Sequence Reversal Prevents Chain Collapse and Yields Heat-Sensitive Intrinsic Disorder

Lance R. English, Alexander Tischer, [...], and Steven T. Whitten

Abstract

Sequence patterns of charge, hydrophobicity, hydrogen bonding, and other amino acid physicochemical properties contribute to mechanisms of protein folding, but how sequence composition and patterns influence the conformational dynamics of the denatured state ensemble is not fully understood. To investigate structure-sequence relationships in the denatured state, we reversed the sequence of staphylococcal nuclease and characterized its structure, thermodynamic character, and hydrodynamic radius using circular dichroism spectroscopy, dynamic light scattering, analytical ultracentrifugation, and size-exclusion chromatography as a function of temperature. The macromolecular size of “Retro-nuclease” is highly expanded in solution with characteristics similar to biological intrinsically disordered proteins. In contradistinction to a disordered state, Retro-nuclease exhibits a broad sigmoid transition of its hydrodynamic dimensions as temperature is increased, indicating a thermodynamically controlled compaction. Counterintuitively, the magnitude of these temperature-induced hydrodynamic changes exceed that observed from thermal denaturation of folded unaltered staphylococcal nuclease. Undetectable by calorimetry and intrinsic tryptophan fluorescence, the lack of heat capacity or fluorescence changes throughout the thermal transition indicate canonical hydrophobic collapse did not drive the Retro-nuclease structural transitions. Temperature-dependent circular dichroism spectroscopy performed on Retro-nuclease and computer simulations correlate to temperature sensitivity in the intrinsic sampling of backbone conformations for polyproline II and α-helix. The experimental results indicate a role for sequence direction in mediating the collapse of the polypeptide chain, whereas the simulation trends illustrate the generality of the observed heat effects on disordered protein structure.

Introduction

Intrinsic disorder in biological proteins correlates with sequence composition and sequence patterns. Unlike the heterogeneous composition of amino acids and weak repetition in the sequences of folded proteins (1, 2), intrinsically disordered proteins (IDPs) and disordered protein domains generally have a lower sequence complexity (3) with strong preferences for hydrophilic and charged amino acid side chains over aromatic and hydrophobic side chains (4, 5). The lack of structure in IDPs as compared to the stable secondary and tertiary structures of folded proteins therefore appears to arise from fundamental disparities between properties of the primary sequence.

Folded proteins can be destabilized by relatively minor structural perturbations, for instance, by the deletion of a few terminal residues (6, 7) or reduction of disulfide bonds (8). These modifications can provide experimental access to denatured states in the absence of chemical denaturants. Denatured states under normal conditions are compact (6, 7, 9) relative to chemically denatured proteins (10) and show evidence of secondary (11, 12) and “residual” structure (13) that can be native like (14). Residual structure and hydrophobic clusters can persist in some denatured state ensembles at even high concentrations of chemical denaturants (15, 16, 17), suggesting inherent sequence-encoded preferences for structural ordering. An understanding of the relationships linking sequence to conformation, thermodynamic stability, and dynamics is needed to describe protein folding mechanistically, and an understanding of the changes to these structure-energy relationships as a sequence transitions to IDP characteristics is necessary to describe folding activity associated with IDP biology (18, 19).

An additional strategy to destabilize protein folds and gain access to denatured states is sequence reversal. Early experiments with protein sequence reversal were performed to test a “side-chain-only” hypothesis, whereby the pattern of hydrophobic and hydrophilic positions would be maintained in the reversed sequence to promote folding (20). Modeling of peptide structures indicated that reversing the direction of the protein backbone could yield a topological mirror of the parent (21). Supporting the side-chain-only hypothesis, computer simulation with a lattice model showed that reversing the sequence of the B domain of staphylococcal protein A could result in a stable, folded structure (22). Subsequent studies using circular dichroism (CD) and NMR spectroscopies, however, demonstrated instead a disordered protein (20). Reversed sequences from the SH3 domain of α-spectrin, the B1 domain from staphylococcal protein G, and rubredoxin also did not fold (20, 23). For the GCN4 leucine zipper, which contains a palindromic hydrophobicity profile and adopts a coiled-coil helical structure, reversing the direction of L-amino acids in a 35-residue fragment produced a “retro-leucine zipper” that folds into a stable four-helix bundle tetramer (24). Thus, although sequence reversal usually yields protein structures that lack thermodynamic stability (20, 23, 25, 26), the retro-leucine zipper demonstrates this observation is not absolute. Proteins with reversed sequence directionality have been reported to exhibit residual structure (26), form stable oligomers (27), and assemble into amyloid fibrils (28).

More recently, molecular dynamics methods and structure prediction algorithms were combined to simulate folding in reversed sequences from small proteins (29). Relative to random heteropolymer sequences, the preservation of patterns (e.g., polar, apolar, charge) and other properties (e.g., secondary structural propensities, hydrophobicity) in the protein sequence are thought to support folding. Results of that computational study indicated reversed sequences tend to retain preferences for native structural motifs, such as α-helices and β-hairpins, but suffer from steric effects on folding introduced by reversing the order of chiral residues.

Here, we utilize protein sequence reversal to gain experimental access to the denatured state ensemble to uncover intrinsic conformational properties of disordered sequences. Using staphylococcal nuclease as an experimental model, we obtained "Retro-nuclease" by reversing the sequence of the unaltered wild-type protein, "WT," to maintain an identical composition of L-amino acids and pattern of side chains despite the reversed sequence directionality. WT adopts a stable native structure consisting of three α-helices and a five-stranded, barrel-shaped β-sheet (30) and has been used as a model in protein stability and folding studies (31). Heat-denatured WT has good solubility and is not prone to aggregate (32), suggesting Retro-nuclease will be soluble at conditions required for structural characterization.

We report that Retro-nuclease is an elongated monomer and exhibits the structural characteristics of intrinsic disorder rather than adopting a compact denatured state. Unexpectedly for a disordered protein, Retro-nuclease exhibits sigmoid behavior in the temperature dependence of its hydrodynamic size with apparent baselines, a transition region, and a magnitude to the temperature-induced changes in mean hydrodynamic size that rivals heat denaturation of folded WT. Calorimetric and fluorescence examination of heat effects on Retro-nuclease structure report negligible hydration changes with temperature and argue against folding or hydrophobic collapse as the origin to the sigmoidicity. Sigmoid dependence of protein structure on an intensive property (e.g., temperature) usually indicates a cooperative phase change linked to folding/unfolding of a protein or protein domain (33). Results from CD difference spectroscopy and computer simulation suggest an alternative for the intrinsically disordered Retro-nuclease protein: a correlation to heat effects on intrinsic backbone conformational propensities.

Materials and Methods

Expression and purification of recombinant protein

Recombinant Retro-nuclease was expressed in bacterial cells and isolated to >95% purity from cell lysate using affinity chromatography as described elsewhere for the expression and purification of intrinsically disordered human p53(1–93) (32). WT nuclease was expressed in bacteria and isolated to >95% purity from cell lysate using ethanol precipitation and ion exchange chromatography (34). Bovine carbonic anhydrase, chicken albumin, and horse myoglobin were purchased from Sigma-Aldrich (St. Louis, MO) and further processed by ion-exchange chromatography to remove residual contaminants.

NMR spectroscopy

^{1}H-^{15}N heteronuclear single-quantum coherence (HSQC) NMR spectra were collected on a Bruker Avance III 400-MHz (Bruker, Billerica, MA) instrument using a double resonance probe equipped with a pulsed field z gradient. Spectra were processed using the Bruker TopSpin 3.5 software (Bruker). Measurements were at 28°C, the ambient probe temperature of the instrument. Samples contained 200 μM Retro-nuclease in 10 mM sodium phosphate, 100 mM sodium chloride, pH 7 or pH 6, as indicated.

Dynamic light scattering

Dynamic light scattering (DLS) readings used noninvasive backscatter optics and were measured with a Zetasizer Nano ZS using Peltier temperature control from Malvern Instruments (Malvern, Worcestershire, UK). Protein samples were buffered at pH 7 in 10 mM sodium phosphate, 100 mM sodium chloride, and filtered using 0.2 μm polyvinylidene difluoride syringe-driven filters. The sample temperature was cycled (5–65°C in 10°C steps and back) to establish the reversibility of heat effects on mean R_h. Samples were equilibrated at each temperature for 10 min before measurement. Fresh samples were prepared daily from frozen stock, because protein degradation was observed after ~48 h even in samples maintained at 4°C. All measurements used 1-cm pathlength quartz cuvettes. Mean R_h was calculated from the diffusion coefficient D, solvent viscosity η, and the Stokes-Einstein relationship $R_h = kT/6\pi\eta D$, where k is the Boltzmann constant and T is temperature in Kelvin. Solvent viscosity was calculated by the solvent builder program provided by Malvern, which uses Sednterp (35) to estimate η from the solution contents. The SD of the mean R_h from N number of measurements was calculated by $\boldsymbol{\sigma} = \sqrt{\sum_{i=1}^{N} (\boldsymbol{mean\ R_{h,i} - average\ mean\ R_h})^2/N}$, given in column 4 of Table S1. The SD of the size distribution (in terms of R_h) from N number of measurements was calculated by $\boldsymbol{propagated\ standard\ deviation/average\ mean} = \sqrt{\sum_{i=1}^{N} (\boldsymbol{standard\ deviation\ of\ distribution_i / mean\ of\ distribution_i})^2/N}$, given in column 5 of Table S1.

