Supporting Information

Tuning the Continuum of Structural States in the Native Ensemble of a Regulatory Protein

*Abhishek Narayan and Athi N. Naganathan**

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

e-mail: athi@iitm.ac.in

Materials and Methods

Wako-Saitô-Muñoz-Eaton (WSME) Model The Ising-like WSME model,[1,2] with energy terms for packing, solvation,[3,4] all-to-all native electrostatics[4] – all of which are lumped into a stabilization free energy term (ΔG^{stab}) - and structure-dependent conformational entropy,[5] was employed to characterize the differential scanning calorimetry (DSC) profile of Cnu following the protocol prescribed before.[6] A 5 Å heavy atom cut-off including nearest neighbors was used to generate the contact-map that is provided as an input to the model. The final parameters as noted in the recent work on Cnu are[6]: van der Waals interaction (vdW) energy per native heavy-atom contact (ξ = 78.9 J mol^{-1}), entropic cost for fixing a helical residue (as identified by STRIDE) in the native conformation (ΔS_{conf} = -14.5 J mol^{-1} K^{-1} per residue) and ΔS_{conf} = -20.6 J mol^{-1} K^{-1} per residue for all other residues (i.e. an excess conformational entropy of ~6.1 J mol^{-1} K^{-1} per residue). The ionic strength was fixed to 0.17 M and the protonation state of charged residues was fixed to that of pH 7.0, mimicking the experimental conditions.

The chemical denaturant dependence was introduced as before following the conventional linear denaturant dependence of the folding free energy.[7] In short, the stabilization free energy of the WSME model that includes contributions from the various energetic terms is assumed to depend linearly on urea concentration[8,9] with a proportionality constant determined by m_{cont} (=0.0017 kJ mol^{-1} M^{-1} per native contact) that reproduces the apparent C_m of ~5.3 M at 298 K. A more detailed description of the WSME model, its energy terms and the procedure employed to calculate the total partition function and free energy profiles can be found in recent works.[4,5,8-10]

Protein Expression and Purification Cnu and H-NS_{1-59} were over-expressed in *E. coli* BL21 DE3 Star (Invitrogen), purified and then lyophilized following the protocol described before.[6]

Equilibrium Spectroscopic Measurements All experiments were performed at pH 8.0 and 170 mM ionic strength conditions (20 mM phosphate buffer + 112 mM NaCl). Far- and near-UV CD temperature-wavelength spectra were collected on a JASCO J-815 spectropolarimeter coupled to a Peltier unit in 1 mm and 10 mm quartz cuvettes, respectively. Protein intrinsic fluorescence was measured by exciting Cnu at 274 nm in a Chirascan-plus qCD instrument (Applied Photophysics Ltd.) using a 10x10 mm cuvette coupled with a Peltier system for temperature control. The protein quantum yields (QY) were calculated using NATA quantum yield (excitation wavelength 274 nm) at 298 K and pH 7.0 as a reference.

Simulations The all-atom coarse-grained model of Onuchic and co-workers (SMOG[11]) was used to generate the starting files for simulations in GROMACS as reported before.[6] Briefly, a heavy-atom contact map of Cnu was generated with a 6 Å cut-off and excluding up to 3 nearest neighbors; this was fed into the SMOG server for generating a native-centric potential with default parameters. The simulations were run for 10^8 steps with a 0.5 femtosecond time-step (as carbon mass =1 in reduced units) for each of the temperatures in the range between 50 – 150 (in reduced units). The resulting single-molecule trajectories were analyzed with GROMACS commands and in-house codes in MATLAB. The resulting free-energy profiles and the temperature dependence of the different observables can be found in supporting reference 6. The all-atom MD simulations in explicit water were also performed following the protocol in supporting reference 6.

Hydrodynamic Measurements A GE Superdex 200 HR analytical size exclusion chromatographic column was calibrated with five different proteins (Ferritin, Aldolase, Conalbumin, Albumin & Ribonuclease) of known Stokes radii at 278 K.[6] Cnu equilibrated at different urea concentrations was then passed through the column and the elution time was used to calculate the Stokes radius using the standard graph above.

Binding Studies The H-NS_{1-59}-Cnu binary complex formation was monitored by tracking the fluorescence anisotropy changes of the sole tryptophan in Cnu (W67). The experiments were carried out in a Chirascan-plus qCD instrument (Applied Photophysics Ltd.) equipped with a fluorescence polarization (FP) detector and temperature controlled cuvette holder in a 10 x 10 mm quartz cuvette. Cnu was excited at 295 nm (2 nm bandwidth) and a 320 nm cut-off filter was placed in front of the FP detector for filtering out the excitation light. Cnu (4 μM) was titrated with increasing concentrations of H-NS_{1-59} and equilibrated for three minutes with stirring at each addition. The titration curves were fitted to a 1:1 (Cnu:H-NS_{1-59}) thermodynamic model to calculate dissociation constants (K_D).

