Supplementary Information

A Finely Balanced Order-Disorder Equilibrium Sculpts the Folding-Binding Landscape of an Antibiotic Sequestering Protein

Lawanya Natarajan,[1] Maria Laura De Sciscio,[2] Alessandro Nicola Nardi,[2] Ashok Sekhar,[,3]*

Alessandra Del Giudice,[,2] Marco D'Abramo*[,2] & Athi N. Naganathan*[,1]*

[1]Department of Biotechnology, Bhupat & Jyoti Mehta School of Biosciences, Indian Institute of Technology Madras, Chennai 600036, India.

[2]Department of Chemistry, Sapienza University of Rome, 00185, Italy.

[3]Molecular Biophysics Unit, Indian Institute of Science Bangalore, Bengaluru 560 012, India.

AUTHOR INFORMATION

Corresponding Author

e-mail: athi@iitm.ac.in, marco.dabramo@uniroma1.it, alessandra.delgiudice@uniroma1.it, ashoksekhar@iisc.ac.in

Phone: +91-44-2257 4140

Materials and Methods

Purification of TipAS The gene sequence corresponding to TipAS protein – MGINLTPEEKFEVFGDFDPDQYEEEVRERWGNTDAYRQSKEKTASYTKEDWQRIQD EADELTRRFVALMDAGEPADSEGAMDAAEDHRQGIARNHYDCGYEMHTCLGEMY VSDERFTRNIDAAKPGLAAYMRDAILANAVRHTPA, cloned between NdeI and SpeI restrictions sites of the pTXB1 vector was purchased from GenScript USA Inc. TipAS was constructed as a fusion protein, with a C terminal intein tag that enabled affinity based purification and in-column cleavage of the tag via IMPACT^{TM} (New England Biolabs). An additional Ala residue was added at the end of the naturally occurring sequence (with Pro as the last residue), for better efficiency of cleavage, thereby increase the yield of pure TipAS.

The protein was overexpressed in *E. coli* BL21(DE3) cells, cultured in LB broth containing 0.1 mg/ml of ampicillin, at 37 ℃ until OD_{600} = ~1, induced with 0.5 mM IPTG and grown for 4 hours post induction before harvesting the cells by centrifugation. Cell pellets were refrigerated at -80℃ for at least 6 hours before resuspending them into Buffer A – 20 mM Tris, 500 mM NaCl and 1 mM EDTA at pH 8. Cells pelleted from 500 ml of culture broth resuspended in 50 ml of buffer A with 1 mM PMSF (Phenylmethylsulphonyl Fluoride) to inhibit proteases released during cell disruption was subjected to sonication followed by centrifugation at 10,300 rpm for 60 minutes at 4℃. The cleared lysate was loaded into manually packed affinity chromatography columns pre-equilibrated with buffer A without EDTA. The flow rate was maintained between 0.5-0.7ml/min. The columns were washed with at least 10 column volumes each of buffer A and buffer B (20 mM Tris, 50 mM NaCl, pH 8) at a flow rate of 0.5 - 0.7 ml/min. Cleavage of intein tag from the fusion protein was achieved with ~2.5 column volumes of buffer B with 50 mM β-mercaptoethanol. The columns were sealed and incubated at room temperature for 16-18 hours before eluting with buffer B into fractions of 5 ml. All elution fractions from chitin columns were pooled together and applied onto a fresh chitin column pre-equilibrated with buffer B to ensure removal of free intein and fusion protein that might have eluted along with cleaved TipAS. Samples were taken at each stage of purification and analyzed through SDS-PAGE. Fractions containing protein were pooled and concentrated by lyophilization. The lyophilized sample was further purified by size exclusion chromatography, using HiLoad 26/600 Superdex 200 pg column (Cytiva) pre-equilibrated with 150 mM Ammonium acetate buffer at pH 8. The eluted protein fractions in ammonium acetate were analyzed by SDS-PAGE, pooled and lyophilized. ^{15}N labeled-TipAS was purified following the same steps as above but by growing the cells in minimal media (containing ^{15}N-labeled NH_4Cl as the sole nitrogen source), and with the final size-exclusion chromatography step replaced by buffer exchange through a 26/10 HiPrep desalting column.

Lyophilized TipAS was dissolved in 20 mM sodium phosphate buffer, pH 7 (ionic strength of 43 mM) for all experiments unless otherwise mentioned. Once dissolved the solution was filtered with 0.22-micron syringe filter, or (for SAXS measurements) centrifuged for 5 minutes at 13000 rpm. Concentrations were measured through a Jasco V-730 spectrophotometer or a ThermoFischer Nanodrop employing an extinction coefficient of 21430 $M^{-1}cm^{-1}$ at 280 nm.

Equilibrium spectroscopy Far-, near-UV circular dichroism and fluorescence experiments were performed as detailed before[1] at concentrations of ~10, 65 and 11 μM, respectively.

Thermal melts were typically recorded between 278 and 368 K at every 5 K intervals, and an equilibration time of at least 2 minutes at every temperature.

Differential scanning calorimetry Highly concentrated samples of TipAS were filtered and desalted with 26/10 HiPrep desalting column (Cytiva), to eliminate trace residual salts and to enable accurate heat capacity measurements. TipAS samples were also degassed before loading into the calorimetric cells. DSC profiles were acquired in a Microcal VP-DSC instrument at a scan rate of 1.5 K/min at concentrations of 54.1, 94.1 and 131.3 μM, respectively. Multiple buffer-buffer baselines were acquired before and after every protein scan to check for baseline drifts. The apparent heat capacities were then employed to estimate the absolute heat capacity via the method proposed by Sanchez-Ruiz and coworkers.[2]

Time-resolved fluorescence Fluorescence lifetimes of tryptophan residues in TipAS were measured in ChronosBH time-resolved spectrometer, with Ludox solution to measure the instrument response function. TipAS, at a concentration of 12 μM was excited at 300 nm and the time traces were collected from 278 to 353 K in 5 K intervals. The traces were fitted to tri-exponential functions at each temperature point – two lifetimes to account for the differential behavior of the two tryptophan residues, W30 and W51, in folded-like environment, and one mean life-time when they are in an unfolded-like environment. The longer lifetime (τ_1) is fixed at 7 nanoseconds to enable well-defined fits.

NMR spectroscopy TipAS NMR samples were made by dissolving lyophilized TipAS powder in 20 mM sodium phosphate buffer, pH 7 (ionic strength of 43 mM). 10 % D_2O (v/v) was added as the lock solvent to all NMR samples. All NMR spectra were acquired on a Bruker 18.8 T (800 MHz ^1H Larmor frequency) NMR spectrometer or an Agilent DD2 (14.1 T) 600 MHz spectrometer equipped with either a room-temperature (800 MHz) or a cryogenically cooled (600 MHz and 800 MHz) triple resonance single axis gradient probe. NMR spectra were processed with NMRPipe[3] and visualized with NMRDraw or NMRFAM-Sparky.[4] Longitudinal (R_1) and transverse (R_2) relaxation rate constants of the Trp indole sidechain ^{15}N were measured using pulse sequences previously published in literature.[5,6] R_2 values were determined using a rotating-frame relaxation rate constant ($R_{1\rho}$) measurement made with a spin lock field of amplitude $B_{SL} \sim 1.8$ kHz. $R_{1\rho}$ was converted to R_2 using the equation[6]:

$$R_2 = \frac{R_{1\rho} - R_1 \cos^2\theta}{\sin^2\theta}$$

where θ is the angle made by the effective field with the z-axis, given by:

$$\theta = \tan^{-1}\left(\frac{B_{SL}}{\delta}\right)$$

Here, δ is the residue-specific chemical shift offset from the spin-lock carrier frequency in Hz. 11 relaxation delays ranging from 10 ms to 800 ms were used to measure R_1, while 10 durations from 3 ms to 80 ms were used to determine $R_{1\rho}$. Three points in each array were duplicated to estimate the error in the intensities.

Small angle X-ray scattering (SAXS) SAXS measurements of TipAS solutions as a function of temperature were performed both with a lab source and at a synchrotron radiation facility.

SAXS measurements at SAXSLab Sapienza were performed with a Xeuss 2.0 Q-Xoom system (Xenocs SAS, Grenoble, France), equipped with a micro-focus Genix 3D X-ray source with Cu anode (λ = 0.1542 nm) and a two-dimensional Pilatus3 R 300K detector which can be placed at variable distance from the sample (Dectris Ltd., Baden, Switzerland). The beam size was defined to be 0.5 mm × 0.5 mm through the two-pinhole collimation system equipped with "scatterless" slits. Calibration of the scattering vector q range ($q=4\pi\sin\theta/\lambda$ with 2θ the scattering angle and λ the photon wavelength), was performed using silver behenate. The sample-to-detector distance was kept at 0.55 m so that the explored q region was $0.1\ nm^{-1} < q < 6\ nm^{-1}$. The protein sample was loaded into a vacuum-tight quartz capillary cell placed in a instrument sample chamber at reduced pressure (~0.2 mbar) within a thermalized holder. Starting from a temperature of 20°C, 3h acquisitions were collected after equilibrating the temperature to 10, 15, 20, 25, 30, 35, 40 and then back to 30 and 20 °C (temperature uncertainty is ±0.1°C). The two-dimensional scattering patterns were subtracted for the "dark" counts, and then masked, azimuthally averaged, and normalized for transmitted beam intensity, exposure time and subtended solid angle per pixel, using the FOXTROT software developed at SOLEIL. The one-dimensional intensity vs. q profiles were then subtracted for the data of the buffer measured in the same cell and same temperatures and divided by the capillary thickness calibrated using water scattering to obtain intensity in macroscopic scattering cross section units (cm^{-1}). The intensity was further normalized to units of $cm^{-1}/mg\cdot ml^{-1}$ dividing by the protein concentration.