Size-exclusion chromatography

Size-exclusion chromatography (SEC) experiments used Sephadex G-75 (GE Healthcare, Piscataway, NJ) equilibrated in 10 mM sodium phosphate, 100 mM sodium chloride (pH 7). Elution volumes were determined from chromatograms measured using a Bio-Rad BioLogic LP System (Bio-Rad, Hercules, CA) equipped with an ultraviolet absorbance monitor (Bio-Rad). SEC experiments at temperatures other than room temperature used a water-jacketed chromatography column (Bio-Rad) that circulated refrigerated or heated water. Protein samples loaded onto the column had a volume of 80 μL and contained 2–3 mg mL^{-1} of protein in 10 mM sodium phosphate, 100 mM sodium chloride (pH 7). Indicator dyes were loaded separately

from the protein samples and contained 20 μL of 0.3 mg mL^{-1} blue dextran and 0.03 mg mL^{-1} 2,4-dinitrophenyl-L-aspartate to establish the void and column volumes, respectively. Thermodynamic retention factors (K_D) were calculated as $(V_e-V_o)/(V_c-V_o)$, where V_o is the void volume, V_c the column volume, and V_e the elution volume of the protein, determined from volumes at maximal absorbance. Uncertainties associated with K_D were calculated by the standard deviation from three or more measurements.

Fluorescence spectroscopy

Fluorescence measurements were performed using a Horiba Jobin-Yvon Fluorolog 3 spectrofluorometer (Horiba, Kyoto, Japan) equipped with a Wavelength Electronics Model LF1-3751 temperature controller (Wavelength Electronics, Bozeman, MT). Fluorescence emission spectra of 0.1 mg mL^{-1} WT and Retro-nuclease (in 10 mM sodium phosphate, 100 mM sodium chloride (pH 7)) as a function of temperature were measured between 314 and 394 nm with an excitation wavelength of 295 nm. Spectra were recorded between 10 and 80°C under slight stirring. The overall scan rate was 1.6°C min^{-1}, the step width was 2 nm, and the integration time was 0.5 s. The wavelength of maximal intensity was obtained by fitting each individual buffer-corrected spectrum to a Gaussian function as previously described (36). Acrylamide quenching of the tryptophan fluorescence for 0.1 mg mL^{-1} protein used an excitation wavelength of 295 nm. After the addition of buffered acrylamide solution, the protein sample was equilibrated for 5 min under slight stirring, and the fluorescence intensity was recorded for 20 s at 359 nm and averaged.

Differential scanning calorimetry

Differential scanning calorimetry (DSC) of 0.5 mg mL^{-1} WT and of 1 mg mL^{-1} Retro-nuclease was performed at a scan rate of 2°C min^{-1} on a TA Instruments NanoDSC (TA Instruments, New Castle, DE) at a constant pressure of 3 atm. Before any measurement, samples were degassed under moderate stirring. The instrument was equilibrated first with several scans between 10 and 95°C with buffer (10 mM sodium phosphate, 100 mM sodium chloride (pH 7)) loaded into the sample and reference capillaries. Then the protein sample was loaded into the sample capillary between two scans. The molar heat capacity was calculated as previously described (36).

Unfolding parameters determined by DSC for WT (i.e., transition enthalpy, ΔH, and midpoint temperature, T_m) were used to model WT unfolding by integration of the van 't Hoff equation, yielding $lnK = \Delta H/R \times (1/T_m - 1/T)$. The temperature dependence of WT mean R_h was modeled in Fig. 2 as $R_h = (R_{h,native} + R_{h,unfolded} \times K)/(1 + K)$, where $R_{h,native}$ and $R_{h,unfolded}$ were assigned the DLS-measured mean values at 25 and 65°C, respectively.

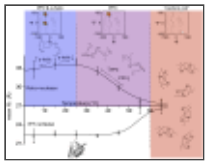

Figure 2

Temperature dependence of WT and Retro-nuclease mean R_h. The open circles show mean R_h measured by DLS for WT (*lower values*) and Retro-nuclease (*upper values*). The solid line for WT is a two-state fit using the transition enthalpy and midpoint temperature ...

CD spectroscopy

CD spectra were recorded using a Jasco (Oklahoma City, OK) J-710 spectropolarimeter equipped with a PFD-425S Peltier unit (Jasco, Easton, MD) and employed a 1-mm pathlength quartz cuvette. Samples were equilibrated at each temperature for 10 min. Spectra were collected with a resolution of 0.5 nm and a scan rate of 20 nm min^{-1} and were the average of eight scans. Reported spectra were baseline corrected for solvent and buffer contributions. Samples used 0.17 mg mL^{-1} protein (unless indicated otherwise for testing protein concentration effects), 10 mM sodium phosphate, 100 mM sodium chloride (pH 7).

Analytical ultracentrifugation

Sedimentation velocity (SV) experiments were performed on samples of Retro-nuclease buffered at pH 7 in 10 mM sodium phosphate and 100 mM sodium chloride. SV was used to measure the sedimentation coefficient (s) and the diffusion coefficient (D) of Retro-nuclease as a function of temperature. Both coefficients are inversely proportional to the frictional coefficient, f, by $\boldsymbol{s = M(1 - \bar{\nu}\rho)/(Nf)}$ and $D = RT/(Nf)$, where M is the molar mass, $\bar{\nu}$ the partial specific volume, ρ the solvent density, R the gas constant, T the temperature in Kelvin, and N is Avogadro's number. The hydrodynamic radius was determined from f and the solvent viscosity (η) using the Stokes-Einstein relationship $R_h = f/6\pi\eta$.

The absorbance of a stock solution was adjusted by dilution to an optical density of 0.6 absorbance units at 230 nm. Duplicate samples were measured at 50,000 rotations per minute and data collected by ultraviolet intensity at 230 nm with 20 s intervals in a Beckman Optima AUC analytical ultracentrifuge (Beckman Coulter Life Sciences, Indianapolis, IN) at the Center for Analytical Ultracentrifugation of Macromolecular Assemblies at the University of Texas Health San Antonio, using an An60Ti rotor and standard two-channel epon centerpieces from Beckman Coulter. All data were analyzed with UltraScan-III version 4.0, release 2420 (37, 38). Solvent density and solvent viscosity were estimated by UltraScan to be 1.00362 g/mL

and 1.00866 cP, respectively, at 20°C. Sedimentation and diffusion transport in the ultracentrifugation cell is described by the Lamm equation, which was solved using adaptive finite element methods (39, 40).

Experimental SV data were preprocessed by two-dimensional spectrum analysis (41) with simultaneous removal of radially and time-invariant noise contributions (42) and to fit the meniscus position as described in (43). The resulting data were further refined by parsimonious regularization using genetic algorithms (44). To obtain error statistics, Monte Carlo analysis was coupled with genetic algorithms in the final refinement (45). Monte Carlo results from duplicate samples were combined and weight averaged. The calculations are computationally intensive and were carried out on high-performance computing platforms (42, 46). All calculations were performed on the Lonestar-5 cluster at the Texas Advanced Computing Center at the University of Texas at Austin and on Comet at the San Diego Supercomputing Center.

Computer generation of disordered ensembles

Polyalanine structures were generated by a random search of conformational space using a hard sphere collision model (32, 47). This model uses van der Waals atomic radii (48, 49) as the only scoring function to eliminate grossly improbable conformations. The procedure to generate a random conformer starts with a unit peptide and all other atoms for a chain are determined by the rotational matrix (50). Backbone atoms are generated from the dihedral angles φ, ψ, and ω and the standard bond angles and bond lengths (51). Backbone dihedral angles are assigned randomly using a random number generator based on Knuth's subtractive method (52). (φ, ψ) is restricted to the allowed Ramachandran regions (53) to sample conformational space efficiently. To model intrinsic conformational propensities (e.g., for polyproline II (PPII) or α-helix), conformational bias is achieved by biasing the random search of backbone torsion angles (54, 55). For peptide bonds, ω was given a Gaussian fluctuation of ±5% about the *trans* form (180°). Of the two possible positions of the Cβ atom, the one corresponding to L-amino acid residues was used. Structures generated for each ensemble were population weighted using a structure-based energy function parameterized to solvent-accessible surface areas (56, 57, 58, 59, 60, 61, 62, 63, 64). Random structures were generated for an ensemble until L (defined in Fig. 6 legend) converged to a stable value (32). The interdependence of PPII and α-helix conformational sampling effects on mean R_h shown in Fig. 6 *B* was determined from simulation of polyalanine ensembles reported previously (55).