Supporting References

(1) Wako, H.; Saito, N. Statistical Mechanical Theory of Protein Conformation .2. Folding Pathway for Protein. *J. Phys. Soc. Japan* **1978**, *44*, 1939-1945.

(2) Muñoz, V.; Eaton, W. A. A simple model for calculating the kinetics of protein folding from three-dimensional structures. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11311-11316.

(3) Bruscolini, P.; Naganathan, A. N. Quantitative Prediction of Protein Folding Behaviors from a Simple Statistical Model. *J. Am. Chem. Soc.* **2011**, *133*, 5372-5379.

(4) Naganathan, A. N. Predictions from an Ising-like Statistical Mechanical Model on the Dynamic and Thermodynamic Effects of Protein Surface Electrostatics. *J. Chem. Theory Comput.* **2012**, *8*, 4646-4656.

(5) Rajasekaran, N.; Gopi, S.; Narayan, A.; Naganathan, A. N. Quantifying Protein Disorder through Measures of Excess Conformational Entropy. *J. Phys. Chem. B* **2016**, *120*, 4341-4350.

(6) 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.

(7) Myers, J. K.; Pace, C. N.; Scholtz, J. M. Denaturant m-Values and Heat-Capacity Changes - Relation to Changes in Accessible Surface-Areas of Protein Unfolding. *Protein Sci.* **1995**, *4*, 2138-2148.

(8) Naganathan, A. N.; Sanchez-Ruiz, J. M.; Munshi, S.; Suresh, S. Are Protein Folding Intermediates the Evolutionary Consequence of Functional Constraints? *J. Phys. Chem. B* **2015**, *119*, 1323-1333.

(9) Munshi, S.; Naganathan, A. N. Imprints of function on the folding landscape: functional role for an intermediate in a conserved eukaryotic binding protein. *Phys. Chem. Chem. Phys.* **2015**, *17*, 11042-11052.

(10) Naganathan, A. N. A Rapid, Ensemble and Free Energy Based Method for Engineering Protein Stabilities. *J. Phys. Chem. B* **2013**, *117*, 4956-4964.

(11) Noel, J. K.; Whitford, P. C.; Sanbonmatsu, K. Y.; Onuchic, J. N. SMOG@ctbp: simplified deployment of structure-based models in GROMACS. *Nuc. Acids Res.* **2010**, *38*, W657-W661.

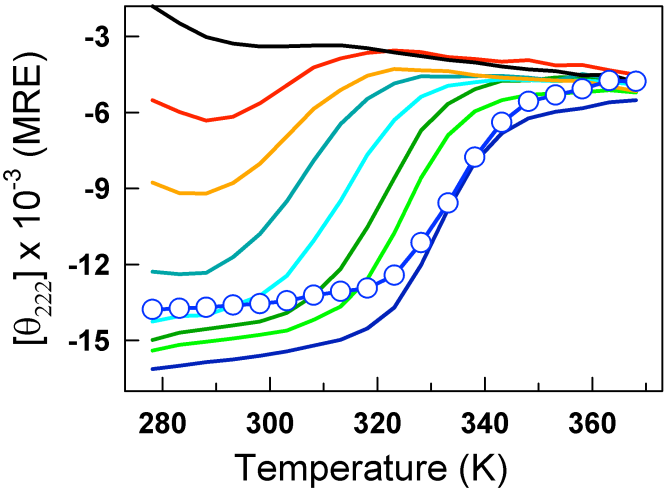

Figure S1. The far-UV CD of Cnu in mean residue ellipticity (MRE) units of deg. cm^2 $dmol^{-1}$ as a function of temperature at various urea concentrations. The urea concentrations are 0 M (blue), 1 M (dark blue), 2 M (light green), 3 M (dark green), 4 M (cyan), 4.75 M (dark cyan), 5.5 M (orange), 6 M (red) and 7 M (black).