SAXS measurements were also performed at the BioSAXS beamline BM29 of the European Synchrotron Radiation Facility (ESRF, Grenoble, France).[7] The SAXS measurements were performed with the automatic sample-changer. The capillary and storage temperatures were set sequentially at all values as mentioned before in the main text. Sample volumes of 50 µl were withdrawn and a set of 10 consecutive one second exposures was acquired during sample flowing in the capillary. The data reduction of 2D images (masking, azimuthal averaging, transmission scaling) was performed through the analysis pipeline.[11] The I(q) profiles obtained from the frames were compared to assess radiation damage and then averaged. The scattering contribution of the capillary filled with buffer was subtracted and the intensity was divided by the protein mass concentration. The intensity was scaled to macroscopic scattering cross section (I(q) in cm^{-1}) using water as a standard. Two repetitions of the measurement procedure for each temperature were run and the data were averaged after checking for overlap. Sample details and data collection parameters are also reported in Supporting Table S1.

Analysis of the scattering profiles was performed with the tools of ATSAS 3.2.[8] The Guinier approximation was employed to estimate the radius of gyration (R_g) and the scattered intensity extrapolated at zero angle (I(0)). The indirect Fourier transform method was applied to obtain the pair distance distribution function (P(r)), with an estimate of the maximum particle dimension (D_{max}), in addition to an independent calculation of I(0) and R_g. For this purpose, we employed both the GNOM algorithm,[9] by imposing a constant D_{max} for all temperatures for consistency of comparison, and the BayesApp method,[10] in which the D_{max} is optimized according to Bayesian inference.

More in detail, all collected data where first compared in Guinier plot representation to estimate the radius of gyration R_g and the scattered intensity extrapolated at zero angle I(0) using the Guinier approximation:

$$ln(I) = ln(I(0)) - \frac{R_g^2}{3} q^2$$

From the I(0) an estimate of the molecular weight can be obtained according to the equation

$$MW[kDa] = \frac{I(q)[cm^{-1}] \cdot N_A[mol^{-1}]}{\Delta\rho_M^2[cm^2 \cdot g^{-2}] \cdot c[mg \cdot cm^{-3}]}$$

$\Delta\rho_M^2$ is the squared scattering contrast per mass of protein, which is related to the protein partial specific volume $\tilde{v}$ (cm^3/g) and the electron density of the water solvent ρ_{solv} (electrons/cm^3) at each temperature:

$$\Delta\rho_M = r_e(\rho_{M,dry} - \rho_{solv} \cdot \tilde{v})$$

With $r_e = 2.8179 \cdot 10^{-13}$ cm (scattering length of the electron) and $\rho_{M,dry} = 3.22 \cdot 10^{23}$ e/g (number of electrons per mass of dry protein), assuming a common $\tilde{v}$ value of 0.732 cm^3/g at all considered temperatures and according to the decrease of water density provides $\Delta\rho_M^2$ from $4.76 \cdot 10^{20}\ cm^2 \cdot g^{-2}$ at 283 K to $4.98 \cdot 10^{20}\ cm^2 \cdot g^{-2}$ at 313 K.

For the data collected in-lab, the plots showed linearity in two distinct q regions (Figure S5A): in a lower q range (0.1-0.68 nm^{-1}) with estimated R_g of the order of 2.4 nm, and in a higher q range (0.74-1.08 nm^{-1}) with estimated R_g of the order of 1.8 nm. The low-q region showed a systematic increase of R_g with increasing temperature from 10 to 40 °C, with reversibility upon cooling back to 30°C and 20°C. The slope in the higher q region was instead substantially constant with temperature. The extrapolated I(0) values were observed to decrease as a function of increasing temperature in both cases (Figure S5D).

By comparing the normalized Kratky plots obtained by using the parameters (R_g and I(0)) from both estimates, we could perceive indeed that the lower values of R_g observed to be constant as a function of temperature (around 1.8 nm) provide normalized Kratky plots with a maximum in the expected position for a fully compact particle ($qR_g=\sqrt{3}$, $[I(q)/I(0)] \cdot (q*R_g)^2 = 3/e \approx 1.104$) (Figure S5I), following the Guinier law in the low q and the Porod law in the high q. On the other hand, the R_g values obtained in the lower q range observed to slightly increase reversibly above 30 °C, provided normalized Kratky plots slightly deviating from this fully compact behavior, with a shifted maximum usually related to an overall more extended object[12] (Figure S5H). From this preliminary evaluation it could be suggested that the second linearity region of the Guinier plot at higher q could be due to a fully compact sub-portion of the TipAS protein, while the overall construct has a small additional portion with a higher degree of flexibility.

The synchrotron data deviated from linearity in the lower q region (Figure S5B), suggesting an appreciably higher amount of protein oligomers both pre-existent and possibly also induced by radiation damage. By restricting the q range for the Guinier fit in the lower q region to 0.2-0.5 nm^{-1} a reasonable linearity could be found, and values of R_g around 3 nm were obtained (Figure S5D).

To better correlate the structural parameters inferred by SAXS analysis to the conformational state of the individual TipAS monomers, we considered as the most reliable data those giving rise to the smallest R_g and I(0) values in the q range < 0.5 nm^{-1} (lab experiment at 1.98 mg/ml) and, since the noise level was rather high for q > 1 nm^{-1}, they were merged with the synchrotron data at the same temperatures and similar protein concentration (1.78 mg/ml), after assessing the overlap starting from q > 0.9 nm^{-1}. Therefore a set of merged data comprising the lab data for 0.11<q<0.9 nm^{-1}, the synchrotron data for 1<q<5 nm^{-1} and considering the average of the two profiles in the overlap region (0.9<q<1 nm^{-1}) were obtained using the merging tool of SAXSUtilities2[13] and used for further analysis.

MW values estimated from the absolute intensity I(0) and a common value of the partial specific volume equal to 0.732 cm^3/g at all temperatures are in the range 25-20 kDa (Table S2), from 1.5 to 1.2 times higher than the expected monomer mass of 16.6 kDa and also show a decreasing trend with increasing temperature (Figure S5D). In order for the predicted MW to be constant and equal to the expected monomer value of 16.6 kDa at all temperatures, $\tilde{v}$ values increasing from 0.68 at 283 K to 0.71 cm^3/g at 313 K should be assumed. Changes of protein hydration with temperature could contribute to the observed intensity and are not straightforward to be taken into account. The presence of a fraction of small oligomers, for example dimers, could also contribute to the scattering in the lowest q-range and partly justify the estimated MW values. However the observed reversible trend with increasing temperature for the same sample (involving an increase of R_g but a decrease of I(0)) would not be compatible with just an increase of oligomer fraction. The molecular weight was also estimated via concentration-independent methods and the tool based on Bayesian inference to estimate a most probable value and a confidence interval was employed [14]. Such estimates resulted in MW values within the interval 18-24 kDa (Table S3) also when applied to the theoretical SAXS profiles calculated from MD simulations (Section on MD in the main text).

Modeling SAXS profiles The ensemble method implemented in EOM 3.0 was applied to model the SAXS profiles[15,16] of TipAS at the representative temperatures of 283 K, 298 K, and 313 K. The quality of the fits was employed to deduce insights about the size distributions and conformational flexibility of TipAS. Pools of 10,000 conformers were generated using the RANCH algorithm. Overall eight pools of models with multiple flexibility criteria were compared, as detailed below:

i) Two conformer pools were generated by fixing the C-terminal domain (residues 55-143) according to the crystallographic coordinates of the Thiostrepton bound form (PDB id: 2MBZ), while the N-terminal domain (residues 1-54) was represented as a chain of dummy residues obeying to specific Quasi-Ramachandran databases for dihedral angles: either fully disordered (pool *ndis*) or assuming helical regions (9-13, 19-26, 32-37, 48-54) like in the folded conformation (pool *nhel*).

ii) In addition, possible conformations in which the region of highest flexibility could represent a loop joining two structured helical subdomains were considered. Within this assumption, a

series of four conformers pools were generated considering the regions 1-30 (N-term) and 46-145 (C-term) as fully structured and the region 31-45 as a flexible loop joining them and allowing for domain reorientation. This selection was based on the observation that in the MD simulation at 360 K the residues classified as belonging to the structured regions still have a RMSF < 0.4, while residues 31-45 show higher RMSF values. The coordinates of the N-term (1-30) and C-term (46-145) structured domains were assigned according to a representative simulation frame, and the C-term was fixed. The linker region 31-45 was generated as a dummy residue chain according to four different assumptions: all residues 31-45 disordered (pool *nddd*), all residues 31-45 helical (pool *nhhh*), residues 31-38 helical and 39-45 disordered (pool *nhhd*) and finally residues 31-35 helical and 36-45 disordered (pool *nhdd*).

iii) An additional pool (*nmix*) comprised conformers generated both with the unstructured N-term assumption (*ndis*) and with the structured and reoriented N-term (*nhhd*) mentioned before.

iv) Finally, a pool was generated constraining the coordinates of the structured N-term to be fixed in the folded conformation (*nfix*), to check the impact on the overall size parameters.

