BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

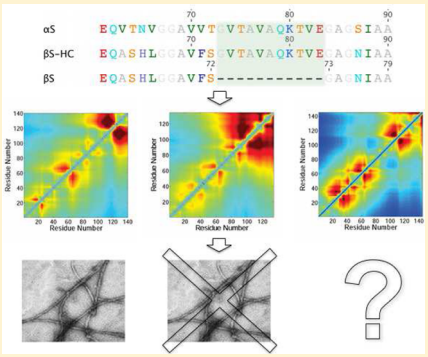
A Relationship between the Transient Structure in the Monomeric State and the Aggregation Propensities of α-Synuclein and β-Synuclein

Jane R. Allison,*[,†] Robert C. Rivers,[‡] John C. Christodoulou,[§] Michele Vendruscolo,* and Christopher M. Dobson*

Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, U.K.

Supporting Information

ABSTRACT: α-Synuclein is an intrinsically disordered protein whose aggregation is implicated in Parkinson's disease. A second member of the synuclein family, β-synuclein, shares significant sequence similarity with α-synuclein but is much more resistant to aggregation. β-Synuclein is missing an 11-residue stretch in the central non-β-amyloid component region that forms the core of α-synuclein amyloid fibrils, yet insertion of these residues into β-synuclein to produce the βS_{HC} construct does not markedly increase the aggregation propensity. To investigate the structural basis of these different behaviors, quantitative nuclear magnetic resonance data, in the form of paramagnetic relaxation enhancement-derived interatomic distances, are combined with molecular dynamics simulations to generate ensembles of structures representative of the solution states of α-synuclein, β-synuclein, and βS_{HC}. Comparison of these ensembles reveals that the differing aggregation propensities of α-synuclein and β-synuclein are associated with differences in the degree of residual structure in the C-terminus coupled to the shorter separation between the N- and C-termini in β-synuclein and βS_{HC}, making protective intramolecular contacts more likely.



Intrinsically disordered proteins (IDPs) are involved in myriad biological processes, including cellular signaling, molecular recognition, and transcriptional regulation.[1−5] Additionally, members of this class of proteins have been implicated in a number of debilitating protein misfolding disorders.[6] For instance, Aβ peptides and α-synuclein (αS) are the primary constituents of the amyloid deposits found in Alzheimer's disease and Parkinson's disease, respectively.[7−9] A description of the native state ensembles of IDPs in terms of the constituent structures and their relative populations is vital to understanding both the function and the aggregation process of these proteins. The absence of persistent secondary and tertiary structure elements in IDPs does not preclude the presence of well-defined conformational preferences. Indeed, residual structure, often in the form of transient long-range contacts, has been detected in many IDPs,[10−22] and some exhibit pockets of structure that have a propensity to bind small molecules.[23,24]

The heterogeneity and broadness of the ensembles of structures characteristic of disordered states of proteins make the determination of the conformational properties of IDPs particularly challenging. For example, nuclear Overhauser effect (NOE)-based nuclear magnetic resonance (NMR) measurements are sensitive only up to separations of ∼0.5 nm, the result being that transient tertiary interactions in disordered states are unlikely to be detected using this approach. Despite this limitation, it has been possible to extract some structural information about disordered states from certain types of X-ray techniques and NMR spectroscopy measurements. For instance, small-angle X-ray scattering (SAXS)[25,26] and diffusion NMR spectroscopy[27] have been used to determine the molecular dimensions of IDPs. NMR observables such as residual dipolar couplings (RDCs) have proven to be a useful source of detailed structural information about disordered states,[10,14,16,28−33] and methods are also emerging for utilizing chemical shifts.[18−22,34−37]

