www.sciencemag.org/content/358/6360/238/suppl/DC1

Supplementary Materials for

Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water

Joshua A. Riback, Micayla A. Bowman, Adam M. Zmyslowski, Catherine R. Knoverek, John M. Jumper, James R. Hinshaw, Emily B. Kaye, Karl F. Freed, Patricia L. Clark,* Tobin R. Sosnick*

*Corresponding author. Email: pclark1@nd.edu (P.L.C.); trsosnic@uchicago.edu (T.R.S.)

Published 13 October 2017, *Science* **358**, 238 (2017)
DOI: 10.1126/science.aan5774

This PDF file includes:

Other supplementary material for this manuscript includes the following:

Movies S1 and S2

Supplemental Materials:

Protein expression and purification.

PNt - A PNt coding sequence optimized for expression in *E. coli* was purchased commercially (Invitrogen), cloned into plasmid pET21b and overexpressed in *E. coli* BL21(DE3)pLysS. PNt was isolated from insoluble inclusion bodies using procedures described previously (*20, 21*) and purified as follows. Inclusion bodies were isolated from cell lysates and solubilized in 6 M Gdn. Denaturant was removed first by dilution to 1.2 M Gdn followed by dialysis against 50 mM Tris pH 7.5 to a final concentration of Gdn <1 mM. The soluble fraction after dialysis was purified by strong anion exchange chromatography (Source 15Q, GE Healthcare) using a linear gradient from 0 to 150 mM NaCl. Fractions containing PNt as confirmed by SDS-PAGE were pooled and supplemented with 5 mM EDTA. Pooled fractions were concentrated and purified further using a HiLoad 10/60 Superdex 200 preparative grade column (GE Healthcare) equilibrated with 50 mM Tris pH 7.5, 5 mM EDTA. Purified PNt (100-200 uM) was stored at 4ºC in 50mM Tris pH 7.5, 5 mM EDTA.

FhuA - The FhuA plug coding sequence (residues 19-161 from pdb:1BY3) from *E. coli* was purchased commercially (Invitrogen) and cloned into plasmid pET21b containing an N-terminal His8-Protein G-TEV sequence tag for expression and purification (*27*). Cells were grown to an optical density of 0.8 at 600 nm in 1 L of LB/ampicillin medium in baffled flasks shaking at 200 RPM at 37 C. Expression of FhuA was induced with 1mM IPTG at 37C for 4h, after which cells were harvested at 5000 g for 15 minutes before storage of cell pellets at -80C. Cell pellets were solubilized in lysis buffer (8 M urea 300 mM NaCl 50 mM TrisHCl pH 8.0) and lysed by sonication. Solubilized inclusion bodies were applied to 10mL of high density nickel iminodiacetic acid (IDA) beads pre-equilibrated with lysis buffer and washed with 5 column volumes (CV) each of lysis buffer then wash buffer (300 mM NaCl 20 mM imidazole 50 mM NaPi pH 8.0). Protein was eluted in 5 mL fractions using elution buffer (300 mM NaCl 250 mM imidazole 50 mM NaPi pH 8.0). Protein containing fractions (measured by absorbance at 280 nm) were pooled, 1 mg of TEV protease in 1 mL of TEV protease buffer (50% glycerol 1 mM EDTA 5 mM DTT 0.01% Triton X-100 50 mM TrisHCl pH 8.0) was added, and the mixture was dialyzed against 2 L of TEV cleavage buffer (1 mM EDTA 50 mM TrisHCl pH 8.0) at room temperature for 12+ hours (overnight). Dialyzed and cleaved protein was passed through the same nickel beads (equilibrated with 5 CV wash buffer) to remove the cleaved His8-DesG tags. Purified and cleaved, FhuA plug was dialyzed against 20 mM $NaHCO_3$ overnight and lyophilized for long term storage.

RNaseA- Protein was prepared as previously described (*14, 23*). Briefly, protein powder was solubilized in DTT and denaturant (>5M Gdn) and equilibrated at room temperature for an hour prior to injection onto the SEC column.

Far-UV CD.

Far-UV CD spectra were collected as previously described (*29*) using a Jasco J-815 spectropolarimeter. Purified PNt was diluted to 5 uM in 25 mM sodium phosphate pH 7.5. Cosolutes were added as indicated.

NMR.

NMR spectra were obtained using 500 and 600 MHz spectrometers with Bruker AVANCE III consoles. HSQC spectra were run under non-denaturing conditions.

SAXS.

Data were taken at the BioCAT beamline at the Advanced Photon Source (Argonne National Lab) using a GE Lifesciences Superdex 200 SEC column with the scattering presented as $q=4\pi \sin\theta/\lambda$, where 2θ is the scattering angle and λ is the X-ray wavelength (1 Å).

Code for simulations and fitting scattering data.

Code and associated files necessary to produce simulations and the analysis can be accessed at https://github.com/SuperScienceJew/SAXSonIDPs. Additionally, our webserver http://sosnick.uchicago.edu/SAXSonIDPs is available for fitting SAXS data with our MFF(ν,R_g).

Simulations and generation of MFF. Calculations were performed at the University of Chicago Research Computing Center using a version our *Upside* molecular dynamics program (*30*) modified for Cβ-level interactions. Each of the five simulations started from different random conformations and run for 11,000 time steps, saving a conformation every 20 time steps. Each simulation was run under a Hamiltonian replica exchange protocol with 30 replicas having different Cβ-Cβ interaction strengths with Monte Carlo (ϕ,ψ) pivot moves. The first 1000 time steps were discarded for equilibration. For the selected 500 individual conformations, side chains were added with TreePack (*31*) and scattering profiles were determined with the FOXS program (*32*) with the command "foxs -s 80 -q 0.4 --excluded_volume 1.0 --water_layer_c2 2.0" accounting for an expected hydration layer of ~0.03 electron/Å^3 above bulk water. Explicit solvent simulations for fixed PNt conformations (*33*) yielded similar scattering curves (**Fig. S6**). For each simulation, an R_g is calculated based on the Cα atom positions. For each Cβ-Cβ interaction strength, the scattering profiles were made dimensionless (transformed into $I(qR_g)/I_0$ with interpolation), and averaged for the five independent simulations and their standard error was determined. To calculate ν for each interaction strength, intra-chain RMS distances between Cα residues were calculated ($|i-j|>15$) and fit to $\langle R_{|i-j|}^2\rangle^{0.5} = R_0|i-j|^{\nu}$ weighted by the errors from the five replicates in Mathematica (**Fig. 3B, lines**). Across the 30 interaction strengths investigated, ν =0.35-0.6. The dependence on ν of the (averaged) dimensionless scattering profile for all the Cβ-Cβ interaction strengths was smoothed with splines in Mathematica, yielding a continuous function, the molecular form factor MFF(qR_g,ν).