The theoretical SAXS profiles for all models of each pool were calculated using the FFMAKER tool, which uses the CRYSOL method, imposing default settings for the hydration layer and atomic excluded volume, 25 spherical harmonics and 301 data points. The genetic algorithm optimization implemented in GAJOE was employed to select from the pool ensembles of conformers whose averaged scattering profile would fit the experimental data. A fixed ensemble size of 50 curves per ensemble without repetitions was imposed. Eight GAJOE runs (each involving 100 automatic repetitions) for each pool were run to estimate standard deviations for the average dimensional parameters ($<R_g>$ and $<D_{max}>$ of selected final ensembles and of overall generated histogram in the 100 repetitions) and for the χ^2 between experimental and calculated SAXS profiles.

Binding experiments Thiostrepton (TSN) was purchased from Sigma Aldrich and dissolved in DMSO to prepare stock solutions. Concentrations of TSN were measured with an extinction coefficient of 27000 M^{-1} cm^{-1} at 280 nm. Equilibrium fluorescence based binding experiments were carried out at 288, 298 and 310 K with 10 μM of TipAS, and a range of TSN concentrations from 0.1 nM to 100 μM, and with the final DMSO concentration being 0.5% (v/v). Time resolved fluorescence binding experiments at 288, 298 and 310 K were performed with 10 μM of TipAS, with TSN concentrations varying from 1 nM to 100 μM, and at final DMSO concentrations of 5% (v/v), to enable better solubility. Fluorescence lifetimes of tryptophan residues were also measured with 1:1 molar ratio of TipAS to TSN, for every 5 K between 278 and 353 K.

Block Wako−Saitô−Muñoz−Eaton (bWSME) model The model is described in great detail in several recent works.[17,18] Briefly, stretches of two consecutive residues (block size of 2) are assumed to exhibit either an unfolded (represented by zero) or folded-like (represent by one) conformational status. Combining this with ensembles defined by single sequence approximation (SSA; a single stretch of folded island), double sequence approximation (DSA; two stretches of folded islands) and DSA allowing for interaction across islands (DSAw/L), we partition the conformational landscape of TipAS into 1,702,915 microstates. The model's energy function includes contributions from stabilizing free-energetic terms - van der Waals interactions identified with a 5 Å spherical cut-off involving heavy atoms and with every

atomic-level interaction worth ξ, Debye-Hückel electrostatics, simplified solvation defined by the heat capacity change per native contact (ΔC_p^{cont}) - and a destabilizing entropic term introduced through the entropic penalty for fixing a residue in a folded conformation from an unfolded-like conformation (ΔS_{conf}). Additionally, an excess entropic penalty of 6.06 J mol^{-1} K^{-1} is attributed to residues identified as a coil by STRIDE,[19] and proline residues are assigned a ΔS_{conf} of zero due to their restricted backbone flexibility. With this description it is straightforward to calculate the statistical weight of every one of the microstates and hence the total partition function. A theoretical heat capacity curve is then generated from derivatives of the total partition function. The experimental heat capacity curve was employed to calibrate the model parameters through non-linear least-squares fitting procedures in MATLAB. While the model captured the weak stability of NTSD *de novo*, the strength of van der Waals interaction for the first 46 residues was reduced to 85.4% of its original number to capture the shape of the experimental heat capacity curve. The final model parameters are: van der Waals interaction energy of -77.3 ± 5.1 J mol^{-1} per native contact, entropic penalty of -15.4 ± 1.0 J mol^{-1} K^{-1} per residue and ΔC_p^{cont} of -0.31 ± 0.04 J mol^{-1} K^{-1} per native contact. To predict the conformational landscape of the orthologs, we employ identical parameters to that of TipAS without modulation of the NTSD contact strength.

MD simulations Molecular dynamics (MD) simulations of TipAS were carried out using the GROMACS 2022 software package[20] using explicit SPC (simple point charge[21]) water model and using the Amber ff99SB-ILDN force-field[22] at temperatures of 280, 300, 320, 340, and 360 K. For each temperature, the initial structure generated by AlphaFold2[23] with the higher predicted local distance difference test (pLDDT) was chosen (cyan line in Figure 9D) and centered in a cubic box and solvated using the SPC water model. Then, 15 Na^+ ions were inserted in the simulation box (each replacing a water molecule) to achieve system electroneutrality. An energy minimization step was further performed using the steepest descent algorithm without position restraints. After the minimization, and for the simulations at temperatures higher than 280 K, the temperature was increased to reach the target temperature through a series of short MD steps in the NVT ensemble, lasting 50 picoseconds with a timestep of 2 femtoseconds. When the target temperature was reached, a series of equilibration steps, lasting 50 picoseconds with a timestep of 2 femtoseconds, were performed in the NVT ensemble; after each run, we checked the pressure of the system and tuned the size of the box in order to reach the correct density value.[24,25] Following this, each system was simulated for 2 microseconds at a fixed temperature (indicated above) and volume, using a timestep of 2 femtoseconds. The temperature was kept constant by using the V-rescale thermostat[26] employing a τ_T of 0.1 ps. Electrostatic interactions were calculated using the particle mesh Ewald method[27] with a cutoff of 1.1 nm, and the same cutoff value was also used for the van der Waals interactions. For each temperature, 4 replicas were performed, using the simulation protocol described above. This results in a total time sampling of 8 μs for each temperature. Helical residues were identified using the DSSP algorithm as implemented in the MDTraj package[28] and a principal component analysis (PCA) was performed on the MD trajectories to obtain information on the most relevant protein motions. In such an approach, a $3N \times 3N$ covariance matrix of atomic positions is built from the MD simulation on a selected group of N atoms (usually C_α, as in the present case) and diagonalized. From the diagonalization, a set of eigenvectors and the corresponding eigenvalues are obtained. The eigenvectors represent the directions of the principal (essential) motions of the system.

Dihedral angle calculations were performed using the "gangle" program in GROMACS. χ_1 was defined by selecting the following atoms for each tryptophan residue: N, C_α, C_β, C_γ. Theoretical SAXS profiles of MD simulation ensembles were generated following the methodology described in literature.[29] Native contact analysis was also performed to evaluate the distribution of folding status across the protein sequence. The native state was defined as the equilibrated structure at 280 K (Q=1). The fraction of native contacts was calculated through MDTraj tool, on all the heavy atoms, using a cutoff of 0.45 nm, a λ factor equal to 0.18 nm and a β value of 50 nm^{-1}, according to Eaton and co-workers.[30] To check the convergence of the results, we report in Figure S4 the analyses of the deviations of root mean square fluctuation (RMSF), the first two principal components from the PCA, and the fraction of native contacts per residue among the performed replicas.

Supporting Text

Analysis of NMR Data In order to probe the local environment at the site of the indole sidechains of W30 and W51 in each sub-state, we measured ^{15}N R_1 (Figure S3B) relaxation rate constants of wt TipAS at 293 K, where all three states are present at comparable populations. Since W30U and W51N are overlapped, their intensities cannot be independently extracted, and ^{15}N R_1 values are reported only for W30N, W30N*, W51N* and W51U. R_1 is sensitive to dynamics on the picosecond-nanosecond (ps-ns) timescale. The ^{15}N R_1 is highest for W51U (1.22 ± 0.08 s^{-1}), consistent with the presence of significant ps-ns dynamics in the unfolded state. Interestingly, the R_1 for W30N (1.18 ± 0.08 s^{-1}) is virtually identical to W51U suggesting that the W30 sidechain in helix 2 has considerable local mobility even in the binding-incompetent N sub-state. Taken in conjunction with the high fluorescence emission maximum of W30 in W51A TipAS (λ_{max} ~348 nm, Figure 2E), this observation indicates that W30 may be solvent accessible in N and not part of a rigid hydrophobic core. In contrast to W30, the fluorescence emission maximum of W51 in W30A TipAS is much lower (λ_{max} ~342 nm at 293 K; Figure 2E). W51 is therefore in a much less polar environment than W30, and we would therefore predict that it has more restricted mobility and a smaller R_1. While the ^{15}N R_1 of W51 in wt TipAS cannot be measured because of peak overlap, the R_1 of W51 in W30A TipAS (0.86 ± 0.06 s^{-1}) is much lesser than that of W51U (1.6 ± 0.2 s^{-1}) in the same variant, demonstrating that ps-ns dynamics is restricted in W51 compared to W30. The R_1 values for both Trp sidechains in the binding-competent N* state, on the other hand, are lower (0.9 ± 0.1 s^{-1}, Figure S3B) indicating that the indole sidechains are not as dynamic on the ps-ns timescale as in the unfolded state because of native or non-native interactions with the TipAS core.