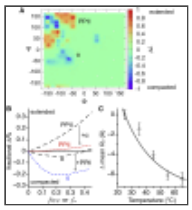

Figure 6

Mean hydrodynamic size dependence on backbone conformational bias. (*A*) The population-weighted size for 25-residue polyalanine was calculated by $L = \sum L_i \times P_i$, where L_i is the maximal Cα-Cα distance (Å) in ...

Predicting R_h from sequence

Equation 1 can be used to predict the mean R_h for disordered proteins from sequence using f_{PPII}, N, and Q (55). f_{PPII} is calculated from $\sum_{i=1}^{N} P_{PPII,i}/N$, where $P_{PPII,i}$ is the experimental PPII propensity for amino acid type i (65), and the summation is over the protein sequence containing N number of amino acids. Q is the net charge, determined by the number of lysine and arginine residues minus the number of glutamic acid and aspartic acid.

The temperature dependence of sequence-predicted mean R_h was calculated by applying a transition enthalpy, ΔH_{PPII}, to $P_{PPII,i}$. Integration of the van 't Hoff equation yields $lnK_{PPII}(T) - lnK_{PPII}(298K) = \Delta H_{PPII}/R \times (1/298 - 1/T)$. K_{PPII} is defined in the main text, and $K_{PPII}(T)$ refers to its temperature dependence.

Calculation of hydrophobicity index

Hydropathy plots were calculated from sequence using ProtScale at the ExPASy Bioinformatics Resource Portal (66).

Results and Discussion

Retro-nuclease sequence and structure characterization

The presence of intrinsic disorder in proteins and protein regions can be predicted from sequence with good confidence (67). To determine if sequence reversal substantially changes disorder predictions for the nuclease sequence, we analyzed the Retro-nuclease and WT sequences using the GeneSilico MetaDisorder service (68). This method generates a consensus prediction based on 13 primary methods that are weighted for accuracy. Analysis of the Retro-nuclease and WT sequences using disorder predictors indicate disorder tendencies that are similar but not identical for the two proteins (Fig. 1 *A*). Residues with a disorder tendency >0.5 are predicted to be disordered. 69 of 149 residues from the Retro-nuclease sequence and 58 of 149 from

the WT report a disorder tendency >0.5. High disorder tendencies are predicted in the N- and C-terminus regions of both sequences and mostly order for the remaining regions. MetaDisorder 3D, which uses fold-recognition methods (68), predicts low disorder tendencies for both sequences.

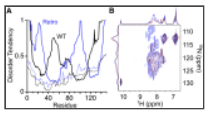

Figure 1

Disorder tendency predicted from WT and Retro-nuclease sequences. (*A*) The disorder tendency calculated using the GeneSilico MetaDisorder service (68) is shown. The solid lines show consensus prediction based on 13 primary methods for WT (*black*) and Retro-nuclease ...

Although disorder predictions for WT and Retro-nuclease sequences have some similarities, the Retro-nuclease protein is not structurally stable. The ^{1}H-^{15}N HSQC NMR spectrum of Retro-nuclease had poor peak dispersion (Fig. 1 *B*), indicating a disordered chain. WT, of course, folds into a stable, globular structure (30). The ^{1}H-^{15}N HSQC spectrum of a 131-residue fragment of nuclease with positions 4–12 and 141–149 excised, causing denaturation under folding conditions (12), shows individually resolved peaks for many amide groups that persist in 6 M urea (15). The correlation spectral differences from protein sequence reversal suggest the conformational ensembles of denatured nuclease and disordered Retro-nuclease have distinctly different structural characters.

The hydrodynamic size of Retro-nuclease is larger than WT under normal conditions (Fig. 2). DLS techniques report 22.4 ± 0.4 Å and 34.0 ± 0.5 Å for WT and Retro-nuclease mean R_h, respectively, at 25°C and pH 7. The error is the SD from repeat measurements (>20) of the mean to demonstrate that variability was small among replicate samples. The standard deviation of the apparent size distributions for WT and Retro-nuclease were broad (~5–7 Å), characteristic of translational Brownian motion used to infer hydrodynamic size (Fig. S1; Table S1). R_h determined by DLS is sensitive to the protein concentration (Fig. S1 *B*). The mean R_h measured at 0.5 mg mL^{-1} for a set of folded proteins agreed with crystal structure estimates of R_h (Fig. S1 *C*, *inset*), suggesting that measurements at this protein concentration offer good evaluation of structural size. Extrapolating the concentration dependence of mean R_h to zero protein concentration gave 33.9 Å for Retro-nuclease (Fig. S1 *D*), roughly matching the measurement at 0.5 mg mL^{-1}. Mean R_h in Fig. 2 for WT and Retro-nuclease were from DLS measurements at 0.5 mg mL^{-1} protein concentration.

Under similar solution conditions (pH 7, 20°C), sedimentation analysis by analytical ultracentrifugation (AUC) indicates Retro-nuclease is monomeric and has a high frictional ratio (~2), suggesting an elongated and nonglobular mean conformation (Fig. S2; Table S1). The mean R_h estimated from global analysis of replicate SV experiments was 33.0 Å. Confirming an extended structure, the K_D measured for Retro-nuclease by SEC indicated a R_h of 33.7 Å at room temperature (~23°C), determined from linear interpolation using globular proteins of known hydrodynamic size (Fig. S1 *C*). These hydrodynamic size measurements for WT and Retro-nuclease were consistent with scaling trends observed for folded proteins and IDPs, respectively. At 25°C, the mean R_h for WT matched R_h reported for other folded proteins of similar residue length N, and Retro-nuclease mean R_h agreed with the IDP trend for R_h dependence on N (Fig. S3).

Temperature effects on Retro-nuclease structure

Fig. 2 gives the temperature dependence of Retro-nuclease mean R_h from 5 to 65°C. Temperature-induced changes in mean R_h were fully reversible. Size distributions had gradual shifts with temperature (Fig. S1 *A*). Temperatures >25°C caused compaction, reducing mean R_h to 27.4 ± 0.6 Å at 65°C. For WT, increasing temperatures caused structural expansion, increasing mean R_h to 26.8 ± 1.0 Å at 65°C. Temperatures <25°C caused a slight compaction in Retro-nuclease structure when monitored by DLS, reducing mean R_h to 32.8 ± 1.3 Å at 5°C. AUC also reported a slight compaction in Retro-nuclease structure with decreasing temperatures in this general temperature range (Table S1). From 28 to 12°C, mean R_h decreased from 35.1 to 32.4 Å. A small increase in mean R_h to 33.2 Å, however, was observed at 4°C. At 37°C, the mean R_h for Retro-nuclease was 34.8 Å by AUC, which was the temperature limit of the instrument.

SEC chromatograms at temperatures from 5 to 65°C (Fig. 3) confirm that Retro-nuclease mean hydrodynamic size 1) depends on temperature, 2) is extended relative to WT under nondenaturing conditions, 3) is comparable to WT under heat-denaturing conditions, and 4) exhibits a broad sigmoidal shape in the K_D dependence on temperature. Briefly, Retro-nuclease K_D increased with temperature from 25 to 65°C, supporting the idea of heat-induced compaction. WT K_D decreased with temperature from 25 to 65°C, indicating structural expansion. At 65°C, chromatograms for WT and Retro-nuclease were similar with comparable K_D, suggesting similar hydrodynamic dimensions. At temperatures within the WT unfolding transition, two elution peaks were observed for WT. The peak of the smaller species gave a K_D at 45°C that was equivalent to folded WT and a K_D at 55°C that was in between folded and heat denatured. The other WT elution peak reported a K_D of approximately zero at both temperatures, indicating a species too large to discern by the column and presumed to be an aggregate. At 65°C, a near-zero K_D peak was not observed from WT elution. At 5°C, WT K_D was comparable to the values at 25, 35, and 45°C (Fig. 3 *B*). For Retro-nuclease, a slight decrease in K_D was observed at 5°C compared to 25°C. SEC thus reports a slight increase in mean R_h rather than a slight decrease, the opposite of the DLS observation. The three techniques, DLS, AUC, and SEC, were consistent in indicating any change in Retro-nuclease mean R_h from 5 to 25°C was small and on the order of ~1 Å. K_D and DLS-measured mean R_h showed some agreement and correlated linearly (Fig. 3 *C*; $R^2 = 0.87$), when compared at common temperatures and excluding the near-zero WT K_D.

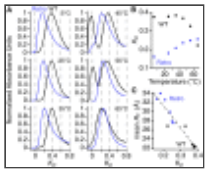

Figure 3

SEC analysis of hydrodynamic size. (*A*) Elution chromatograms of WT (*black*) and Retro-nuclease (*blue*) at temperatures from 5 to 65°C are shown. Proteins and indicator dyes were each run separately to resolve individual elution profiles. To correct ...