Fitting the MFF. Data were fit to I_0*MFF(qR_g,ν) using NonlinearModelFit in Mathematica over the range $0.01\le q\le 0.15$Å^{-1} except for 4 and 8 M Gdn, which were fit over $0.01\le q\le 0.1$Å^{-1} and $0.01\le q\le 0.075$Å^{-1}, respectively. When fitting MFFs to simulations, simulations were interpolated over $0.01\le q\le 0.15$Å^{-1} with q^{-1} weighting, to mimic the q-dependence of the experimental errors.

Scattering for a SARW from renormalization theory. Calculations utilized the analytical form (Eq. 4.2 in reference (*34*)) coded into Mathematica using numerical integration routine for the conversion from β to q using the Series approximation (original form, $\beta*(1-13/96)\sim \beta* 0.864583=q^2R_g^2$; revised form, $\beta* 0.864453=q^2R_g^2$).

Bioinformatics. The non-redundant dataset of folded proteins from the PDB was curated with the PISCES (*35*) server using a percentage identity cutoff of 25%, a resolution cutoff of 3.0 Å, the R-factor cutoff of 0.3, and selecting chain lengths >40 residues. The reported hydrophobicity is presented as a Z-score according to the Hopp-Woods scale (*36*). The fraction of charge residues was calculated using the number of Lys, Arg, Asp, and Glu residues. The

boundary of region predicted to be a poor solvent(*7*) (hydrophobicity value above -0.2, **Fig 1C**) was obtained by conversion to the Hopp-Woods scale using linear regression between the two scales from proteins in the curated set. The boundary value from simulation was a fraction total charge below 0.25 (*26*).

Author Contributions: Author contributions: JAR, MAB, JH, PLC, TRS designed research; JAR, MAB, AZ, CRK, JH, EBK, PLC, TRS performed research; JAR, MAB, CRK, EBK, JJ, KFF, PLC, TRS contributed reagents/analytic tools; JAR, MAB, CRK, JH, PLC, TRS analyzed data; and JAR, MAB, PLC, TRS, wrote the paper.

Supplementary Text:

Comparison to existing polymer models. We compared our data to a variety of polymer models (**Fig. S2A**). The Debye random walk (RW) model has the same ν=0.5 value as a chain in a θ solvent. This model often is applied to describe unfolded chains even though it represents an unphysical situation where the polypeptide chain can pass through itself (*37*). The swollen Gaussian coil (SGC) model treats self-avoidance by assuming the residue-residue distance distribution is affected by ν: $P(r_{ij}) \propto e^{\wedge}[3r^2_{ij}/(2\langle r^2_{ij}\rangle]$ where $\langle r^2_{ij}\rangle \propto |i-j|^{2\nu}$ (*7, 38, 39*). A third procedure uses conformational space renormalization (RN) to account for excluded volume in a SARW (*34, 40*). As far as we are aware, it has not yet been applied to the scattering of disordered proteins. We also include the Guinier Approximation, where the scattering curve at low q is modeled as a Gaussian, $I(q)=I_o e^{\wedge}(-1/3(qR_g)^2)$. Although this procedure is more appropriate for globular objects, it often is used to describe disordered states using a limited data range $qR_g \leq 1$.

Significant deviations are observed between the experimental data and both the Debye and the SGC models, even at low q ($qR_g \sim 1.5$) (**Fig. S2A**). The scattering in water closely matches the RN model out to $qR_g \sim 3$, where the data exhibits a downturn. In high denaturant, the RN model performs well even out to $qR_g \sim 5$. The close agreement with the RN model supports the common view that chemically denatured proteins are well approximated as a SARW (*1*).

Alternative methods to determine solvent quality. The SGC model described above has the potential to determine ν. However, the model does a poor job of describing the SAXS data (**Fig. S2A).** The only other method we are aware of to obtain ν from a single SAXS measurement involves dimensional analysis. Here, the chain is assumed to be a fractal aggregate with $I(q) \sim q^{-1/Dm}$ at high q and fractal dimension $D_m \sim 1/\nu$ (*14*). To test the validity of this approach, we analyzed the scattering profiles from the simulations as we know their true ν value. We find both that D_m generally is not equal to $1/\nu$, and it often has an error of ± 0.05 units; further, the fitted D_m is very sensitive to the q range (**Fig S2B**). Hence, we do not recommend this approach, especially as our MFF can easily provide ν with an error of only ~0.01.

Nonlinearity in denaturation measurements. Chemical denaturation is often used to measure protein stability, tracking signals such as fluorescence or CD. The free energy difference between the native state and the DSE often is approximated with a linear dependence on denaturant, $\Delta G([den]) = \Delta G^o - m^o[den]$, with the slope proportional to the decrease in accessible surface area (ASA). Such data often is fit with a constant m^o, suggesting that the ΔASA upon folding is unchanged at different denaturant concentrations. However, the observed decrease in ν at lower denaturant indicates that the ASA of the DSE does decrease at lower denaturant. To quantify this effect, we calculated the ASA for our simulated ensembles at different values of ν,

but the effect is small. The ASA decreases by 8% going from water (ν=0.54) to 2 M Gdn (ν=0.59) and 2% from 2 to 4 M Gdn, and the impact on linearity of ΔG with denaturant is minimal (**Fig. S5C**).