Table S1. Summary of SAXS data acquisition information, sample details, and data analysis software used.

(a) Sample details for the SAXS experiments	
	TipAS
Concentration range (mg ml^{-1})	1.79-3.88
Storage and dilution buffer composition	20 mM Na-phosphate, pH 7.2
Extinction coefficient for concentration determination from UV absorbance at 280 nm ($M^{-1}cm^{-1}$)	21555

(b) SAXS data collection parameters - Lab	
Source, instrument	Xenocs Xeuss 2.0-QXoom
Detector	PILATUS 300K
Beam geometry (mm^2)	0.5 x 0.5
Wavelength (Å)	1.5418
sample-to-detector distance (m)	0.55
q-measurement range (nm^{-1})	0.11-6.3
Absolute scaling method	water scattering I(0)= 0.01632 cm^{-1} at 20°C
Capillary path length (mm)	1.5
Sample volume (µl)	100
Exposure time (s)	1800
Number of exposures	6
Sample temperature (ºC)	20°C, 10°C, 15°C, 20°C, 25°C, 30°C, 35°C, 40°C, 30°C, 20°C

(c) SAXS data collection parameters - Synchrotron	
Source, instrument	BM29 (ESRF)
Detector	PILATUS 2M
Beam geometry (mm^2)	0.2 x 0.2
Wavelength (Å)	0.99186
sample-to-detector distance (m)	2.813
q-measurement range (nm^{-1})	0.06-5.2
Absolute scaling method	water scattering I(0)= 0.01632 cm^{-1} at 20°C
Capillary path length (mm)	1
Sample volume (µl)	50
Exposure time (s)	1

Number of exposures	10
Extra flow time (s)	15
Sample temperature (ºC)	20°C, 10°C, 15°C, 20°C, 25°C, 30°C, 35°C, 40°C, 30°C, 20°C

(d) Software employed for SAS data reduction, analysis, and interpretation	
Solvent subtraction, averaging and basic analysis (Guinier fit, P(R), MW)	ATSAS 3.2
Theoretical intensity calculations	WAXSiS GROMACS module
Ensemble optimization	EOM 3.0
Molecular graphics	PyMOL 1.8

Table S2. Summary of model independent SAXS data analysis of TipAS in the temperature range 283-313 K.

Temperature (K)	283	288	293	298	303	308	313	303*	293*
Guinier fit									
q range 0.11-0.68 nm^{-1}									
R_g (nm)	2.37	2.39	2.37	2.44	2.46	2.59	2.67	2.35	2.42
$\sigma(R_g)$ (nm)	0.03	0.03	0.02	0.03	0.03	0.03	0.03	0.03	0.03
I(0) ($cm^{-1}/mg \cdot ml^{-1}$)	1.98E-02	1.92E-02	1.85E-02	1.80E-02	1.73E-02	1.70E-02	1.62E-02	1.78E-02	1.67E-02
σ(I(0))	2E-04	2E-04	2E-04	2E-04	2E-04	2E-04	2E-04	2E-04	2E-04
MW from I(0) (kDa)	25.1	24.3	23.2	22.4	21.4	20.7	19.6	22.3	20.6
Indirect Fourier transform (GNOM)									
q range 0.11-3 nm^{-1}									
Dmax imposed for P(r) (nm)	8								
R_g (nm)	2.33	2.32	2.37	2.28	2.34	2.50	2.53	2.28	2.34
$\sigma(R_g)$ (nm)	0.03	0.02	0.02	0.04	0.03	0.03	0.03	0.03	0.04
I(0) ($cm^{-1}/mg \cdot ml^{-1}$)	1.87E-02	1.78E-02	1.74E-02	1.61E-02	1.58E-02	1.55E-02	1.44E-02	1.64E-02	1.55E-02
σ(I(0))	3E-04	2E-04	2E-04	3E-04	3E-04	3E-04	2E-04	2E-04	3E-04
MW from I(0) (kDa)	23.7	22.5	22.0	20.4	20.0	19.7	18.2	20.7	19.6
Indirect Fourier transform (BayesApp)									
q range 0.4-3 nm^{-1}									
R_g (nm)	2.19	2.11	2.30	2.14	2.19	2.57	2.61	2.14	2.30
$\sigma(R_g)$ (nm)	0.01	0.01	0.02	0.01	0.01	0.05	0.04	0.01	0.01
I(0) ($cm^{-1}/mg \cdot ml^{-1}$)	1.79E-02	1.67E-02	1.73E-02	1.54E-02	1.51E-02	1.57E-02	1.48E-02	1.58E-02	1.54E-02
σ(I(0))	5E-05	6E-05	1E-04	5E-05	8E-05	3E-04	2E-04	5E-05	7E-05
D_{max} (nm)	7.01	6.72	7.91	6.86	7.12	8.76	8.47	6.81	7.34
$\sigma(D_{max})$ (nm)	0.05	0.07	0.15	0.05	0.08	0.19	0.18	0.04	0.07
MW Bayesian									
credibility interval (kDa)	20.8 24.0	20.8 24.0	22.7 24.6	23.4 25.9	20.8 24.0	22.7 25.2	22.1 25.2	19.6 22.7	20.8 24.0
interval probability (%)	0.95	0.97	0.91	0.97	0.94	0.92	0.91	0.96	0.94

*Cooled back after heating.

Table S3. Summary of size parameters obtained from the theoretical SAXS profiles calculated from MD simulations of TipAS.

Temperature (K)	280	300	320	340
Guinier fit				
q range 0-0.6 nm^{-1}				
R_g (nm)	2.17	2.18	2.22	2.35
$\sigma(R_g)$ (nm)	0.02	0.02	0.03	0.02
I(0) ($cm^{-1}/mg \cdot ml^{-1}$)*	0.0775	0.0704	0.0486	0.0557
MW Bayesian				
credibility interval	19.6	18.4	19.6	20.9
(kDa)	22.8	22.1	23.4	24.0
interval probability (%)	0.93	0.91	0.92	0.91

*Scaling factor from intensity in electrons^2 $= 0.001 \cdot r_e^2 \cdot N_A/MW$, where $MW = 16.6\ kDa$, $N_A = 6.022 \cdot 10^{23} mol^{-1}$ and $r_e = 2.8179 \cdot 10^{-13} cm$.

Table S4. Summary of the EOM fitting attempts of the TipAS SAXS experimental data. (a) Average structural parameters of the generated random pools of conformers: the radius of gyration (R_g), the maximum particle size (D_{max}), and the flexibility index (R_{flex}). The R_{flex} index (between 0 and 1) is a metric obtained from the entropy (information content) of the distribution and can be used to approximate the flexibility. Using this metric the size distributions of the selected ensembles after fitting can be numerically compared to that of the initial random pool, which is supposed to represent the reference for the randomness of the distribution under the specific structural constraints. (b) Structural parameters and fitting results obtained by eight repetitions of the fitting algorithm applied to the SAXS data of TipAS at temperatures 283 K, 298 K, 313 K considering each of the generated pools. The average value and standard deviation for each quantity are reported. For each pool these parameters are shown from left to right: average R_g of the selected pool of conformers; average D_{max} of the selected pool of conformers; flexibility parameter of the selected distribution; check parameter representing the ratio between the standard deviations of the selected and starting size histograms distributions; experimental vs. calculated SAXS intensity agreement index χ^2.

a)

Pool	**Pool average**		**Flexibility index**
	R_g (nm)	**D_{max} (nm)**	**R_{flex}**
ndis	2.39	8.12	0.87
nhel	2.39	8.23	0.87
nhhd+ndis	2.28	7.74	0.84
nddd	2.23	7.57	0.80
nhhh	2.19	7.73	0.90
nhhd	2.17	7.37	0.90
nhdd	2.16	7.25	0.88
nfix	1.60	4.93	0.52

b)