The increase in WT mean R_h from 45 to 65°C coincides with heat denaturation monitored by CD (32) and fluorescence (34) spectroscopies. WT has a single tryptophan residue at position 140. WT intrinsic tryptophan fluorescence has an emission maximum at ∼345 nm from excitation at 295 nm, indicating that the tryptophan side chain is buried (Fig. 4 *A*). The wavelength of emission maximum increased to ∼363 nm at temperatures >50°C for WT, consistent with denaturation-induced tryptophan solvation. For Retro-nuclease, intrinsic tryptophan fluorescence from excitation at 295 nm gave an emission maximum at a wavelength of ∼361 nm that was mostly independent of temperature from 15 to 80°C, demonstrating tryptophan solvation did not discernibly change with heat-induced changes in Retro-nuclease hydrodynamic size. Confirming that the tryptophan side chain in Retro-nuclease is solvated, the Stern-Volmer constant, K_{SV}, which quantifies solute quenching of fluorescence emission, was 12.0, 13.4, 12.7, and 12.6 M^{-1} at 1, 20, 40, and 60°C, respectively (Fig. S4). WT K_{SV} was 5.8, 7.4, 9.8, and 13.7 M^{-1} at 1, 20, 40, and 60°C, respectively, suggesting tryptophan is partially exposed in WT under native conditions. K_{SV} is similar for both proteins at 60°C, confirming similar environments for the tryptophan sidechain at denaturing temperatures. K_{SV} for N-acetyl tryptophanamide was 21.8, 23.9, 26.7, and 29.0 M^{-1} at 1, 20, 40, and 60°C, respectively. The comparable intrinsic fluorescence properties, tryptophan solvent exposure, and mean hydrodynamic size for WT and Retro-nuclease at 65°C suggests that heat-induced compaction of Retro-nuclease structure involves a transition from a highly expanded and disordered state at 25°C to a more compact but similarly disordered state at 65°C that has some resemblance to heat-denatured WT.

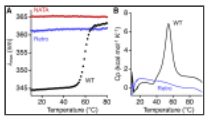

Figure 4

Temperature dependence of intrinsic tryptophan fluorescence and constant pressure heat capacity. (*A*) N-acetyl tryptophanamide (NATA) was used to represent solvated tryptophan. Shown is the wavelength of emission maximum (λ_{max}) from the intrinsic ...

The temperature dependence of the constant pressure heat capacity (Cp) of Retro-nuclease, measured by DSC, exhibited a mostly constant slope with minimal inflection or excess heat capacity over the temperature range of 10–95°C (Fig. 4 *B*). Because Cp is the temperature derivative of the enthalpy, a temperature-induced change in state cannot occur without a concomitant change in the heat capacity curve (69). Cp changes resulting from protein conformational change arise primarily from hydration contributions (61). The lack of a perceptible shift in the slope of the Cp dependence on temperature from 25 to 65°C, coinciding with temperature-induced R_h compaction, thus shows that hydration changes with temperature were small for Retro-nuclease and, overall, absent an apparent change in thermodynamic state. In contrast, WT demonstrated an excess heat capacity, with a transition enthalpy of 54.9 kcal mol^{-1} at 54°C. Comparison of the temperature dependence of WT mean R_h to a two-state model using the van 't Hoff equation, $d(lnK)/d(1/T) = -\Delta H/R$, and the DSC-determined values for the transition enthalpy (ΔH) and midpoint temperature is provided in Fig. 2, showing good agreement. In the van 't Hoff equation, T is temperature, R the gas constant, and K the two-state equilibrium constant for the native and thermally denatured states.

The structure of Retro-nuclease was also analyzed by CD spectroscopy. At 25°C, the Retro-nuclease spectrum is similar in character to spectra reported for IDPs (4), with molar residue ellipticity (MRE) at 221 nm close to zero (Fig. 5 *A*). Based on a survey of CD spectra from ∼100 proteins, it has been argued that IDPs can be categorized as "mostly unfolded" if MRE (× 10^{-3}) at 200 and 222 nm are from −21.7 to −16.1 and −2.4 to −1.0, respectively, or consistent with the "existence of some residual secondary structure" if the values at 200 and 222 nm are from 12.0 to −19.4 and −5.0 to −2.8 (4). Using that dichotomy, the Retro-nuclease CD spectrum at 25°C is showing signs of some residual secondary structure.

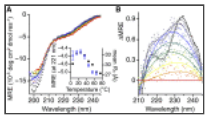

Figure 5

Temperature dependence of Retro-nuclease CD spectrum. (*A*) The spectrum at 5°C (*black*), 15°C (*gray*), 25°C (*purple*), 35°C (*blue*), 45°C (*blue-green*), 55°C (*green*), 65°C (*yellow*), 75°C ...

The temperature dependence of Retro-nuclease mean R_h is supported by temperature effects on its CD spectrum. Changes in MRE at 221 nm track with changes in DLS-measured mean R_h (Fig. 5 *A inset*), again exhibiting sigmoid dependence with temperature. The Retro-nuclease spectrum and its temperature dependence were independent of protein concentration from 0.11 to 0.23 mg mL^{-1}. CD measurements at different protein concentrations each demonstrate similar agreement between changes in DLS-measured mean R_h and changes in MRE at 221 nm (Fig. S5). Spectrum analysis was performed by fitting the experimental data to secondary structure basis components using the BeStSel Web Server (70). This analysis predicts negligible changes in secondary structure content for Retro-nuclease from 5 to 85°C (Fig. S6). The main structure determination by this method was

for the "other" classification that contains 3,10-helix, π-helix, β-bridge, bends, loops/irregular, and invisible regions. Deconvolution of CD spectra into specific structural elements is difficult for Retro-nuclease because no satisfactory reference set exists for disordered proteins, owing to the spectral similarity of denatured structures (71).

Additional inspection of CD results by monitoring difference spectra relative to the 85°C spectrum is provided in Fig. 5 *B*. From 25 to 85°C, Retro-nuclease shows a local CD maximum at ~225 nm that decreases in intensity with increasing temperature. A local CD maximum at ~225 nm is seen in proteins with PPII structures (72). PPII conformational sampling is prominent in unfolded proteins (72) and known to decrease with increasing temperature (73). The temperature dependence of the Retro-nuclease CD maximum at ~225 nm thus indicates PPII conformational sampling at normal temperatures that is reduced in magnitude as it is heated, a common observation for disordered proteins (32, 74).

Lowering the temperature to 5°C induced a local CD minimum at ~222 nm (Fig. 5 *B*). Local minima at ~222 and ~208 nm are observed in proteins with α-helix (75). 3,10-helix, type I, and type III β-turns also show a local CD minimum at ~222 nm (76). Conformational parameters for residues in 3,10-helix, type I, and type III β-turn structures can be close to the α-helix when comparing backbone dihedral angles φ and ψ. (φ, ψ) close to the α-helix mean (77) is not uncommon for internal residues in type I/III turn structures and 3,10-helix (78). The decrease in Retro-nuclease MRE at 208 nm and 5°C relative to 25°C (Fig. S7) suggests the temperature-induced changes in structure at 5°C are not from an increase in type II β-turn conformational sampling. Type II β-turns produce a positive band at ~208 nm (79). The local minimum at ~222 nm observed in the difference spectrum of Retro-nuclease at 5°C (and 15°C) accordingly suggests low-temperature sampling of backbone conformations similar to α-helix. Based on the spectral indications of temperature-sensitive PPII and α-helix content (Fig. 5 *B*), as well as correlated temperature-induced changes in MRE at 221 nm and R_h (Fig. 5 *A*, *inset*), we investigated by computer simulation the question of whether temperature-dependent backbone sampling of PPII and α-helix conformations could describe heat effects on Retro-nuclease hydrodynamic size.

Correlation of mean R_h to backbone conformational propensities for PPII and α-helix

The conformational ensemble for disordered and/or denatured proteins is characterized by rapid conformational sampling. Bias in the transient sampling of backbone dihedral angles (φ, ψ) has the ability to modulate mean hydrodynamic size (54, 55). This idea is illustrated in Fig. 6 *A*, using ensembles of disordered states simulated by a hard sphere collision model that generates protein structures randomly (32, 47). Polyalanine was used to minimize the influence of charge in the simulations and focus attention on backbone conformational bias effects. Shown are results from varying the region in a Ramachandran plot where a sampling bias was artificially applied, increasing the population weight to 10% in targeted 20 × 20° areas. Red and blue colored regions indicate (φ, ψ) sampling biases that generated ensembles of extended or compacted structures, respectively, relative to a polyalanine ensemble that had unbiased (φ, ψ) sampling.

In the simulations, hydrodynamic size for a computer-generated structure was calculated as the maximal Cα-Cα distance (L), used to approximate the axial length of its tumbling volume. For folded proteins, one-half L (i.e., L/2) calculated from crystal structures agrees with experimental R_h (Fig. S1 *C*, *inset*). For IDPs, L/2 calculated from simulated ensembles provides good agreement with experimental R_h when intrinsic chain propensities for PPII (54) and net charge effects (55) are accounted for.