Figures:

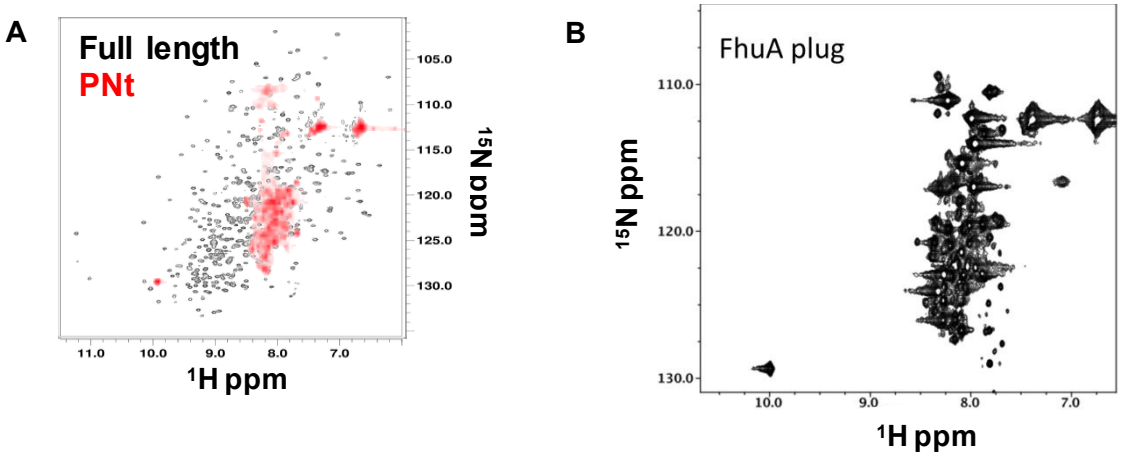


Fig. S1. NMR on disordered proteins. ^{1}H ^{15}N HSQC spectrum of **(A)** PNt and **(B)** FhuA, showing disordered peaks characteristic of unfolded protein.

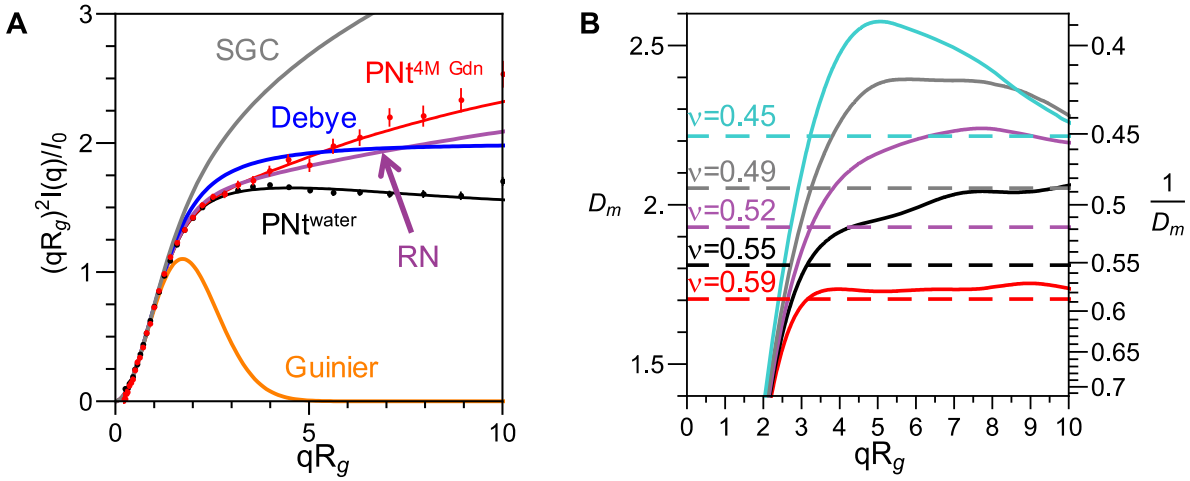


Fig. S2. Comparison to polymer models and alternative methods to extract solvent quality. (**A**) Comparison of PNt data to Swollen Gaussian Coil (SGC) (*39*), Renormalization (RN) (*34*), Debye and Guinier models (*37*). (**B**) Fractal dimension calculated for five simulations shown in **Fig. 3B** as a function of fitted qR_g range [qR_g-1, qR_g+1] (solid lines). The ν value of each simulation is shown as a dotted line at $D_m=1/\nu$. Only for the SARW simulation at $qR_g > 3$ does this method reliably recapitulate the correct ν. See **Supplementary Text** for further discussion.

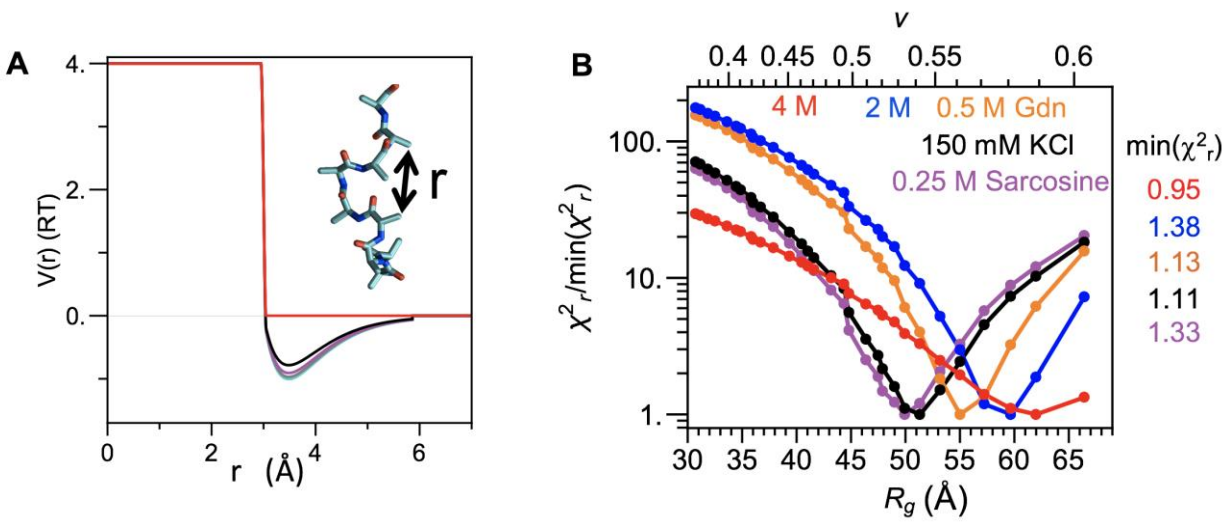


Fig. S3. Simulations potential and match to data. (**A**) Distance dependence for five of the 30 different Cβ-Cβ interaction potentials used in the simulations; one hard sphere and four attractive potentials are shown. (**B**) Error in the fit to the data (colors correspond to experimental data in **Fig 2A**) to each of the 30 simulations obtained with different Cβ-Cβ potentials (represented as points). The y-axis is the normalized error, the ratio of the error of fitting the data to each simulation to the minimum error ($\min(\chi^2_r)$). Minimum error is listed on the right.