Fitting data	Selected pool average				Flexibility index		Distribution width ratio		Goodness of fit	
	$\langle R_g \rangle$ (nm)	σ	$\langle D_{max} \rangle$ (nm)	σ	$\langle R_{flex} \rangle$	σ	$\langle R_\sigma \rangle$	σ	$\langle \chi^2 \rangle$	σ
283 K										
ndis	2.12	0.002	7.10	0.01	0.798	0.001	0.92	0.005	3.44	0.12
nhel	2.12	0.002	7.14	0.01	0.793	0.001	0.90	0.008	3.67	0.08
nhhd+ndis	2.04	0.004	6.78	0.02	0.772	0.003	0.91	0.015	3.86	0.07
nddd	1.99	0.002	6.57	0.00	0.733	0.001	0.93	0.008	4.59	0.08
nhhh	1.99	0.002	6.77	0.01	0.916	0.002	1.18	0.004	5.97	0.17
nhhd	1.95	0.001	6.39	0.01	0.821	0.002	0.87	0.006	4.34	0.05
nhdd	1.94	0.002	6.33	0.01	0.798	0.002	0.84	0.012	4.11	0.12
nfix	1.62	0.000	4.93	0.00	0.491	0.004	0.93	0.018	8.90	0.01
298 K										
ndis	2.13	0.004	7.00	0.021	0.791	0.003	0.89	0.014	2.10	0.03
nhel	2.12	0.002	7.03	0.006	0.787	0.002	0.88	0.005	2.26	0.07
nhhd+ndis	2.04	0.003	6.67	0.015	0.764	0.002	0.87	0.012	2.18	0.06
nddd	1.99	0.002	6.48	0.007	0.730	0.002	0.91	0.012	2.62	0.10
nhhh	2.00	0.001	6.71	0.003	0.913	0.001	1.17	0.005	3.37	0.15
nhhd	1.95	0.002	6.33	0.011	0.818	0.002	0.85	0.009	2.51	0.03
nhdd	1.94	0.003	6.27	0.009	0.795	0.002	0.82	0.011	2.42	0.03
nfix	1.60	0.000	4.93	0.000	0.563	0.001	1.51	0.004	8.51	0.03
313 K										
ndis	2.17	0.001	7.16	0.01	0.768	0.002	0.77	0.013	2.82	0.03
nhel	2.16	0.003	7.21	0.02	0.764	0.003	0.76	0.016	3.08	0.04
nhhd+ndis	2.08	0.002	6.84	0.01	0.739	0.002	0.74	0.005	2.87	0.01
nddd	2.02	0.001	6.62	0.01	0.699	0.002	0.76	0.011	3.11	0.02
nhhh	2.02	0.001	6.80	0.01	0.893	0.001	1.02	0.008	3.56	0.04
nhhd	2.00	0.002	6.48	0.01	0.796	0.003	0.72	0.014	3.04	0.02
nhdd	1.99	0.0004	6.44	0.01	0.775	0.002	0.70	0.009	3.03	0.02
nfix	1.59	0.0004	4.93	0.00	0.597	0.001	1.73	0.013	11.31	0.01

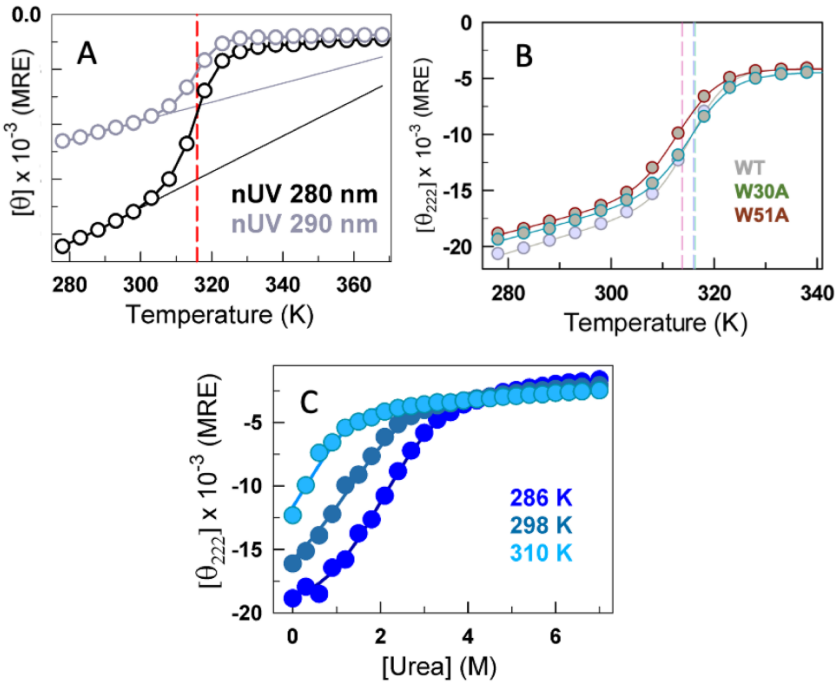

Figure S1 (A) Signal changes at 280 and 290 nm as a function of temperature monitored by near-UV CD on WT TipAS. The curves through the points are two-state fits while the continuous lines are folded baselines. The dashed line represents the melting temperature. Note the steep changes in signal at low temperatures indicative of structural modulations in the native ensemble at temperatures <310 K. (B) Far-UV CD monitored thermal unfolding curves for the WT and the two mutant variants. The W30A mutant has identical melting temperature as the WT (dashed green line) while the W51A mutant has a 2 K lower melting temperature (dashed dark red line). (C) Changes in far-UV CD signal of WT TipAS at 222 nm as a function of urea concentration. Note the absence of pre-transition regime even at 286 K, indicating that the structure of TipAS is sensitive to minor perturbations.

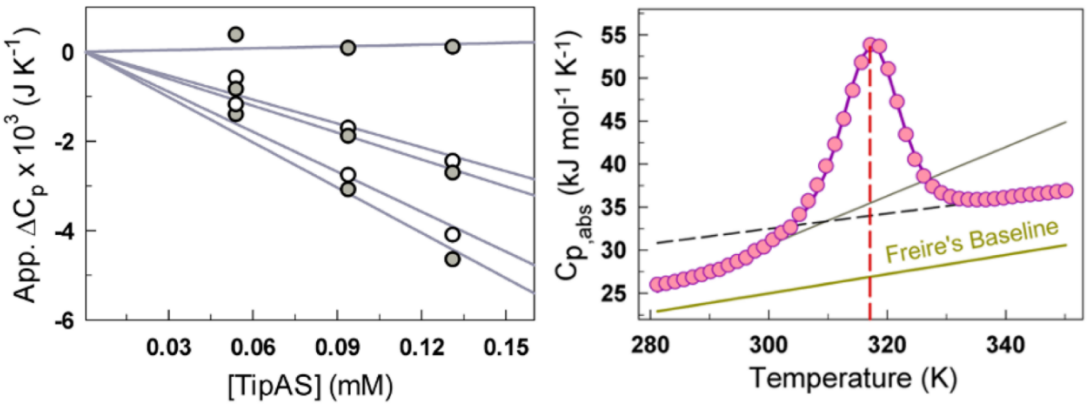

Figure S2 Apparent heat capacity as a function of protein concentration at select temperatures (left). The panel to the right plots the absolute heat capacity of TipAS fit to a two-state model (curves through the point) together with the folded (continuous line) and unfolded baselines (dashed line). Freire's baseline represents the expected enthalpic fluctuations of a fully folded protein. The vertical dashed line indicates the melting temperature.

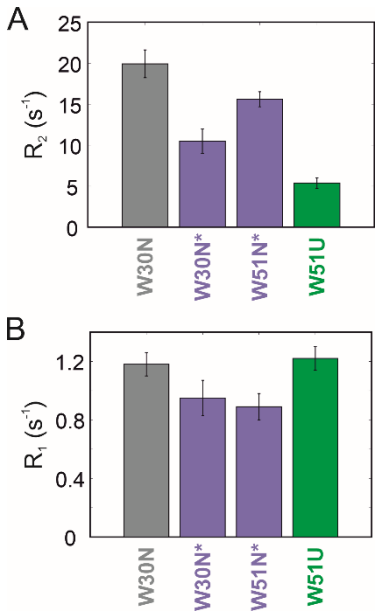

Figure S3. (A) Transverse relaxation rate constants (R_2) and (B) longitudinal relaxation rate constants (R_1) for W30 and W51 indole ^{15}N nuclei in various states. R_1 and R_2 values are reported only for states whose resonances are not overlapped.