Retro-nuclease difference CD spectra (Fig. 5 *B*) suggest temperature-dependent sampling of PPII and α-helix backbone conformations. IDPs are known to transiently populate α-helix (80), and PPII is the dominant backbone conformation in unfolded peptides (72); thus, the Retro-nuclease CD spectra are reasonable for a disordered protein. Applying a backbone sampling bias for PPII when simulating ensembles, with (φ, ψ) for PPII defined as (−75 ± 10°, 145 ± 10°) (81), produced the black dashed line labeled "PPII" in Fig. 6 *B*. Increasing the sampling bias for PPII generates disordered ensembles with increasingly larger mean hydrodynamic size (54). The simulation-derived relationship between mean R_h and backbone conformational bias for PPII, modeled as an intrinsic PPII propensity, was tested experimentally using IDPs (54, 55, 82). Under normal conditions, experimental mean R_h from IDPs was shown to follow

$$R_h = (2.16\text{Å}) \times N^{(0.503-0.11\times ln(1-f_{PPII}))} + 0.25 \times |Q| - 0.31 \times N^{0.5}, \quad (1)$$

where N is the number of residues, Q the net charge, and f_{PPII} the population-weighted fractional number of residues in the PPII backbone conformation (55).

The first term in Eq. 1 is structural expansion with increasing PPII propensity, reproducing the black dashed line labeled "PPII" in Fig. 6 *B*. The remaining terms, $0.25 \times |Q| - 0.31 \times N^{0.5}$, represent apparent net charge effects on mean R_h (55). Q is estimated from sequence as the number of lysine and arginine residues minus the number of glutamic and aspartic acid residues. f_{PPII} is estimated from sequence using experimental PPII propensities that were measured from substitution-induced changes in the binding energetics of the Son of Sevenless (Sos) peptide with its substrate, the *Caenorhabditis elegans* Sem-5 SH3 domain. A noninteracting residue (83) that adopts PPII in the bound conformation (84) was chosen as the peptide substitution site for host-guest studies that measured intrinsic PPII propensities for the common amino acids (65). The average error predicting mean R_h from sequence using Eq. 1 was ±1.7 Å for a database of IDPs (55). Retro-nuclease mean R_h predicted from sequence was 33.1 Å when using Eq. 1, within the IDP average error when compared to mean R_h measured for Retro-nuclease by DLS, SEC, and AUC at normal or near-normal

temperatures. Accordingly, mean R_h measured experimentally for Retro-nuclease was consistent with its sequence length and composition based on IDP-established trends.

PPII conformational sampling in short, unfolded peptides is quantified by a transition enthalpy (73). To model temperature effects on PPII conformational sampling in Retro-nuclease, and by extrapolation the mean R_h, we again use the van 't Hoff equation, $dlnK_{PPII}/d(1/T) = -\Delta H_{PPII}/R$. Here, K_{PPII} is the equilibrium defined by the intrinsic PPII propensity, $K_{PPII} = (1 - P_{PPII})/P_{PPII}$, and ΔH_{PPII} is the transition enthalpy. In this model, P_{PPII} is the experimental intrinsic PPII propensity (65). K_{PPII} varies with residue position, depending on amino acid type. ΔH_{PPII} is constant. Changes in P_{PPII} with temperature, modeled at each residue position by the van 't Hoff equation, predict changes in mean R_h by Eq. 1, assuming net charge effects are temperature independent. This model does not correct for nearest-neighbor effects on intrinsic P_{PPII} (85), and its use is simply to demonstrate that the magnitude of mean R_h changes observed for Retro-nuclease, although large, is consistent with known heat effects on peptide conformation. To isolate heat effects on PPII, temperatures <25°C were not included in this analysis because the Retro-nuclease difference CD spectrum implies population of α-helix, or (φ, ψ) conformational sampling closely resembling α-helix, in addition to PPII conformational sampling at those temperatures. At low temperatures, additivity effects that combine PPII and α-helix conformational sampling should be considered, as discussed below.

Fig. 6 *C* shows that the temperature dependence of Retro-nuclease mean R_h at $T > 25$°C can be modeled using 10 kcal mol^{-1} for ΔH_{PPII}. Kallenbach and colleagues calculated $\Delta H_{PPII} \sim 10$ kcal mol^{-1} for alanine peptides from CD measurements and ~13 kcal mol^{-1} from NMR measurements (73), comparable to our results from modeling the temperature dependence of Retro-nuclease mean R_h. Kallenbach argued that assigning temperature-dependent behavior exclusively to heat effects on PPII sampling will overestimate ΔH_{PPII}, which is an important consideration for this analysis, too.

Extrapolating the temperature dependence of PPII backbone conformational sampling to low temperatures, using 10 kcal mol^{-1} for ΔH_{PPII}, predicts that Retro-nuclease mean R_h will increase by 12 Å from 25 to 5°C, which was not observed experimentally (Fig. 2). This demonstrates the PPII-coil transition is unable to describe the full temperature dependence of mean R_h from low to high temperatures and misses the observed sigmoidicity. At low temperatures, Retro-nuclease CD spectra indicate α-helix backbone conformational sampling in addition to PPII (Fig. 5 *B*). Returning to Fig. 6 *B*, a (φ, ψ) sampling bias at (−64 ± 10°, −41 ± 10°), the α-helix mean (77), in ensemble simulations produced the black dashed line labeled "α" (55). Additivity effects are shown by red and blue dashed lines and demonstrate a key result. Specifically, the effect on mean hydrodynamic size from α-helix and PPII conformational sampling biases predicts interdependent behavior that is correlated to the magnitude of the biases. Simple additivity, whereby the combined effect of PPII and α-helix conformational sampling on hydrodynamic size is the sum of the biases applied in isolation (i.e., the sum of values taken from the *black dashed lines* in Fig. 6 *B*), was not observed.

Ensemble simulations show that a modest increase in α-helix conformational sampling to ~0.05 from the background rate of ~0.02 reduced the size dependence on PPII bias (Fig. 6 *B*, *red line*). α-helix conformational sampling at even higher rates reversed the size dependence on PPII bias, whereby increasing PPII conformational sampling correlated with decreasing mean hydrodynamic size (55). For the reciprocal relationship, increasing PPII conformational sampling from its background rate of ~0.02 to ~0.4, matching the average PPII propensity in an IDP database (55), enhanced the sensitivity of mean hydrodynamic size to intrinsic α-helix bias (Fig. 6 *B*, *blue line*). Overall, the ensemble simulations show that α-helix conformational sampling reduces the ability of PPII bias to elongate disordered structures (Fig. 6 *B*, *red line*) and compacts mean R_h (Fig. 6 *B*, *blue line*), providing a reasonable hypothesis to describe the effects of cold temperatures on Retro-nuclease mean R_h at which a leveling off or slight compaction was observed.

Conclusions

Reversing sequence order can destabilize a protein that otherwise would fold (20). To investigate the folding behavior of a reversed protein sequence, we reversed the sequence of staphylococcal nuclease, a protein that folds into a stable, globular structure (30). Spectroscopic, calorimetric, and hydrodynamic techniques reveal Retro-nuclease to be an elongated monomer and intrinsically disordered. At normal temperatures, Retro-nuclease is best described as an ensemble of disordered structures with backbone sampling preferences for PPII (Fig. 2). At low temperatures, an additional preference for (φ, ψ) sampling resembling α-helix was observed. At high temperatures, Retro-nuclease resembles a "random coil," with this term referencing disordered chains with no strongly preferred backbone conformations.

The burial of apolar surfaces provides a large source of free energy that favors protein folding under normal conditions (86). The formation of intramolecular hydrogen bonds between peptide groups is also a key contributor to the net folding free energy (87), as is peptide solvation that includes hydrogen bonds between water and peptide -NH and -C=O groups (88). The energetic penalty of an unsatisfied hydrogen bond is estimated to be 6 kcal mol^{-1}, indicating that free hydrogen bonding groups are unlikely (88). From polymer physics (89), solvent quality is classified as "poor" if the solution promotes intramolecular interactions in the polymer, like the hydrogen bond, to yield compacted structures. "Good" solvents favor intermolecular interactions and extended polymer conformations. The structural size of a polymer in solution thus reports on the net balance of intra- and intermolecular interactions for a polymer-solvent combination. Our experiments show that Retro-nuclease does not experience a hydrophobic collapse of structure under normal conditions, consistent with the idea that water is generally a good solvent for disordered proteins (90). Unlike Retro-nuclease, many nuclease mutants that maintain the forward sequence direction adopt compact denatured states (6, 91), indicating that

hydrophobic collapse in the WT and structural compaction in denatured nuclease depend on sequence directionality. Sequence patterns of charge, hydrophobicity, hydrogen bonding, and other amino acid physicochemical properties contribute to the folding mechanisms and structural collapse of the polypeptide chain. The observation that, under normal conditions, WT folds and Retro-nuclease adopts a highly extended and disordered structure also demonstrates a role for sequence direction.