Paramagnetic relaxation enhancement (PRE) experiments overcome the limitations of NOE measurements in probing the conformational properties of IDPs by utilizing the longer-range dipolar interactions between unpaired electrons in paramagnetic probes and atomic nuclei, which can be detected experimentally at distances up to ∼2.0 nm. The paramagnetic probe is often a free radical, typically a nitroxide spin-label covalently attached to a cysteine residue introduced into the protein of interest by site-specific mutagenesis. $^{1}H-^{15}N$ HSQC spectra are then recorded with the spin-label in its paramagnetic (oxidized) and diamagnetic (reduced) states. The enhancement of the transverse relaxation of each proton due to the free electron of the oxidized spin-label can be quantified by comparing the intensities of each proton resonance measured for each spin-label state.[38] From the resultant intensity ratios (I_{ox}/I_{red}), the r^{-6} average of the distance between the free

Received: July 29, 2014
Revised: October 25, 2014
Published: October 28, 2014

   dx.doi.org/10.1021/bi5009326 | *Biochemistry* 2014, 53, 7170−7183

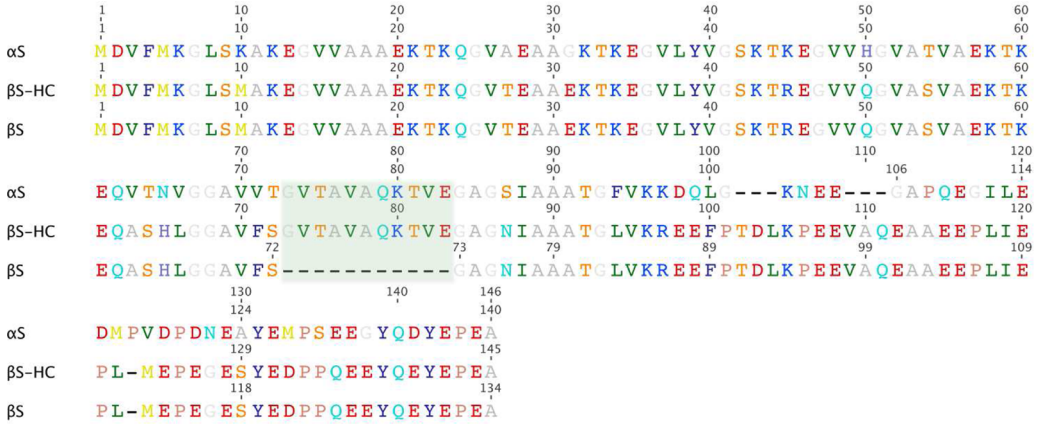


Figure 1. Alignment of the amino acid sequences of αS, βS, and βS_{HC}. Amino acids are colored according to the chemical nature of their side chains. The region shaded in Cambridge blue indicates the 11 residues from αS that were inserted into βS to form βS_{HC}.

electron and each proton (typically the backbone amide hydrogen) in the protein can be deduced.[38−40] The fact that this distance is a time and ensemble average over the duration of the experiment and the ensemble of molecules present is an important consideration in the analysis of PRE data. When PRE−NMR experiments are conducted with the spin-label attached at a number of different positions in the protein, sufficient distances for characterizing key features of the conformational ensemble of the protein can be obtained.[17]

IDPs have also been characterized using molecular dynamics (MD) simulations,[41−46] although such techniques are hampered by the need to explore very large regions of conformational space. Because this is computationally expensive, implicit solvent models are often used.[17,47−53] Regardless of whether implicit or explicit solvent models are used, however, compact structures tend to be favored relative to more extended conformational states, most likely because most force fields have been parametrized to reproduce structural data for natively folded proteins. This is, however, an area of intense research in which rapid progress can be expected. Conducting the simulations at high temperatures allows more expanded structures to be sampled, but with the concomitant risk of compromising the physical relevance of the structures that are explored. Adding restraints derived from experimental data can, however, help to overcome this problem, while simultaneously biasing sampling toward relevant structures and restricting the conformational space that is explored, thereby reducing the simulation time and computational expense required for converged simulations.[17,54,55] Such restrained MD simulations can also be seen to aid the interpretation of experimental data in terms of structures and their populations, particularly for IDPs where the experimental data are in general averages over many disparate structures. Because of this factor, it is important to apply the restraints as averages, which can be achieved by averaging over time[56,57] or space,[58−61] i.e., over ensembles of structures. Care must be taken, however, to ensure that there are sufficient data to warrant the additional degrees of freedom that result from averaging over multiple replicas or time points. Additionally, nonlinearly averaged restraints can result in over- or underestimation of the population of structures with short distances.[62,63] In this context, it has been recently recognized that the use of replica averaging represents an implementation of the maximum entropy principle to incorporate the experimental information into the molecular dynamics simulations.[62−65]