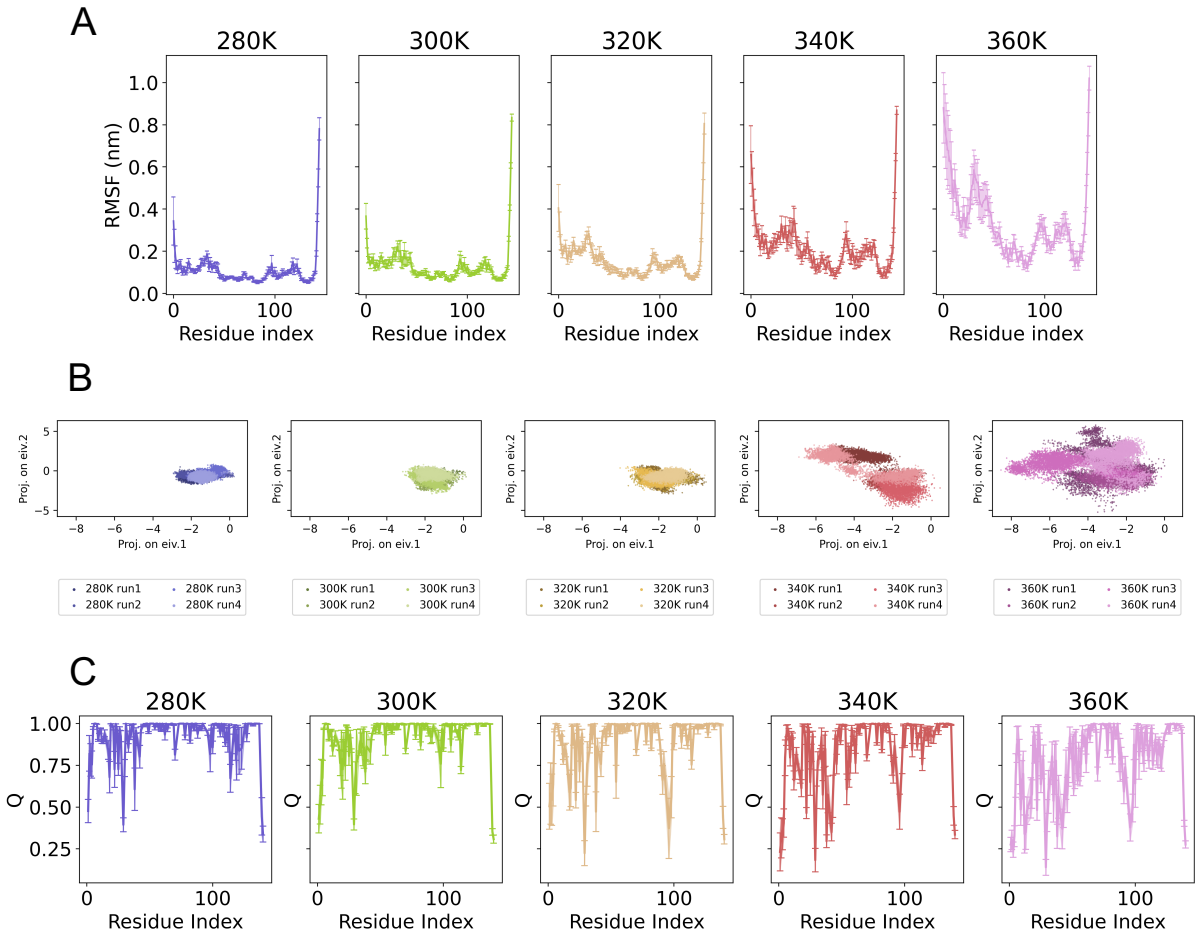

Figure S4 Convergence analysis of MD simulations replicas for the five investigated temperatures. (A) Average C_α RMSF and the corresponding deviation as error bars. (B) Projection of all the frames onto the essential subspace defined by the first two eigenvectors (representing around 50% of the total variance), obtained by a PCA on the concatenated trajectories of all the replicas at all the temperatures. (C) Per-residue fraction of native contacts (Q) averaged for each temperature with the corresponding error bar.

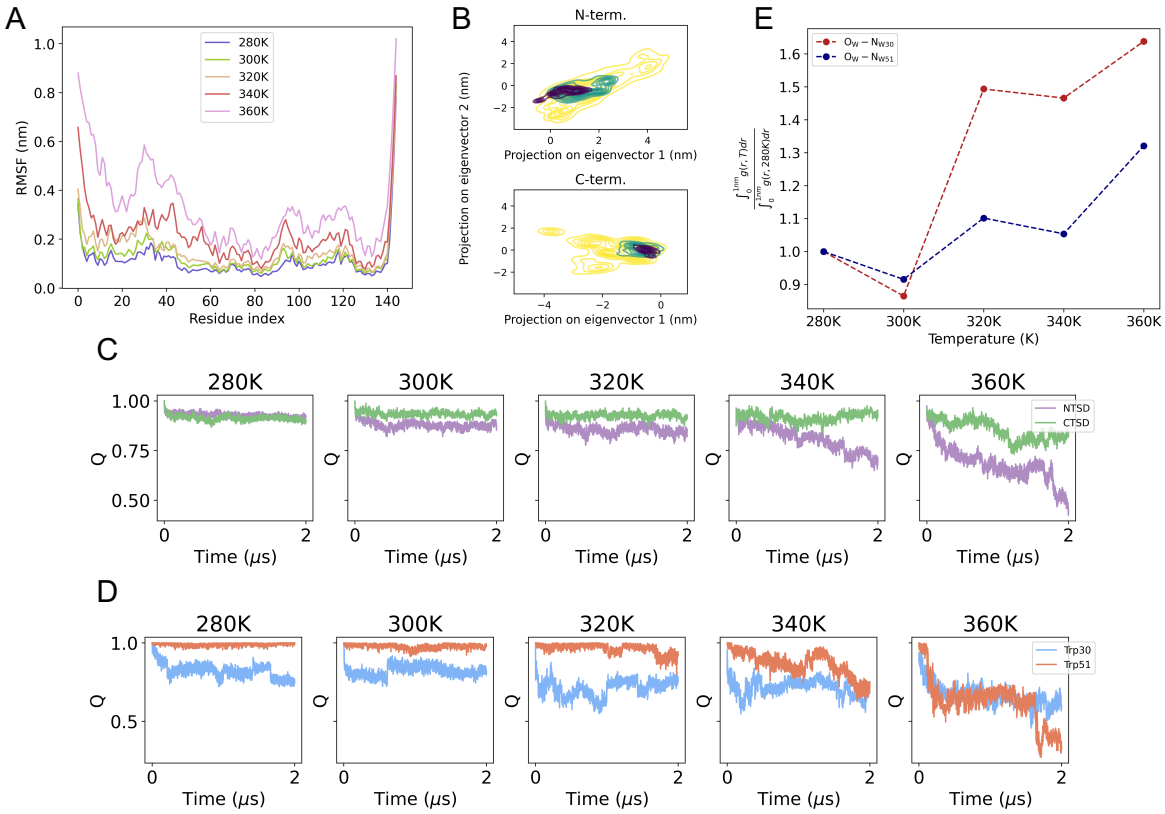

Figure S5 (A) C_α root-mean-square fluctuations (RMSF) as a function of residue index from MD simulations. (B) Kernel Density Estimation of the projected conformations onto the essential space as obtained by PCA of the generated ensembles at 280 K (violet), 320 K (green) and 360 K (yellow) for NTSD (N-term., top panel) and CTSD (C-term., bottom panel). (C) Fraction of native contacts in the NTSD (N-term., purple) and CTSD (C-term., green) as a function of simulation time, calculated using the equilibrated structure at 280 K as reference. (D) Fraction of native contacts for the two fluorescent residues, Trp30 and Trp51, calculated on Trps' side chain dummy atoms and their spatially neighboring atoms, i.e. the (dummy) atoms within 0.5 nm from one or more atoms of the indole group, while the initial structure at 280K was used as reference. (E) Integrals of the radial distribution function, g(r), of the distance between the oxygen atom of the SPC water molecules (O_w) and the nitrogen atom of the tryptophan residues W30 and W51 (N_{W30} and N_{W51}, respectively), obtained from the MD simulations at different temperatures. The integration was performed from 0 to 1 nm for each radial distribution function and the resulting areas were normalized with respect to the one obtained at 280 K for both the W30 and W51 residues.