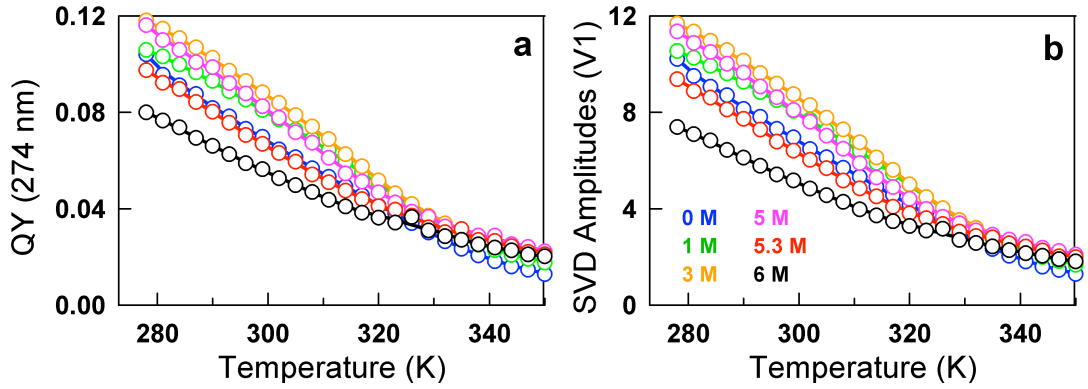

Figure S2. a) Temperature dependence of the protein quantum yield upon excitation of Cnu at 274 nm. b) The SVD amplitudes of the first component (Figure 3b in the main text) is near-identical in relative magnitude and trend to the QY data in panel (a) indicating that they capture the intrinsic temperature dependence of fluorescence quite nicely.

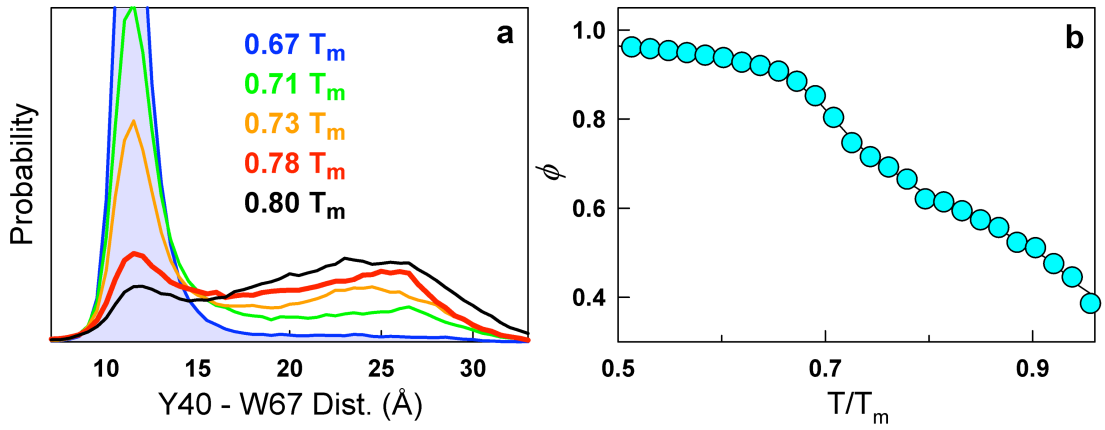

Figure S3. The distribution of C_α-C_α distances between Y40 and W67 (panel a) and the probability (ϕ) of a fully formed H-NS binding interface involving helices 3 and 4 (*i.e.* a well-folded protein; panel b), from all-atom coarse-grained simulations at each temperature. Note that the changes in the structure happen much earlier than the thermodynamic midpoint (i.e. temperatures < 1.0 T_m) indicating that these represent conformational fluctuations in the native ensemble that moves continuously in response to perturbations.

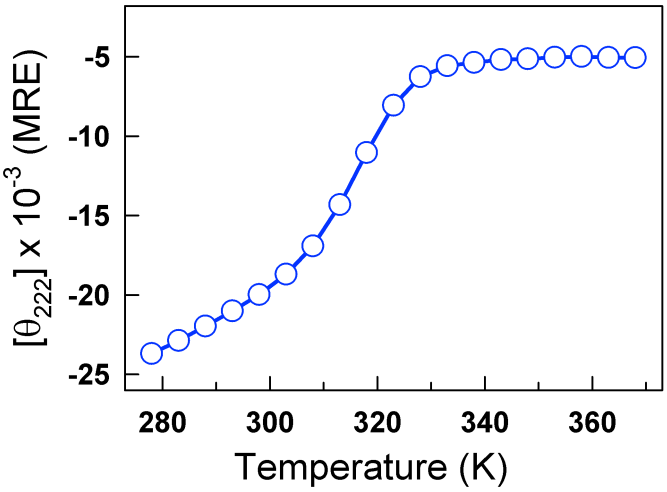

Figure S4. Temperature dependent mean residue ellipticity of the H-$NS_{1\text{-}59}$ fragment at 222 nm as monitored by far-UV CD. The line is shown to guide the eye.