In this work, we consider two related and similarly sized (127 and 140 residues) IDPs, α-synuclein (αS) and β-synuclein (βS).[66−68] Despite their significant sequence similarity (Figure 1), these two proteins differ considerably in their behavior and medical significance. In particular, while αS aggregates to form the Lewy bodies characteristic of Parkinson's disease, βS does not appear to aggregate *in vivo* and has even been shown to inhibit fibril formation by αS.[69,70] To explore the reasons for such differences between αS and βS, we introduced a construct of βS, βS_{HC},[71] which incorporates residues 71−82 of the non-β-amyloid component (NAC) region of αS (Figure 1) to determine whether this highly hydrophobic 11-residue region, which is absent in the sequence of βS, is sufficient to induce αS-like aggregation behavior in βS. Despite the fact that the NAC region is thought to be the primary determinant of αS aggregation[72] and to be necessary for fibril formation, particularly residues 63−74,[73,74] the aggregation properties of βS_{HC} are closer to those of βS,[71] which is much less aggregation prone than αS.

To fully understand the reasons for these differing aggregation behaviors, it is necessary to characterize the ensemble of structures sampled by each protein under the same conditions under which aggregation occurs. αS in solution has been the subject of very many experimental,[75−86] computational,[83,87] and hybrid[10−12,17,19,20,88−92] investigations. While each study has highlighted different structural features, there is a general agreement that in solution, the C-terminal region of αS appears to provide some protection to the remainder of the protein, including the aggregation prone central NAC region. To date, the structural propensities of βS have been characterized only experimentally.[13,15,93] To build upon this, PRE−NMR experiments were conducted on βS and βS_{HC}, and distances derived from the experimental data were used as replica-averaged restraints in MD simulations to generate ensembles of structures representative of the native states of these proteins. Comparison of these ensembles and the ensemble of structures previously generated for αS reveals specific differences in the structural preferences of the three proteins and allows the effects of the hydrophobic core on the structural properties of these forms of synuclein, including their different aggregation propensities, to be examined at a molecular level.

dx.doi.org/10.1021/bi5009326 | *Biochemistry* 2014, 53, 7170−7183

METHODS

Protein Preparation. ^{15}N-labeled βS and βS_{HC} were expressed and purified as described previously.[71] A QuickChange (Stratagene) site-directed mutagenesis kit was used to engineer cysteine mutations at positions A30, S42, S64, F89, A102, S118, and A134 in βS and A30, S42, S64, A113, and A145 in βS_{HC}. Mutation sites were selected to minimize structural perturbations and to correspond as closely as possible to the αS mutation sites (Q24, S42, Q62, S87, N103, and N122). The nitroxide spin-label MTSL (1-oxyl-2,2,5,5-tetramethyl-3-pyrroline-3-methylmethanethiosulfonate) (Toronto Research Chemicals Inc.) was attached to the introduced cysteine residue in each variant in a thiol-specific reaction. The cysteine variants were first reduced with 5 mM DTT, which was subsequently removed using a 20 mL HiTrap desalting column (Amersham-Pharmacia) connected to an Akta fast protein liquid chromatography instrument (Amersham-Pharmacia). Immediately following DTT removal, the protein solution was incubated overnight with a 10-fold molar excess of MTSL. After incubation, unreacted MTSL was removed with a HiTrap desalting column. Uniform labeling was confirmed using mass spectrometry. Analysis of the spin-labeled variants using circular dichroism showed no evidence of any conformational changes. Moreover, the amide proton and nitrogen chemical shifts in the HSQC spectra were not significantly altered, even for residues in the vicinity of the spin-label (Figures S1 and S2 of the Supporting Information).