Some collapse of Retro-nuclease structure was observed at high temperatures. Mean R_h for Retro-nuclease compacted by ∼20% from 25 to 65°C, representing a 50% reduction in hydrodynamic volume. Despite these temperature-induced effects on hydrodynamic size, DSC and fluorescence spectroscopy indicate hydration changes with temperature were small for Retro-nuclease (i.e., in comparison to WT). The hydrodynamic size of Retro-nuclease at 65°C is indistinguishable from thermally denatured WT (Figs. 2 and 3), arguing against heat-induced folding and supporting the idea of a well-solvated high-temperature state. Of note, a hydropathy plot of the Retro-nuclease sequence is similar to other well-studied folded proteins, and its charge per residue is somewhat higher (Fig. S8). The average per-residue hydrophobicity index from using the Kyte-Doolittle scale (92) was −0.84 ± 1.03 for Retro-nuclease and −0.42 ± 0.95 for lysozyme, myoglobin, dihydrofolate reductase, and trypsin inhibitor combined. Because the nuclease sequence appears normal for a folded protein, some Retro-nuclease properties and structural features may be found in other sequence-reversed proteins. Indeed, sequence reversal of bacteriophage T5 endolysin produced a protein that, like Retro-nuclease, did not fold or undergo hydrophobic collapse (25).

The temperature dependence of Retro-nuclease mean hydrodynamic size exhibits a sigmoid shape normally associated with cooperative phase transitions and a magnitude to the mean R_h changes that exceeds heat denaturation of WT. The temperature dependence of the heat capacity did not reveal an excess enthalpy for Retro-nuclease, suggesting noncooperative structure changes. The possibility of cooperative Retro-nuclease folding cannot be eliminated. α-helix formation in peptides report transition enthalpies of ∼1 kcal mol^{-1} res^{-1} (93), which may be below detection limits of our DSC experiments. The Retro-nuclease CD difference spectrum indicated (φ, ψ) conformational sampling resembling α-helix at low temperatures. The change in Retro-nuclease mean R_h from 25 to 65°C, however, was consistent with intrinsic conformational propensities for PPII measured in peptides, in amount (65) and apparent transition enthalpy (73). Fig. 6 *C* shows that heat-induced compaction can be modeled quantitatively as a transition from PPII-biased disordered structures to a more uniform sampling of the accessible (φ, ψ) regions at high temperatures. The ability of intrinsic PPII propensities to predict mean R_h for Retro-nuclease and a database of IDPs (54, 55) and reasonably approximate heat sensitivity from sequence (Fig. 6 *C*) is evidence that the structural features of Retro-nuclease have a noncooperative origin at temperatures ≥25°C and sample backbone conformations according to intrinsic chain properties.

Studies using other proteins and peptides support the hypothesis that cold temperatures stabilize α-helix, whereas normal temperatures generally disfavor intrinsic α-helix conformational sampling. For instance, the C-peptide (residues 1–13 of ribonuclease A, containing the N-terminal α-helix) forms α-helix at low temperatures that otherwise is absent (94). Short alanine-based peptides with charge substitutions form a stable helix at 1°C and transition to coil with increasing temperature (95). Cold denaturation of ubiquitin is noncooperative and produces partially folded structures that adopt α-helix at very low temperatures (96). Also, simulation of disordered ensembles shows that α-helix conformational sampling at rates ≥0.05 and hydrodynamic size that scales by Eq. 1 are mutually exclusive (Fig. 6 *B*, *red line*). A survey of many IDPs established that mean R_h indeed scales by Eq. 1 under normal conditions (55), indicating low α-helix conformational sampling rates. Combined, these observations support two interesting conclusions. First, backbone conformational bias correlates with mean hydrodynamic size in disordered proteins. This indicates an accounting of conformational bias can be inferred from equilibrium measurements of mean R_h. Second, the conformational ensembles for Retro-nuclease and IDPs are highly similar under normal conditions. This indicates that sequence compositions from folded proteins can yield disorder properties that mimic biological IDPs.

Author Contributions

L.R.E., A.T., B.D., and S.T.W. performed research. A.K.D., B.D., and S.T.W. analyzed data. S.T.W. designed research. S.T.W. wrote the manuscript.

Acknowledgments

We thank Dr. Ben A. Shoulders (Texas State University) for technical assistance in collecting NMR spectra of Retro-nuclease.

Supported by National Institutes of Health grants R15GM115603 (S.T.W.) and GM120600 (B.D.), and National Science Foundation grants DMR-1205670 (WJ Brittain), NSF-ACI-1339649 (B.D.), and TG-MCB070039 (B.D.). Sedimentation experiments were supported by National Cancer Institute grant P30 CA054174 to the San Antonio Cancer Institute.

Notes

Editor: James Cole.

Footnotes

Eight figures and one table are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(18)30681-7.

Supporting Citations

References (97, 98, 99) appear in the Supporting Material.

Supporting Material

Document S1. Figs. S1–S8 and Table S1:
Click here to view.(1.1M, pdf)

Document S2. Article plus Supporting Material:
Click here to view.(2.5M, pdf)

Article information

Biophys J. 2018 Jul 17; 115(2): 328–340.
Published online 2018 Jul 17. doi: 10.1016/j.bpj.2018.06.006
PMCID: PMC6050754
PMID: 30021108

Lance R. English,[1] Alexander Tischer,[2] Aysha K. Demeler,[3] Borries Demeler,[4] and Steven T. Whitten[1,*]

[1]Department of Chemistry and Biochemistry, Texas State University, San Marcos, Texas
[2]Department of Internal Medicine, Mayo Clinic, Rochester, Minnesota
[3]Department of Chemistry, Bergische Universität Wuppertal, Wuppertal, North Rhine-Westphalia, Germany
[4]Department of Chemistry and Structural Biology, University of Texas Health San Antonio, San Antonio, Texas
Steven T. Whitten: sw50@txstate.edu
*Corresponding author sw50@txstate.edu

Received 2018 Feb 27; Accepted 2018 Jun 4.

References

1. Wootton J.C., Federhen S. Statistics of local complexity in amino acid sequences and sequence databases. Comput. Chem. 1993;17:149–163. [Google Scholar]

2. Wootton J.C. Non-globular domains in protein sequences: automated segmentation using complexity measures. Comput. Chem. 1994;18:269–285. [PubMed] [Google Scholar]

3. Romero P., Obradovic Z., Dunker A.K. Sequence complexity of disordered protein. Proteins. 2001;42:38–48. [PubMed] [Google Scholar]

4. Uversky V.N. Natively unfolded proteins: a point where biology waits for physics. Protein Sci. 2002;11:739–756. [PMC free article] [PubMed] [Google Scholar]

5. Dunker A.K., Obradovic Z., Brown C.J. Intrinsic protein disorder in complete genomes. Genome Inf. Ser. Workshop Genome Inform. 2000;11:161–171. [PubMed] [Google Scholar]

6. Flanagan J.M., Kataoka M., Engelman D.M. Truncated staphylococcal nuclease is compact but disordered. Proc. Natl. Acad. Sci. USA. 1992;89:748–752. [PMC free article] [PubMed] [Google Scholar]

7. Kohn J.E., Gillespie B., Plaxco K.W. Non-sequence-specific interactions can account for the compaction of proteins unfolded under “native” conditions. J. Mol. Biol. 2009;394:343–350. [PMC free article] [PubMed] [Google Scholar]

8. Nöppert A., Gast K., Damaschun G. Reduced-denatured ribonuclease A is not in a compact state. FEBS Lett. 1996;380:179–182. [PubMed] [Google Scholar]

9. Choy W.Y., Mulder F.A., Kay L.E. Distribution of molecular size within an unfolded state ensemble using small-angle X-ray scattering and pulse field gradient NMR techniques. J. Mol. Biol. 2002;316:101–112. [PubMed] [Google Scholar]

10. Kohn J.E., Millett I.S., Plaxco K.W. Random-coil behavior and the dimensions of chemically unfolded proteins. Proc. Natl. Acad. Sci. USA. 2004;101:12491–12496. [PMC free article] [PubMed] [Google Scholar]

11. Bezsonova I., Singer A., Forman-Kay J.D. Structural comparison of the unstable drkN SH3 domain and a stable mutant. Biochemistry. 2005;44:15550–15560. [PubMed] [Google Scholar]

12. Alexandrescu A.T., Abeygunawardana C., Shortle D. Structure and dynamics of a denatured 131-residue fragment of staphylococcal nuclease: a heteronuclear NMR study. Biochemistry. 1994;33:1063–1072. [PubMed] [Google Scholar]

13. McCarney E.R., Kohn J.E., Plaxco K.W. Is there or isn't there? The case for (and against) residual structure in chemically denatured proteins. Crit. Rev. Biochem. Mol. Biol. 2005;40:181–189. [PubMed] [Google Scholar]

14. Ohnishi S., Lee A.L., Shortle D. Direct demonstration of structural similarity between native and denatured eglin C. Biochemistry. 2004;43:4064–4070. [PubMed] [Google Scholar]

15. Wang Y., Shortle D. The equilibrium folding pathway of staphylococcal nuclease: identification of the most stable chain-chain interactions by NMR and CD spectroscopy. Biochemistry. 1995;34:15895–15905. [PubMed] [Google Scholar]

16. Kazmirski S.L., Wong K.B., Daggett V. Protein folding from a highly disordered denatured state: the folding pathway of chymotrypsin inhibitor 2 at atomic resolution. Proc. Natl. Acad. Sci. USA. 2001;98:4349–4354. [PMC free article] [PubMed] [Google Scholar]