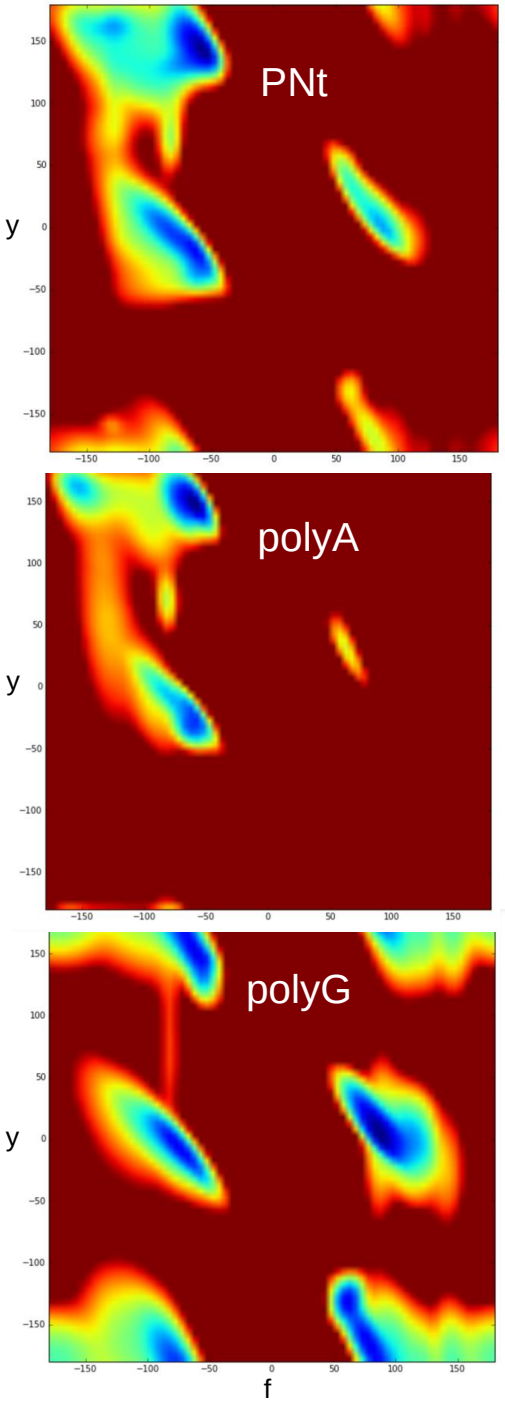


Fig. S4. Ramachandran maps utilized. These maps are used as inputs for backbone dihedral potentials in homopolymer simulations. Plotted on a $\log_e$ scale from most favorable to 5 $\log_e$ units above.

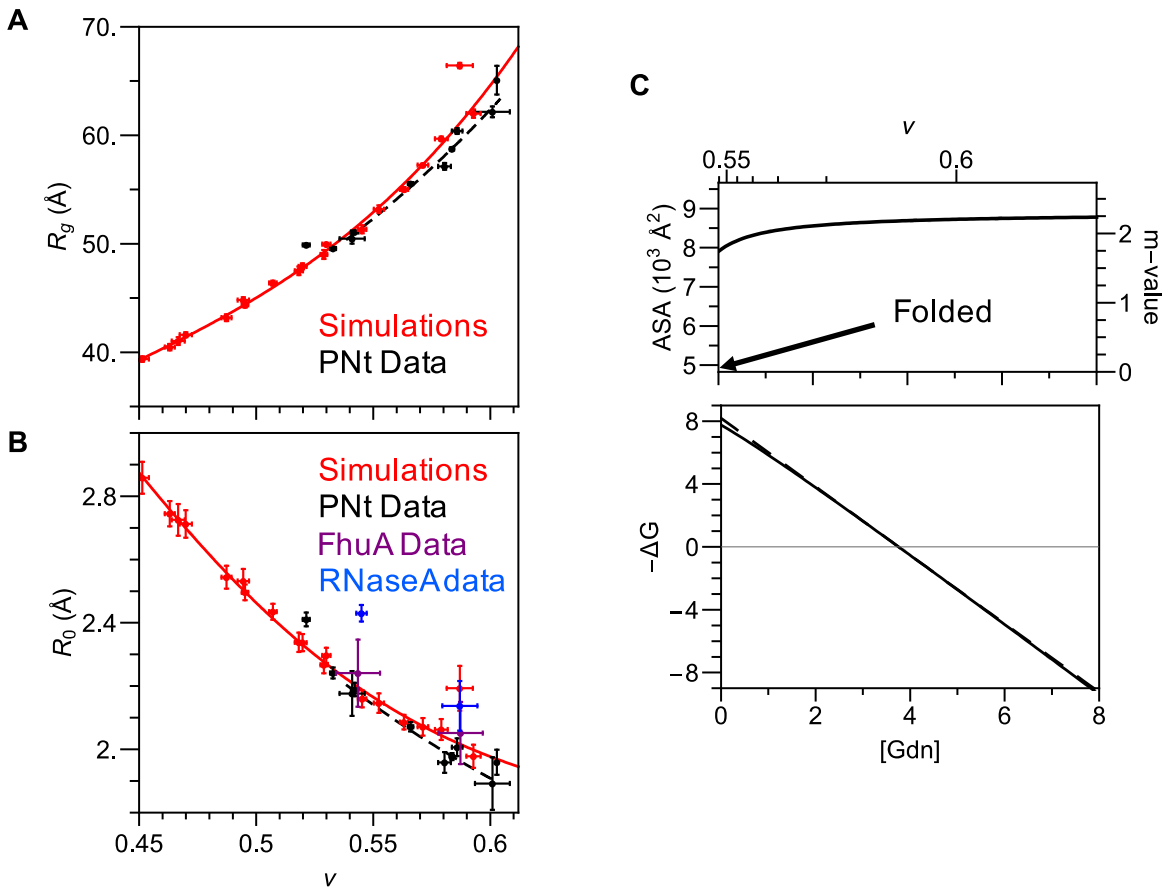


Fig. S5. Extrapolation of simulation to different lengths and ASA values. **A)** R_g and ν pairs from PNt data (black) and from 30 simulations (red) with quartic trend line. **B)** Corresponding plot for the pre-factor $R_0 = R_g N^{-\nu}$, along with data for FhuA and RNase A. **C)** Top, expected increase in the accessible surface area (ASA) and *m*-value of unfolded ubiquitin relative to the folded state. Bottom, expected stability of ubiquitin based on expected ASA decrease and linear extrapolation method. Ubiquitin thermodynamics based on (*12*).