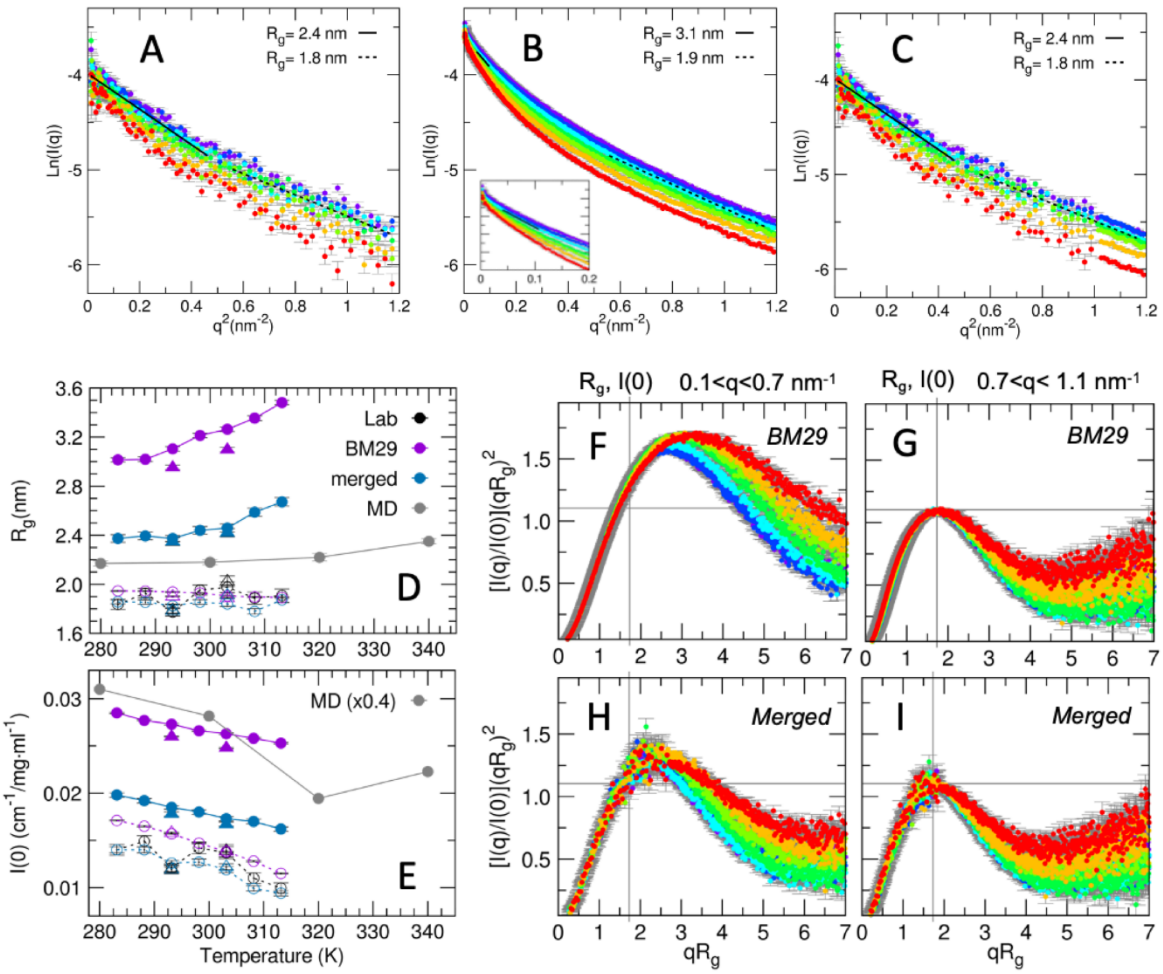

Figure S6 Preliminary evaluation of the SAXS data of TipAS in the temperature range 283-313 K. (A) Guinier plot of the SAXS data collected with a lab instrument ("Lab", TipAS 1.96 mg/ml). (B) Guinier plot of the SAXS data collected at a synchrotron beamline ("BM29", TipAS 1.78 mg/ml). (C) Guinier plot of the SAXS data obtained by merging "Lab" and "BM29" data for the same temperatures. (D, E) R_g and I(0) values obtained by Guinier fit in the lower q range (filled symbols and solid lines) and the higher q range (empty symbols and dashed lines), as a function of temperature. Results from the lab instrument ("Lab", black) and the measurements at a synchrotron beamline ("BM29", purple) are compared with those from the merged data (blue, in the low q they coincide with the "Lab"). Values from the SAXS data collected upon cooling back at 303 K and 293 K are shown as triangles. The parameters obtained from the theoretical SAXS profiles calculated from the MD simulations are also shown, using a scaling factor of 0.4 for the I(0) values. It can be noticed that the SAXS intensity extrapolated at zero angle, I(0), has a rather linear decrease as a function of increasing temperature (also in Figure 7C in the main text), and this may be related to changes in protein excluded volume and structure of the solvation layer. (F-I) Dimensionless Kratky plots of the SAXS data ("Merged" and "BM29"); the R_g and I(0) values obtained by Guinier fit in the lower q range (F,H) and in the higher q range (G,I) were applied. The temperatures are 283 K (purple), 288 K (blue), 293 K (cyan), 298 K (light green), 303 K (dark green), 308 K (orange) and 313 K (red).

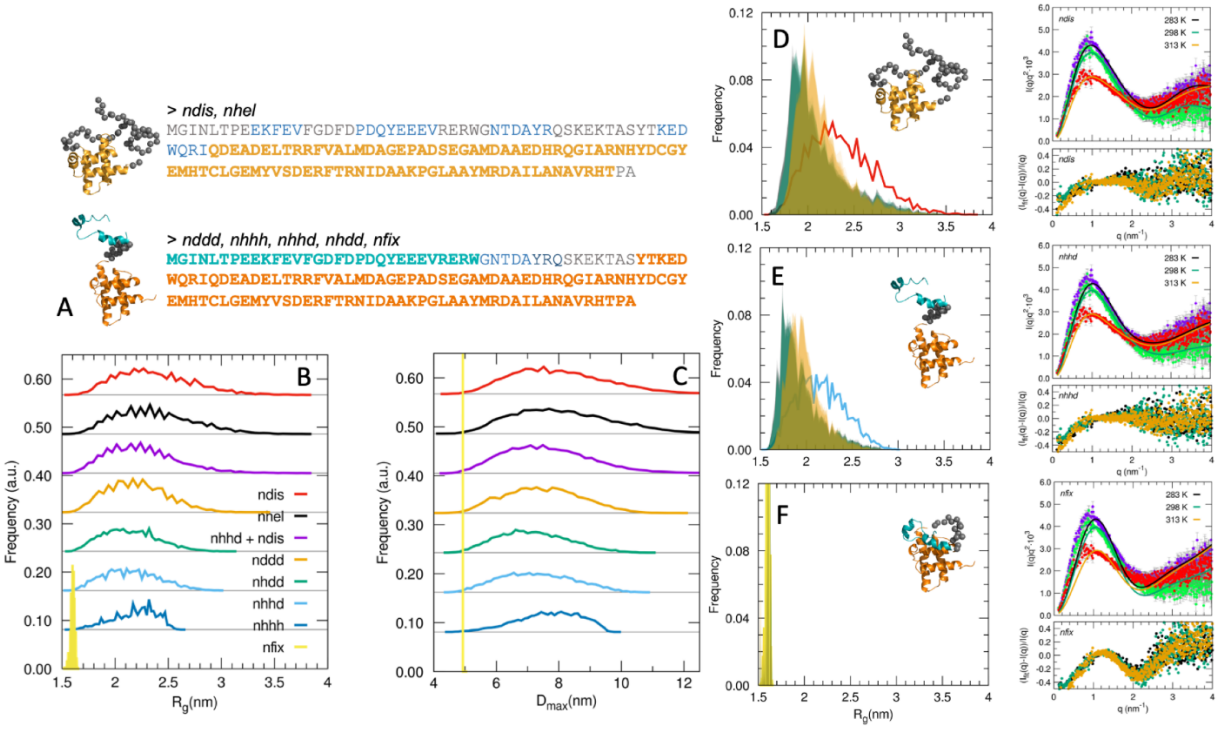

Figure S7 (A) Graphical representation of the structured domains considered in generating TipAS conformer pools for SAXS data modeling through the ensemble optimization method (EOM). On the left, one example of the generated conformers (10,000) is shown for the pools *ndis* and *nhhd* as representatives for two different assumptions: N-terminal fully flexible, or structured N-terminal connected to C-terminal via a flexible linker. The dummy residues are shown as grey spheres, whereas the structured portions described according to the coordinates from the crystallographic structure or a representative MD frame are shown in orange (C-term) and cyan (N-term). The TipAS sequence (on the right) is colored with the same color code used for the corresponding models. The *nfix* pool was generated by constraining the structured N-term to remain in a fixed position as from the original MD frame. For further details, see the Methods and Supporting Methods section. (B, C) R_g (panel B) and D_{max} (panel C) distributions for the different pools mentioned in panel A and for the additional pool comprising both conformers of the *nhhd* and *ndis* pools (20,000 models). The distributions are shifted vertically for better visualization. (D-E) Left column: R_g frequency distributions for the initial pool (solid line) and for the final selected pools to fit SAXS data of TipAS at 283 K (grey shaded area), 298 K (green shaded area), 313 K (orange shaded area). Results from eight optimizations are shown superimposed. Right column: best-fit theoretical SAXS profile obtained from the EOM optimization (solid line) superimposed to the experimental data points in Kratky plot representation (dots). Below, the corresponding normalized residuals $(I_{calc}(q) - I_{exp}(q))/I_{exp}(q)$ are shown. These results are shown for the three pools *ndis* (D), *nhhd* (E) and *nfix* (F).

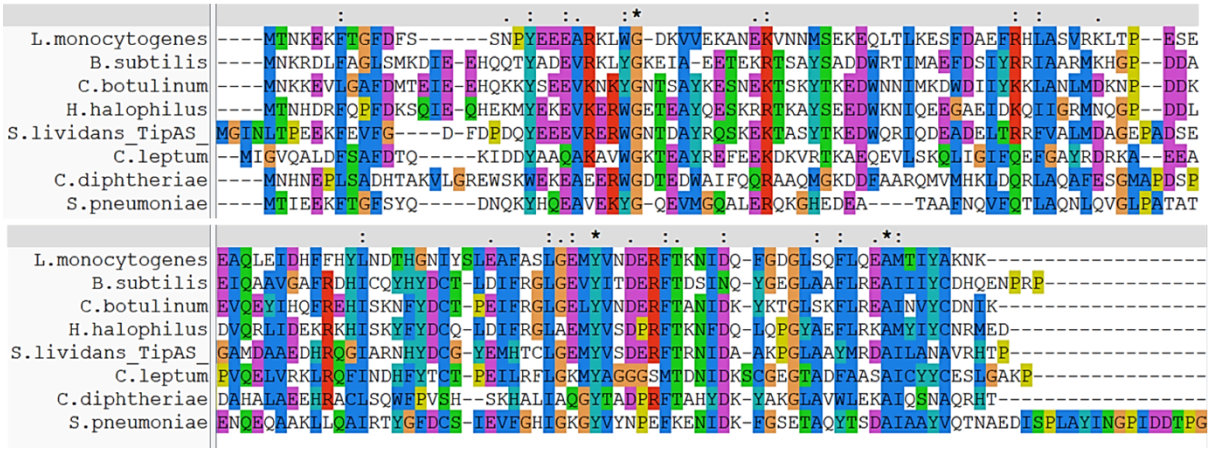

Figure S8 Sequence alignment of the seven TipAS orthologs studied in this work.