17. Klein-Seetharaman J., Oikawa M., Schwalbe H. Long-range interactions within a nonnative protein. Science. 2002;295:1719–1722. [PubMed] [Google Scholar]

18. Sugase K., Dyson H.J., Wright P.E. Mechanism of coupled folding and binding of an intrinsically disordered protein. Nature. 2007;447:1021–1025. [PubMed] [Google Scholar]

19. Bah A., Vernon R.M., Forman-Kay J.D. Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. Nature. 2015;519:106–109. [PubMed] [Google Scholar]

20. Lacroix E., Viguera A.R., Serrano L. Reading protein sequences backwards. Fold. Des. 1998;3:79–85. [PubMed] [Google Scholar]

21. Guptasarma P. Reversal of peptide backbone direction may result in the mirroring of protein structure. FEBS Lett. 1992;310:205–210. [PubMed] [Google Scholar]

22. Olszewski K.A., Kolinski A., Skolnick J. Does a backwardly read protein sequence have a unique native state? Protein Eng. 1996;9:5–14. [PubMed] [Google Scholar]

23. Ruvo M., Fassina G. Folding behaviour of retro-rubredoxin. Protein Pept. Lett. 1996;3:241–246. [Google Scholar]

24. Mittl P.R., Deillon C., Grütter M.G. The retro-GCN4 leucine zipper sequence forms a stable three-dimensional structure. Proc. Natl. Acad. Sci. USA. 2000;97:2562–2566. [PMC free article] [PubMed] [Google Scholar]

25. Kutyshenko V.P., Mikoulinskaia G.V., Uversky V.N. Structure and dynamics of the retro-form of the bacteriophage T5 endolysin. Biochim. Biophys. Acta. 2016;1864:1281–1291. [PubMed] [Google Scholar]

26. Kutyshenko V.P., Prokhorov D.A., Uversky V.N. Dancing retro: solution structure and micelle interactions of the retro-SH3-domain, retro-SHH-'Bergerac' J. Biomol. Struct. Dyn. 2014;32:257–272. [PubMed] [Google Scholar]

27. Ahmed S., Shukla A., Guptasarma P. Folding behavior of a backbone-reversed protein: reversible polyproline type II to beta-sheet thermal transitions in retro-GroES multimers with GroES-like features. Biochim. Biophys. Acta. 2008;1784:916–923. [PubMed] [Google Scholar]

28. Shukla A., Raje M., Guptasarma P. A backbone-reversed all-beta polypeptide (retro-CspA) folds and assembles into amyloid nanofibres. Protein Eng. 2003;16:875–879. [PubMed] [Google Scholar]

29. Zhang Y., Weber J.K., Zhou R. Folding and stabilization of native-sequence-reversed proteins. Sci. Rep. 2016;6:25138. [PMC free article] [PubMed] [Google Scholar]

30. Hynes T.R., Fox R.O. The crystal structure of staphylococcal nuclease refined at 1.7 A resolution. Proteins. 1991;10:92–105. [PubMed] [Google Scholar]

31. Shortle D. The denatured state (the other half of the folding equation) and its role in protein stability. FASEB J. 1996;10:27–34. [PubMed] [Google Scholar]

32. Langridge T.D., Tarver M.J., Whitten S.T. Temperature effects on the hydrodynamic radius of the intrinsically disordered N-terminal region of the p53 protein. Proteins. 2014;82:668–678. [PubMed] [Google Scholar]

33. Lumry R., Biltonen R. Validity of the “two-state” hypothesis for conformational transitions of proteins. Biopolymers. 1966;4:917–944. [PubMed] [Google Scholar]

34. Whitten S.T., García-Moreno E.B. pH dependence of stability of staphylococcal nuclease: evidence of substantial electrostatic interactions in the denatured state. Biochemistry. 2000;39:14292–14304. [PubMed] [Google Scholar]

35. Laue T.M., Shah B.D., Pelletier S.L. Computer-aided interpretation of analytical sedimentation data for proteins. In: Harding S., Rowe A., Horton J.C., editors. Analytical Ultracentrifugation in Biochemistry and Polymer Science. Royal Society of Chemistry; 1992. pp. 90–125. [Google Scholar]

36. Tischer A., Auton M. Urea-temperature phase diagrams capture the thermodynamics of denatured state expansion that accompany protein unfolding. Protein Sci. 2013;22:1147–1160. [PMC free article] [PubMed] [Google Scholar]

37. Demeler, B., G. Gorbet, …, W. Cao. 2016. UltraScan-III version 3.5: A comprehensive data analysis software package for analytical ultracentrifugation experiments. http://www.ultrascan3.uthscsa.edu/.

38. Demeler B., Gorbet G. Analytical ultracentrifugation data analysis with ultraScan-III. In: Uchiyama S., Stafford W.F., Laue T., editors. Analytical Ultracentrifugation: Instrumentation, Software, and Applications. Springer; 2016. pp. 119–143. [Google Scholar]

39. Cao W., Demeler B. Modeling analytical ultracentrifugation experiments with an adaptive space-time finite element solution of the Lamm equation. Biophys. J. 2005;89:1589–1602. [PMC free article] [PubMed] [Google Scholar]

40. Cao W., Demeler B. Modeling analytical ultracentrifugation experiments with an adaptive space-time finite element solution for multicomponent reacting systems. Biophys. J. 2008;95:54–65. [PMC free article] [PubMed] [Google Scholar]

41. Brookes E., Cao W., Demeler B. A two-dimensional spectrum analysis for sedimentation velocity experiments of mixtures with heterogeneity in molecular weight and shape. Eur. Biophys. J. 2010;39:405–414. [PubMed] [Google Scholar]

42. Schuck P., Demeler B. Direct sedimentation analysis of interference optical data in analytical ultracentrifugation. Biophys. J. 1999;76:2288–2296. [PMC free article] [PubMed] [Google Scholar]

43. Demeler B. Methods for the design and analysis of sedimentation velocity and sedimentation equilibrium experiments with proteins. Cur. Protoc. Prot. Sci. 2010;Chapter 7 Unit 7.13. [PMC free article] [PubMed] [Google Scholar]

44. Brookes, E. H., and B. Demeler. 2007. Parsimonious regularization using genetic algorithms applied to the analysis of analytical ultracentrifugation experiments. *Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation.* ACM 978–1-59593–697–4, pp. 361–368.

45. Demeler B., Brookes E. Monte Carlo analysis of sedimentation experiments. Colloid Polym. Sci. 2008;286:129–137. [Google Scholar]

46. Brookes E.H., Demeler B. Parallel computational techniques for the analysis of sedimentation velocity experiments in UltraScan. Colloid Polym. Sci. 2008;286:138–148. [Google Scholar]

47. Whitten S.T., Yang H.W., Hilser V.J. Exploring the impact of polyproline II (PII) conformational bias on the binding of peptides to the SEM-5 SH3 domain. Protein Sci. 2008;17:1200–1211. [PMC free article] [PubMed] [Google Scholar]

48. Ramachandran G.N., Ramakrishnan C., Sasisekharan V. Stereochemistry of polypeptide chain configurations. J. Mol. Biol. 1963;7:95–99. [PubMed] [Google Scholar]

49. Iijima H., Dunbar J.B., Jr., Marshall G.R. Calibration of effective van der Waals atomic contact radii for proteins and peptides. Proteins. 1987;2:330–339. [PubMed] [Google Scholar]

50. Jeffreys H., Jeffreys B.S. Cambridge University Press; New York: 1950. Methods of Mathematical Physics; pp. 122–123. [Google Scholar]

51. Momany F.A., McGuire R.F., Scheraga H.A. Energy parameters in polypeptides. VII. Geometric parameters, partial atomic charges, nonbonded interactions, hydrogen bond interactions, and intrinsic torsional potentials for the naturally occurring amino acids. J. Phys. Chem. 1975;79:2361–2381. [Google Scholar]

52. Knuth D.E. Second Edition. Addison-Wesley; Reading, MA: 1981. The Art of Computer Programming, Volume 2: Seminumerical Algorithms. [Google Scholar]

53. Mandel N., Mandel G., Dickerson R.E. Tuna cytochrome c at 2.0 A resolution. III. Coordinate optimization and comparison of structures. J. Biol. Chem. 1977;252:4619–4636. [PubMed] [Google Scholar]

54. Tomasso M.E., Tarver M.J., Whitten S.T. Hydrodynamic radii of intrinsically disordered proteins determined from experimental polyproline II propensities. PLoS Comput. Biol. 2016;12:e1004686. [PMC free article] [PubMed] [Google Scholar]

55. English L.R., Tilton E.C., Whitten S.T. Intrinsic α helix propensities compact hydrodynamic radii in intrinsically disordered proteins. Proteins. 2017;85:296–311. [PMC free article] [PubMed] [Google Scholar]

56. Baldwin R.L. Temperature dependence of the hydrophobic interaction in protein folding. Proc. Natl. Acad. Sci. USA. 1986;83:8069–8072. [PMC free article] [PubMed] [Google Scholar]