NMR Spectroscopy. Two-dimensional gradient-enhanced ^{1}H−^{15}N HSQC of βS and βS_{HC} was conducted following protocols described previously[12] at the EPSRC-supported biomolecular NMR facility (Department of Chemistry, University of Cambridge) on a Bruker Avance 700 MHz spectrometer operating at 10 °C. Experimental samples contained 100 μM uniformly ^{15}N-labeled protein with MTSL attached in 10 mM sodium phosphate (pH 7.4), 100 mM NaCl, and 10% D_2O. Control samples contained 100 μM ^{15}N-labeled protein and 100 μM spin-labeled protein. The uniformity of the I_{ox}/I_{red} calculated from the control spectra showed that there were no complications arising from the reduction method and that aggregation did not occur. Backbone NMR assignments for αS and βS were obtained by standard triple-resonance methods as previously described.[71,75] Assignment of βS_{HC} was obtained with truncated triple-resonance CBCA(CO)NH and HNCO experiments, and an overlay of the ^{1}H−^{15}N HSQC spectra confirmed that the chemical shifts of the added and original residues conformed to the chemical shifts of these residues in αS and βS, respectively (Figure S3 of the Supporting Information). For each spin-labeled mutant, an HSQC spectrum was first acquired with the label in its oxidized state. A 5-fold molar excess of sodium ascorbate was then added from a concentrated stock solution to reduce the spin-label without altering significantly the sample volume or pH. After incubation for at least 20 h, a second HSQC spectrum was acquired with all parameters remaining unchanged. HSQC spectra were collected using 16 scans per increment, with 1024 complete points for the direct dimension and 128 complex points for the indirect dimension. NMR data were processed with NMRPipe[94] and analyzed with Sparky.[95] Harsh resolution enhancing functions were not used to avoid nonuniform effects on cross-peak intensities, and cross-peaks exhibiting severe overlap were omitted from further analysis.

Distance Calculations. The electron−proton distances were calculated from the intensity ratios (I_{ox}/I_{red}) as described previously,[12] including the modifications introduced by Allison et al.[17] as detailed below. Residue-specific values of R_2 were used where available; otherwise, the average over all residues was used.

Examination of the effect of introducing uncertainty of up to 15% in I_{ox}/I_{red} on the calculated distance showed that variation of up to 10% in I_{ox}/I_{red} results in propagated uncertainties of less than −0.19 or 0.38 nm in the calculated distance, which is a tolerable level. Distances were therefore used as restraints only if the difference between each replicate value of I_{ox}/I_{red} and the average value was less than 10% of the average I_{ox}/I_{red} value. The fraction of the experimental data that was discarded in this way for each protein is listed in Table 3 along with the total number of distance restraints for each protein. For each protein, 20% of the distances were removed from the "working" data set to be used for independent cross-validation.

During ensemble-averaged simulations using PRE-derived distance restraints, the calculated distance, $d_{ij}^{calc}(t)$, is allowed to vary freely within $d_{ij}^{exp}(t) - L$ and $d_{ij}^{exp}(t) + U$, where L and U are the distances to the lower and upper bounds, respectively, of the flat bottom of the harmonic square well. Detailed investigations using synthetic data have shown that the optimal choices for L and U to best reproduce the distribution of distances as well as the r^{-6} average are 0.1 and 0.8 nm, respectively.[17,96]

As a general rule, I_{ox}/I_{red} values of <0.15 are unreliable,[38] as any experimental uncertainty is large relative to the size of the measured I_{ox}/I_{red}. Distances calculated from experimental I_{ox}/I_{red} values of <0.15 were therefore assigned only an upper bound corresponding to $d_{ij}^{0.15} + U$, where $d_{ij}^{0.15}$ is the distance calculated from an I_{ox}/I_{red} value of 0.15. The nature of the equations used to calculate the distances means that for a high I_{ox}/I_{red}, a small change in I_{ox}/I_{red} results in a large change in the calculated distance. Thus, I_{ox}/I_{red} values of >0.85 were used as "negative" restraints by assigning only a lower bound corresponding to $d_{ij}^{0.85} - L$, where $d_{ij}^{0.85}$ is the distance calculated from an I_{ox}/I_{red} value of 0.85.