Supplementary References

(1) Narayan, A.; Campos, L. A.; Bhatia, S.; Fushman, D.; Naganathan, A. N. Graded Structural Polymorphism in a Bacterial Thermosensor Protein. *J. Am. Chem. Soc.* **2017**, *139*, 792–802.

(2) Guzman-Casado, M.; Parody-Morreale, A.; Robic, S.; Marqusee, S.; Sanchez-Ruiz, J. M. Energetic Evidence for Formation of a PH-Dependent Hydrophobic Cluster in the Denatured State of Thermus Thermophilus Ribonuclease H. *J. Mol. Biol.* **2003**, *329* (4), 731–743.

(3) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes. *J. Biomol. NMR* **1995**, *6* (3), 277–293.

(4) Lee, W.; Tonelli, M.; Markley, J. L. NMRFAM-SPARKY: Enhanced Software for Biomolecular NMR Spectroscopy. *Bioinformatics* **2015**, *31* (8), 1325–1327. https://doi.org/10.1093/bioinformatics/btu830.

(5) Farrow, N. A.; Zhang, O.; Forman-Kay, J. D.; Kay, L. E. A Heteronuclear Correlation Experiment for Simultaneous Determination of 15N Longitudinal Decay and Chemical Exchange Rates of Systems in Slow Equilibrium. *J. Biomol. NMR* **1994**, *4* (5), 727–734. https://doi.org/10.1007/BF00404280.

(6) Korzhnev, D. M.; Skrynnikov, N. R.; Millet, O.; Torchia, D. A.; Kay, L. E. An NMR Experiment for the Accurate Measurement of Heteronuclear Spin-Lock Relaxation Rates. *J. Am. Chem. Soc.* **2002**, *124* (36), 10743–10753. https://doi.org/10.1021/ja0204776.

(7) Tully, M. D.; Kieffer, J.; Brennich, M. E.; Cohen Aberdam, R.; Florial, J. B.; Hutin, S.; Oscarsson, M.; Beteva, A.; Popov, A.; Moussaoui, D.; Theveneau, P.; Papp, G.; Gigmes, J.; Cipriani, F.; McCarthy, A.; Zubieta, C.; Mueller-Dieckmann, C.; Leonard, G.; Pernot, P. BioSAXS at European Synchrotron Radiation Facility - Extremely Brilliant Source: BM29 with an Upgraded Source, Detector, Robot, Sample Environment, Data Collection and Analysis Software. *J. Synchrotron Radiat.* **2023**, *30* (Pt 1), 258–266. https://doi.org/10.1107/S1600577522011286.

(8) Manalastas-Cantos, K.; Konarev, P. V; Hajizadeh, N. R.; Kikhney, A. G.; Petoukhov, M. V; Molodenskiy, D. S.; Panjkovich, A.; Mertens, H. D. T.; Gruzinov, A.; Borges, C.; Jeffries, C. M.; Svergun, D. I.; Franke, D. ATSAS 3.0: Expanded Functionality and New Tools for Small-Angle Scattering Data Analysis. *J.App.Cryst.* **2021**, *54* (Pt 1), 343–355. https://doi.org/10.1107/S1600576720013412.

(9) Svergun, D. I. Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria. *J. App. Cryst.* **1992**, *25*, 495–503.

(10) Hansen, S. BayesApp : A Web Site for Indirect Transformation of Small-Angle Scattering Data. *J. App. Cryst.* **2012**, *45*, 566–567.

(11) Kieffer, J.; Brennich, M.; Florial, J. B.; Oscarsson, M.; De Maria Antolinos, A.; Tully, M.; Pernot, P.; Jeng, U. New Data Analysis for BioSAXS at the ESRF. *J. Synchrotron Radiat.* **2022**, *29* (Pt 5), 1318–1328. https://doi.org/10.1107/S1600577522007238.

(12) Receveur-Brechot, V.; Durand, D. How Random Are Intrinsically Disordered Proteins? A Small Angle Scattering Perspective. *Curr. Protein Pept. Sci.* **2012**, *13* (1),

55–75. https://doi.org/10.2174/138920312799277901.

(13) Michael, S. SAXSutilities2: A Graphical User Interface for Processing and Analysis of Small-Angle X-Ray Scattering Data. **2021**. https://doi.org/10.5281/ZENODO.5825707.

(14) Hajizadeh, N. R.; Franke, D.; Jeffries, C. M.; Svergun, D. I. Consensus Bayesian Assessment of Protein Molecular Mass from Solution X-Ray Scattering Data. *Sci. Rep.* **2018**, *8* (1), 1–13. https://doi.org/10.1038/s41598-018-25355-2.

(15) Bernado, P.; Mylonas, E.; Petoukhov, M. V; Blackledge, M.; Svergun, D. I. Structural Characterization of Flexible Proteins Using Small-Angle X-Ray Scattering. *J. Am. Chem. Soc.* **2007**, *129* (17), 5656–5664.

(16) Tria, G.; Mertens, H. D. T.; Kachala, M.; Svergun, D. I. Advanced Ensemble Modelling of Flexible Macromolecules Using X-Ray Solution Scattering. *IUCrJ* **2015**, *2* (Pt 2), 207–217. https://doi.org/10.1107/S205225251500202X.

(17) Gopi, S.; Aranganathan, A.; Naganathan, A. N. Thermodynamics and Folding Landscapes of Large Proteins from a Statistical Mechanical Model. *Curr. Res. Struct. Biol.* **2019**, *1*, 6–12.

(18) Anantakrishnan, S.; Naganathan, A. N. Thermodynamic Architecture and Conformational Plasticity of GPCRs. *Nat. Commun.* **2023**, *14* (1), 128. https://doi.org/10.1038/s41467-023-35790-z.

(19) Heinig, M.; Frishman, D. STRIDE: A Web Server for Secondary Structure Assignment from Known Atomic Coordinates of Proteins. *Nuc. Acids Res.* **2004**, *32*, W500–W502.

(20) Páll, S.; Zhmurov, A.; Bauer, P.; Abraham, M.; Lundborg, M.; Gray, A.; Hess, B.; Lindahl, E. Heterogeneous Parallelization and Acceleration of Molecular Dynamics Simulations in GROMACS. *J. Chem. Phys.* **2020**, *153* (13), 134110. https://doi.org/10.1063/5.0018516.

(21) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. Interaction Models for Water in Relation to Protein Hydration; 1981; pp 331–342. https://doi.org/10.1007/978-94-015-7658-1_21.

(22) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber Ff99SB Protein Force Field. *Proteins: Struct. Funct. Bioinf.* **2010**, *78* (8), 1950–1958. https://doi.org/10.1002/prot.22711.

(23) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. https://doi.org/10.1038/s41586-021-03819-2.

(24) Del Galdo, S.; Marracino, P.; D'Abramo, M.; Amadei, A. In Silico Characterization of Protein Partial Molecular Volumes and Hydration Shells. *Phys. Chem. Chem. Phys.*

2015, *17* (46), 31270–31277. https://doi.org/10.1039/c5cp05891k.

(25) De Sciscio, M. L.; Nardi, A. N.; Parisi, G.; Bulfaro, G.; Costanzo, A.; Gugole, E.; Exertier, C.; Freda, I.; Savino, C.; Vallone, B.; Montemiglio, L. C.; D'Abramo, M. Effect of Salts on the Conformational Dynamics of the Cytochrome P450 OleP. *Molecules* **2023**, *28* (2). https://doi.org/10.3390/molecules28020832.

(26) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126* (1), 14101. https://doi.org/10.1063/1.2408420.

(27) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.

(28) McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; Swails, J. M.; Hernández, C. X.; Schwantes, C. R.; Wang, L.-P.; Lane, T. J.; Pande, V. S. MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophys. J.* **2015**, *109* (8), 1528–1532. https://doi.org/10.1016/j.bpj.2015.08.015.

(29) Chen, P.-C.; Hub, J. S. Validating Solution Ensembles from Molecular Dynamics Simulation by Wide-Angle X-Ray Scattering Data. *Biophys. J.* **2014**, *107* (2), 435–447. https://doi.org/10.1016/j.bpj.2014.06.006.

(30) Best, R. B.; Hummer, G.; Eaton, W. A. Native Contacts Determine Protein Folding Mechanisms in Atomistic Simulations. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110* (44), 17874–17879. https://doi.org/10.1073/pnas.1311599110.