57. Murphy K.P., Freire E. Thermodynamics of structural stability and cooperative folding behavior in proteins. Adv. Protein Chem. 1992;43:313–361. [PubMed] [Google Scholar]

58. Murphy K.P., Bhakuni V., Freire E. Molecular basis of co-operativity in protein folding. III. Structural identification of cooperative folding units and folding intermediates. J. Mol. Biol. 1992;227:293–306. [PubMed] [Google Scholar]

59. Lee K.H., Xie D., Amzel L.M. Estimation of changes in side chain configurational entropy in binding and folding: general methods and application to helix formation. Proteins. 1994;20:68–84. [PubMed] [Google Scholar]

60. Xie D., Freire E. Structure based prediction of protein folding intermediates. J. Mol. Biol. 1994;242:62–80. [PubMed] [Google Scholar]

61. Gómez J., Hilser V.J., Freire E. The heat capacity of proteins. Proteins. 1995;22:404–412. [PubMed] [Google Scholar]

62. D'Aquino J.A., Gómez J., Freire E. The magnitude of the backbone conformational entropy change in protein folding. Proteins. 1996;25:143–156. [PubMed] [Google Scholar]

63. Habermann S.M., Murphy K.P. Energetics of hydrogen bonding in proteins: a model compound study. Protein Sci. 1996;5:1229–1239. [PMC free article] [PubMed] [Google Scholar]

64. Luque I., Mayorga O.L., Freire E. Structure-based thermodynamic scale of alpha-helix propensities in amino acids. Biochemistry. 1996;35:13681–13688. [PubMed] [Google Scholar]

65. Elam W.A., Schrank T.P., Hilser V.J. Evolutionary conservation of the polyproline II conformation surrounding intrinsically disordered phosphorylation sites. Protein Sci. 2013;22:405–417. [PMC free article] [PubMed] [Google Scholar]

66. Gasteiger E., Hoogland C., Bairoch A. Protein identification and analysis tools on the ExPASy server. In: Walker J.M., editor. The Proteomics Protocols Handbook. Humana Press; 2005. pp. 571–607. [Google Scholar]

67. Meng F., Uversky V.N., Kurgan L. Comprehensive review of methods for prediction of intrinsic disorder and its molecular functions. Cell. Mol. Life Sci. 2017;74:3069–3090. [PubMed] [Google Scholar]

68. Kozlowski L.P., Bujnicki J.M. MetaDisorder: a meta-server for the prediction of intrinsic disorder in proteins. BMC Bioinformatics. 2012;13:111. [PMC free article] [PubMed] [Google Scholar]

69. Privalov P.L. Stability of proteins: small globular proteins. Adv. Protein Chem. 1979;33:167–241. [PubMed] [Google Scholar]

70. Micsonai A., Wien F., Kardos J. Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. Proc. Natl. Acad. Sci. USA. 2015;112:E3095–E3103. [PMC free article] [PubMed] [Google Scholar]

71. Sreerama N., Venyaminov S.Y., Woody R.W. Estimation of protein secondary structure from circular dichroism spectra: inclusion of denatured proteins with native proteins in the analysis. Anal. Biochem. 2000;287:243–251. [PubMed] [Google Scholar]

72. Shi Z., Chen K., Kallenbach N.R. Conformation of the backbone in unfolded proteins. Chem. Rev. 2006;106:1877–1897. [PubMed] [Google Scholar]

73. Chen K., Liu Z., Kallenbach N.R. The polyproline II conformation in short alanine peptides is noncooperative. Proc. Natl. Acad. Sci. USA. 2004;101:15352–15357. [PMC free article] [PubMed] [Google Scholar]

74. Kjaergaard M., Nørholm A.B., Kragelund B.B. Temperature-dependent structural changes in intrinsically disordered proteins: formation of alpha-helices or loss of polyproline II? Protein Sci. 2010;19:1555–1564. [PMC free article] [PubMed] [Google Scholar]

75. Saxena V.P., Wetlaufer D.B. A new basis for interpreting the circular dichroic spectra of proteins. Proc. Natl. Acad. Sci. USA. 1971;68:969–972. [PMC free article] [PubMed] [Google Scholar]

76. Toniolo C., Polese A., Kamphuis J. Circular dischroism spectrum of a peptide 3.10-helix. J. Am. Chem. Soc. 1996;118:2744–2745. [Google Scholar]

77. Kabsch W., Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers. 1983;22:2577–2637. [PubMed] [Google Scholar]

78. Venkatachalam C.M. Stereochemical criteria for polypeptides and proteins. V. Conformation of a system of three linked peptide units. Biopolymers. 1968;6:1425–1436. [PubMed] [Google Scholar]

79. Brahms S., Brahms J., Brack A. Identification of β,β-turns and unordered conformations in polypeptide chains by vacuum ultraviolet circular dichroism. Proc. Natl. Acad. Sci. USA. 1977;74:3208–3212. [PMC free article] [PubMed] [Google Scholar]

80. Borcherds W., Theillet F.X., Daughdrill G.W. Disorder and residual helicity alter p53-Mdm2 binding affinity and signaling in cells. Nat. Chem. Biol. 2014;10:1000–1002. [PubMed] [Google Scholar]

81. Cowan P.M., McGavin S. Structure of poly-L-proline. Nature. 1955;176:501–503. [Google Scholar]

82. Perez R.B., Tischer A., Whitten S.T. Alanine and proline content modulate global sensitivity to discrete perturbations in disordered proteins. Proteins. 2014;82:3373–3384. [PMC free article] [PubMed] [Google Scholar]

83. Ferreon J.C., Hilser V.J. The effect of the polyproline II (PPII) conformation on the denatured state entropy. Protein Sci. 2003;12:447–457. [PMC free article] [PubMed] [Google Scholar]

84. Lim W.A., Richards F.M., Fox R.O. Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. Nature. 1994;372:375–379. [PubMed] [Google Scholar]

85. Pandey A.K., Thomas K.M., Zondlo N.J. Tunable control of polyproline helix (PPII) structure via aromatic electronic effects: an electronic switch of polyproline helix. Biochemistry. 2014;53:5307–5314. [PMC free article] [PubMed] [Google Scholar]

86. Kauzmann W. Some factors in the interpretation of protein denaturation. Adv. Protein Chem. 1959;14:1–63. [PubMed] [Google Scholar]

87. Bolen D.W., Rose G.D. Structure and energetics of the hydrogen-bonded backbone in protein folding. Annu. Rev. Biochem. 2008;77:339–362. [PubMed] [Google Scholar]

88. Fleming P.J., Rose G.D. Do all backbone polar groups in proteins form hydrogen bonds? Protein Sci. 2005;14:1911–1917. [PMC free article] [PubMed] [Google Scholar]

89. Flory P.J. The configuration of real polymer chains. J. Chem. Phys. 1949;17:303–310. [Google Scholar]

90. Riback J.A., Bowman M.A., Sosnick T.R. Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water. Science. 2017;358:238–241. [PMC free article] [PubMed] [Google Scholar]

91. Flanagan J.M., Kataoka M., Engelman D.M. Mutations can cause large changes in the conformation of a denatured protein. Biochemistry. 1993;32:10359–10370. [PubMed] [Google Scholar]

92. Kyte J., Doolittle R.F. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 1982;157:105–132. [PubMed] [Google Scholar]

93. Scholtz J.M., Marqusee S., Bolen D.W. Calorimetric determination of the enthalpy change for the alpha-helix to coil transition of an alanine peptide in water. Proc. Natl. Acad. Sci. USA. 1991;88:2854–2858. [PMC free article] [PubMed] [Google Scholar]

94. Brown J.E., Klee W.A. Helix-coil transition of the isolated amino terminus of ribonuclease. Biochemistry. 1971;10:470–476. [PubMed] [Google Scholar]

95. Marqusee S., Robbins V.H., Baldwin R.L. Unusually stable helix formation in short alanine-based peptides. Proc. Natl. Acad. Sci. USA. 1989;86:5286–5290. [PMC free article] [PubMed] [Google Scholar]

96. Babu C.R., Hilser V.J., Wand A.J. Direct access to the cooperative substructure of proteins and the protein ensemble via cold denaturation. Nat. Struct. Mol. Biol. 2004;11:352–357. [PubMed] [Google Scholar]

97. Zahran Z.N., Chooback L., Richter-Addo G.B. Crystal structures of manganese- and cobalt-substituted myoglobin in complex with NO and nitrite reveal unusual ligand conformations. J. Inorg. Biochem. 2008;102:216–233. [PMC free article] [PubMed] [Google Scholar]

98. Saito R., Sato T., Tanaka N. Structure of bovine carbonic anhydrase II at 1.95 A resolution. Acta Crystallogr. D Biol. Crystallogr. 2004;60:792–795. [PubMed] [Google Scholar]

99. Stein P.E., Leslie A.G., Carrell R.W. Crystal structure of ovalbumin as a model for the reactive centre of serpins. Nature. 1990;347:99–102. [PubMed] [Google Scholar]