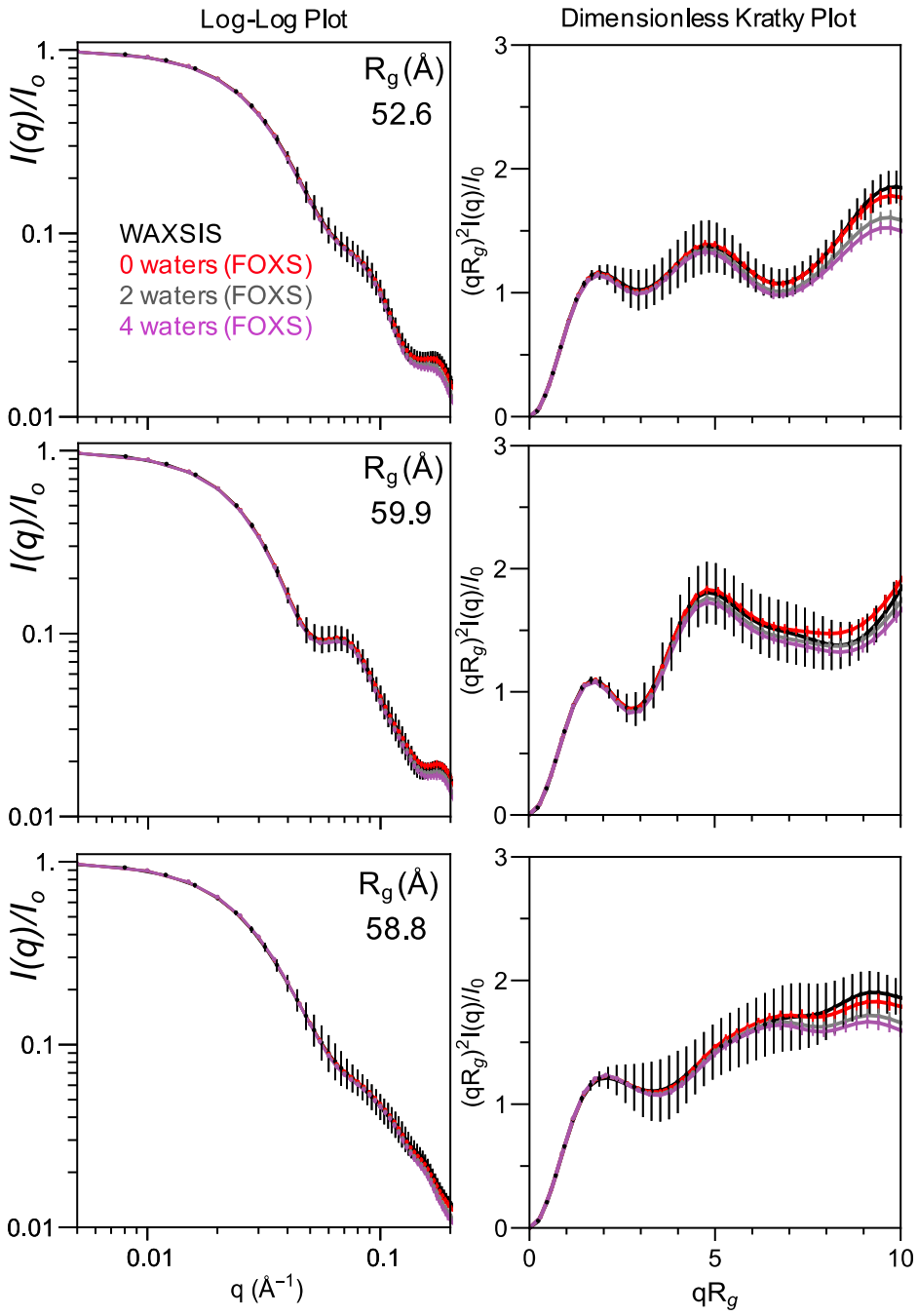


Fig. S6. SAXS profiles are insensitive to the solvent model over pertinent data range. Comparison between the scattering profiles for three PNt SARW conformations determined with explicit solvent from WAXSIS (*33*) and the profiles determined with implicit solvent from FOXS (*32*). The variation of the hydration layer shell parameter in FOXS was yielded only minor changes to scattering profiles. The scattering pattern in each row is for a different conformation.

Tables:

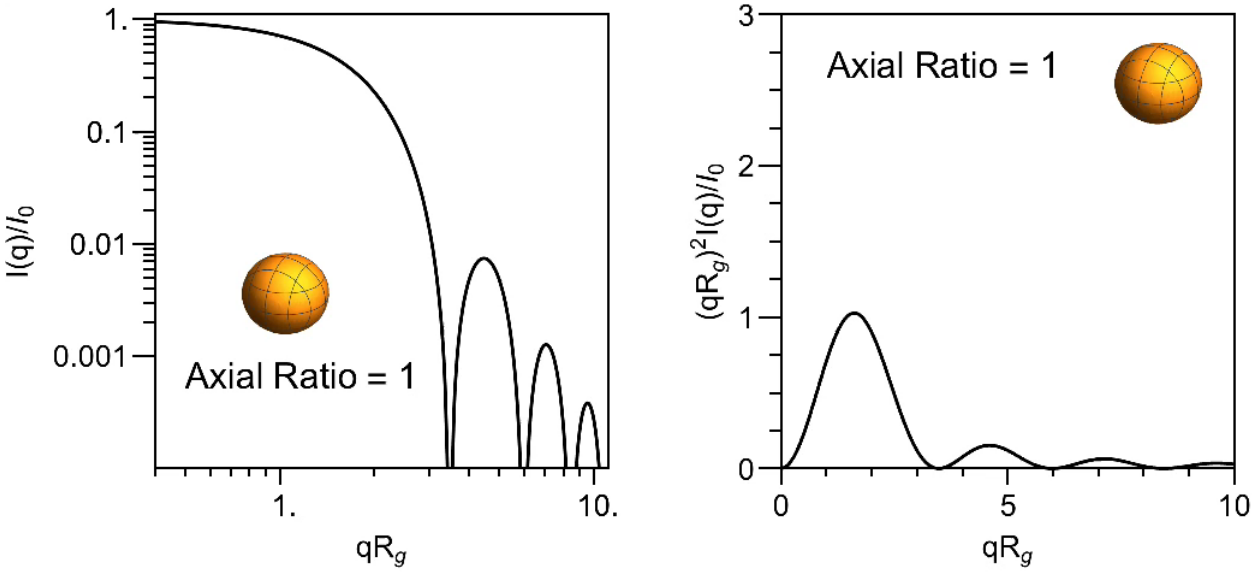
Table S1: Fitting SAXS patterns of ensembles from **Fig. 3** with MFFs. The upper and lower values in each box are the R_g (Å) and ν determined by fitting the associated ensemble to the listed MFF. $MFF_{PNt,334}$ is based on simulations using individual Ramachandran maps for PNt's sequence with uniform interactions at all Cβs, as described in text. The four additional MFFs were generated for 334 residue chains having either i) the average Ramachandran maps over PNt's sequence ($MFF_{av,PNt,334}$), ii) Ramachandran maps for poly(Ala) ($MFF_{ala,334}$) or iii) poly(Gly) ($MFF_{PNt,334}$), (**Fig. S4)** and iv) a heteropolymeric version of the PNt sequence where only the hydrophobic residues were attractive (AMVILFYW) ($MFF_{hetero,334}$). To investigate length dependence, we also calculated a sixth MFF for a 144 residues chain, matched to the length of the FhuA plug domain ($MFF_{PNt,334}$).

Ensembles from Fig3	$MFF_{PNt,334}$	$MFF_{av.PNt,334}$	$MFF_{ala,334}$	$MFF_{gly,334}$	$MFF_{hetero,334}$	MFF_{144}
66.4 0.587	66.6 0.598	67.1 0.588	67.0 0.587	66.6 0.595	66.5 0.603	65.7 0.615
53.2 0.552	53.3 0.552	54.0 0.539	53.7 0.532	53.3 0.545	52.3 0.548	52.3 0.570
47.5 0.518	47.4 0.517	48.1 0.507	47.7 0.495	47.4 0.511	45.8 0.509	46.3 0.534
43.2 0.487	42.9 0.485	43.6 0.477	43.2 0.464	43.0 0.480	40.8 0.472	41.8 0.499
39.4 0.451	39.4 0.452	40.1 0.448	39.7 0.436	39.6 0.449	36.8 0.431	39.4 0.451

Table S2: Application of MFFs to fit to data in **Fig 2**. The three numbers in each box are (top to bottom) the R_g (Å), ν, and χ^2_r determined by fitting the experimental data to the listed MFF. Descriptions of MFFs are the same as in **Table S1**.

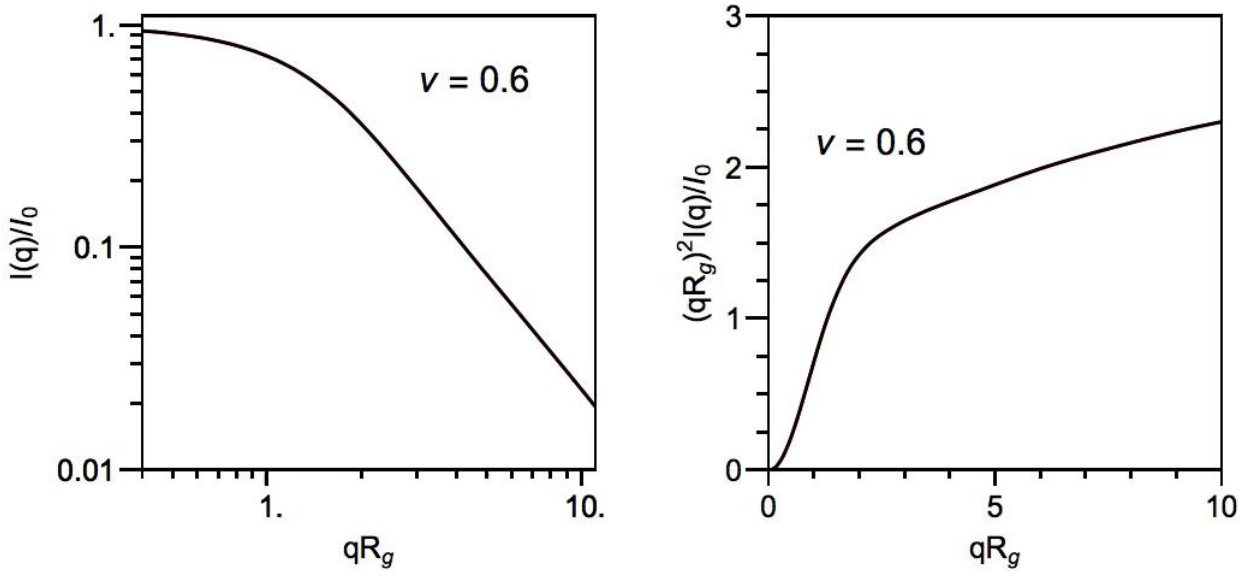
Experiment Condition	$MFF_{PNt,334}$	$MFF_{av.PNt,334}$	$MFF_{ala,334}$	$MFF_{gly,334}$	$MFF_{hetero,334}$	MFF_{144}
250 mM Sarcosine	49.9 0.521 1.12	50.7 0.511 1.11	50.4 0.501 1.18	50.0 0.516 1.20	48.5 0.516 1.31	48.9 0.541 1.33
Water (150mM KCl)	51.1 0.542 1.11	51.9 0.530 1.10	51.6 0.523 1.16	51.2 0.537 1.19	50.1 0.540 1.24	50.2 0.563 1.29
0.5M Gdn	55.5 0.566 1.04	56.3 0.554 1.01	56.1 0.551 1.04	55.7 0.561 1.06	55.0 0.567 1.07	54.8 0.587 1.14
2M Gdn	58.7 0.584 0.986	59.4 0.572 0.985	59.2 0.572 1.01	58.9 0.580 1.03	58.5 0.587 1.01	58.0 0.603 1.14
4M Gdn	62.2 0.601 0.946	62.8 0.596 0.946	62.8 0.594 0.946	62.4 0.603 0.945	62.3 0.608 0.946	61.5 0.618 0.950

Movies:



(Initial picture of movie found as attached file 'MovieS1.avi')

Movie S1: Prolate ellipsoid MFF. Running video changes the axial ratio. For ellipsoid, the R_g held constant as the axial ratio is changed for readers benefit.



