):1836–1840.
5. Schultz SG, Solomon AK (1961) Determination of the effective hydrodynamic radii of small molecules by viscometry. *J Gen Physiol* 44(6):1189–1199.