(Initial picture of movie found as attached file 'MovieS2.avi')

Movie S2: Disordered protein polymer MFF. Running video changes ν of black curve. Red curve is held at ν=0.6. Note that the x-axis on the left plot is dimensionless and is thus scaled during fitting process.

References

1. J. E. Kohn, I. S. Millett, J. Jacob, B. Zagrovic, T. M. Dillon, N. Cingel, R. S. Dothager, S. Seifert, P. Thiyagarajan, T. R. Sosnick, M. Z. Hasan, V. S. Pande, I. Ruczinski, S. Doniach, K. W. Plaxco, Random-coil behavior and the dimensions of chemically unfolded proteins. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 12491–12496 (2004). doi:10.1073/pnas.0403643101 Medline

2. K. A. Merchant, R. B. Best, J. M. Louis, I. V. Gopich, W. A. Eaton, Characterizing the unfolded states of proteins using single-molecule FRET spectroscopy and molecular simulations. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 1528–1533 (2007). doi:10.1073/pnas.0607097104 Medline

3. G. Ziv, D. Thirumalai, G. Haran, Collapse transition in proteins. *Phys. Chem. Chem. Phys.* **11**, 83–93 (2009). doi:10.1039/B813961J Medline

4. A. Dasgupta, J. B. Udgaonkar, Evidence for initial non-specific polypeptide chain collapse during the refolding of the SH3 domain of PI3 kinase. *J. Mol. Biol.* **403**, 430–445 (2010). doi:10.1016/j.jmb.2010.08.046 Medline

5. G. Haran, How, when and why proteins collapse: The relation to folding. *Curr. Opin. Struct. Biol.* **22**, 14–20 (2012). doi:10.1016/j.sbi.2011.10.005 Medline

6. A. Borgia, W. Zheng, K. Buholzer, M. B. Borgia, A. Schüler, H. Hofmann, A. Soranno, D. Nettels, K. Gast, A. Grishaev, R. B. Best, B. Schuler, consistent view of polypeptide chain expansion in chemical denaturants from multiple experimental methods. *J. Am. Chem. Soc.* **138**, 11714–11726 (2016). doi:10.1021/jacs.6b05917 Medline

7. H. Hofmann, A. Soranno, A. Borgia, K. Gast, D. Nettels, B. Schuler, Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 16155–16160 (2012). doi:10.1073/pnas.1207719109 Medline

8. V. A. Voelz, M. Jäger, S. Yao, Y. Chen, L. Zhu, S. A. Waldauer, G. R. Bowman, M. Friedrichs, O. Bakajin, L. J. Lapidus, S. Weiss, V. S. Pande, Slow unfolded-state structuring in Acyl-CoA binding protein folding revealed by simulation and experiment. *J. Am. Chem. Soc.* **134**, 12565–12577 (2012). doi:10.1021/ja302528z Medline

9. G. Reddy, D. Thirumalai, Collapse precedes folding in denaturant-dependent assembly of ubiquitin. *J. Phys. Chem. B* **121**, 995–1009 (2017). doi:10.1021/acs.jpcb.6b13100 Medline

10. K. W. Plaxco, I. S. Millett, D. J. Segel, S. Doniach, D. Baker, Chain collapse can occur concomitantly with the rate-limiting step in protein folding. *Nat. Struct. Biol.* **6**, 554–556 (1999). doi:10.1038/9329 Medline

11. T. Y. Yoo, S. P. Meisburger, J. Hinshaw, L. Pollack, G. Haran, T. R. Sosnick, K. Plaxco, Small-angle X-ray scattering and single-molecule FRET spectroscopy produce highly divergent views of the low-denaturant unfolded state. *J. Mol. Biol.* **418**, 226–236 (2012). doi:10.1016/j.jmb.2012.01.016 Medline

12. J. Jacob, B. Krantz, R. S. Dothager, P. Thiyagarajan, T. R. Sosnick, Early collapse is not an obligate step in protein folding. *J. Mol. Biol.* **338**, 369–382 (2004). doi:10.1016/j.jmb.2004.02.065 Medline

13. J. J. Skinner, W. Yu, E. K. Gichana, M. C. Baxa, J. R. Hinshaw, K. F. Freed, T. R. Sosnick, Benchmarking all-atom simulations using hydrogen exchange. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 15975–15980 (2014). doi:10.1073/pnas.1404213111 Medline

14. Y. Wang, J. Trewhella, D. P. Goldenberg, Small-angle X-ray scattering of reduced ribonuclease A: Effects of solution conditions and comparisons with a computational model of unfolded proteins. *J. Mol. Biol.* **377**, 1576–1592 (2008). doi:10.1016/j.jmb.2008.02.009 Medline

15. S. V. Kathuria, C. Kayatekin, R. Barrea, E. Kondrashkina, R. Graceffa, L. Guo, R. P. Nobrega, S. Chakravarthy, C. R. Matthews, T. C. Irving, O. Bilsel, Microsecond barrier-limited chain collapse observed by time-resolved FRET and SAXS. *J. Mol. Biol.* **426**, 1980–1994 (2014). doi:10.1016/j.jmb.2014.02.020 Medline

16. J. Jacob, R. S. Dothager, P. Thiyagarajan, T. R. Sosnick, Fully reduced ribonuclease A does not expand at high denaturant concentration or temperature. *J. Mol. Biol.* **367**, 609–615 (2007). doi:10.1016/j.jmb.2007.01.012 Medline

17. T. Konuma, T. Kimura, S. Matsumoto, Y. Goto, T. Fujisawa, A. R. Fersht, S. Takahashi, Time-resolved small-angle X-ray scattering study of the folding dynamics of barnase. *J. Mol. Biol.* **405**, 1284–1294 (2011). doi:10.1016/j.jmb.2010.11.052 Medline

18. A. K. Svensson, O. Bilsel, E. Kondrashkina, J. A. Zitzewitz, C. R. Matthews, Mapping the folding free energy surface for metal-free human Cu,Zn superoxide dismutase. *J. Mol. Biol.* **364**, 1084–1102 (2006). doi:10.1016/j.jmb.2006.09.005 Medline

19. S. Piana, J. L. Klepeis, D. E. Shaw, Assessing the accuracy of physical models used in protein-folding simulations: Quantitative evidence from long molecular dynamics simulations. *Curr. Opin. Struct. Biol.* **24**, 98–105 (2014). doi:10.1016/j.sbi.2013.12.006 Medline

20. M. Junker, C. C. Schuster, A. V. McDonnell, K. A. Sorg, M. C. Finn, B. Berger, P. L. Clark, Pertactin beta-helix folding mechanism suggests common themes for the secretion and folding of autotransporter proteins. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 4918–4923 (2006). doi:10.1073/pnas.0507923103 Medline

21. J. P. Renn, M. Junker, R. N. Besingi, E. Braselmann, P. L. Clark, ATP-independent control of autotransporter virulence protein transport via the folding properties of the secreted protein. *Chem. Biol.* **19**, 287–296 (2012). doi:10.1016/j.chembiol.2011.11.009 Medline

22. E. Udho, K. S. Jakes, A. Finkelstein, TonB-dependent transporter FhuA in planar lipid bilayers: Partial exit of its plug from the barrel. *Biochemistry* **51**, 6753–6759 (2012). doi:10.1021/bi300493u Medline

23. P. X. Qi, T. R. Sosnick, S. W. Englander, The burst phase in ribonuclease A folding and solvent dependence of the unfolded state. *Nat. Struct. Biol.* **5**, 882–884 (1998). doi:10.1038/2321 Medline

24. J. Song, G. N. Gomes, C. C. Gradinaru, H. S. Chan, An adequate account of excluded volume is necessary to infer compactness and asphericity of disordered proteins by Förster resonance energy transfer. *J. Phys. Chem. B* **119**, 15191–15202 (2015). doi:10.1021/acs.jpcb.5b09133 Medline

25. H. M. Watkins, A. J. Simon, T. R. Sosnick, E. A. Lipman, R. P. Hjelm, K. W. Plaxco, Random coil negative control reproduces the discrepancy between scattering and FRET measurements of denatured protein dimensions. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 6631–6636 (2015). doi:10.1073/pnas.1418673112 Medline

26. R. K. Das, K. M. Ruff, R. V. Pappu, Relating sequence encoded information to form and function of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **32**, 102–112 (2015). doi:10.1016/j.sbi.2015.03.008 Medline

27. J. A. Riback, C. D. Katanski, J. L. Kear-Scott, E. V. Pilipenko, A. E. Rojek, T. R. Sosnick, D. A. Drummond, Stress-triggered phase separation is an adaptive, evolutionarily tuned response. *Cell* **168**, 1028–1040.e19 (2017). doi:10.1016/j.cell.2017.02.027

28. P. Tompa, G. D. Rose, The Levinthal paradox of the interactome. *Protein Sci.* **20**, 2074–2079 (2011). doi:10.1002/pro.747 Medline

29. M. Junker, P. L. Clark, Slow formation of aggregation-resistant beta-sheet folding intermediates. *Proteins* **78**, 812–824 (2010). doi:10.1002/prot.22609 Medline

30. J. M. Jumper, K. F. Freed, T. R. Sosnick, Rapid calculation of side chain packing and free energy with applications to protein molecular dynamics. https://arxiv.org/abs/1610.07277 (2017).

31. J. B. Xu, B. Berger, Fast and accurate algorithms for protein side-chain packing. *J. Assoc. Comput. Mach.* **53**, 533–557 (2006). doi:10.1145/1162349.1162350

32. D. Schneidman-Duhovny, M. Hammel, J. A. Tainer, A. Sali, Accurate SAXS profile computation and its assessment by contrast variation experiments. *Biophys. J.* **105**, 962–974 (2013). doi:10.1016/j.bpj.2013.07.020 Medline

33. C. J. Knight, J. S. Hub, WAXSiS: A web server for the calculation of SAXS/WAXS curves based on explicit-solvent molecular dynamics. *Nucleic Acids Res.* **43**, W225–W230 (2015). doi:10.1093/nar/gkv309 Medline

34. T. Ohta, Y. Oono, K. F. Freed, Static-coherent-scattering function for a single polymer-chain: Conformational space renormalization of polymers. V. *Phys. Rev. A* **25**, 2801–2811 (1982). doi:10.1103/PhysRevA.25.2801

35. G. Wang, R. L. Dunbrack Jr., PISCES: A protein sequence culling server. *Bioinformatics* **19**, 1589–1591 (2003). doi:10.1093/bioinformatics/btg224 Medline

36. T. P. Hopp, K. R. Woods, Prediction of protein antigenic determinants from amino acid sequences. *Proc. Natl. Acad. Sci. U.S.A.* **78**, 3824–3828 (1981). doi:10.1073/pnas.78.6.3824 Medline

37. P. Debye, Molecular-weight determination by light scattering. *J. Phys. Colloid Chem.* **51**, 18–32 (1947). doi:10.1021/j150451a002 Medline

38. J. A. Capp, A. Hagarman, D. C. Richardson, T. G. Oas, The statistical conformation of a highly flexible protein: Small-angle X-ray scattering of *S. aureus* protein A. *Structure* **22**, 1184–1195 (2014). doi:10.1016/j.str.2014.06.011 Medline

39. B. Hammouda, Sans from homogeneous polymer mixtures—a unified overview. *Adv. Polym. Sci.* **106**, 87–133 (1993). doi:10.1007/BFb0025862

40. T. Ohta, Y. Oono, K. F. Freed, Static scattering function for a polymer-chain in a good solvent. *Macromolecules* **14**, 1588–1590 (1981). doi:10.1021/ma50006a087