Molecular Dynamics Simulations. All simulations were conducted using an in-house version of the CHARMM biomolecular simulation package[97] that has been modified to allow restraints to be applied across multiple replicas. The Newtonian equations of motion were integrated using the Velocity Verlet algorithm,[98] and the Nose-Hoover thermostat[99,100] was employed so that a canonical ensemble was sampled. The CHARMM19 polar hydrogen representation[101] was used, and bond lengths were constrained with the SHAKE algorithm,[102] allowing for an integration time step of 2 fs. A set of unrelated, expanded starting structures for each protein were chosen from high-temperature (500 K) simulations with the EEF1[103] implicit solvent model. The final ensemble for each simulation was obtained by pooling together all of the structures obtained during the production phase; if multiple replicas were used, these were pooled, as well.

Random Coil Model. A reference random coil model for each protein was produced by truncating the nonbonded interactions so that only the repulsive part of the Lennard-Jones potential remained. Molecular dynamics simulations were run in vacuum with no electrostatic interactions. The temperature, T, was typically 500−600 K to enhance the rate of sampling, but the nature of the resulting ensemble was similar at lower

values of T. The coordinates were saved every 20 ps for 200 ns, giving 10000 structures in total.

The intensity ratios expected for a purely random coil were computed by first calculating the r^{-6}-averaged distances between the Cα atoms of the spin-labeled residues and the amide hydrogens of all other residues. These distances were converted into intensity ratios by following the inverse of the procedure used to calculate distances from intensity ratios.

Molecular Dynamics Simulations with Replica-Averaged Distance Restraints. Restrained simulations were conducted using molecular dynamics with replica-averaged distance restraints derived from PRE−NMR measurements. In this approach,[47,50,54,59,104−113] the restraints are applied to multiple independent replicas simulated in parallel. A restraint coordinate, ρ, is defined as the difference between the current average of each observable across all replicas, f_i^{calc}, and the experimentally derived restraint, f_i^{exp}, averaged over all N^{res} restraints:

$$\rho(t) = \frac{1}{N^{res}} \sum_{i=1}^{N^{res}} (f_i^{exp} - f_i^{calc})^2 \quad (1)$$

where f_i^{exp} refers to the r^{-6}-averaged distance d_{ij}^{exp} derived from the experimental PRE−NMR data as detailed above and f_i^{calc} was calculated from the simulated structures according to

$$f_i^{calc} = d_{ij}^{calc}(t) = \left[\frac{1}{N^{rep}} \sum_{i=1}^{N^{rep}} r_{ij,k}^{-6}(t) \right]^{-1/6} \quad (2)$$

where $r_{ij,k}(t)$ is the distance between residues i and j calculated from replica k of the restrained ensemble at time t and N^{rep} is the number of replicas. $r_{ij,k}$ was defined as being between the Cα atom of spin-labeled residue i and the amide hydrogen of residue j. A flat bottom restraint potential was used, meaning that the contribution of a given distance d_{ij} to the restraint coordinate is zero if $d_{ij}^{exp}(t) - L < d_{ij}^{calc}(t) < d_{ij}^{exp}(t) + U$.

An energy penalty of the form

$$\frac{\alpha N^{rep}}{2}[\rho(t) - \rho_0(t)]^2 \quad (3)$$

is added to the potential energy if $\rho(t) > \rho_0(t)$, where

$$\rho_0(t) = \min_0 \leq \tau \leq \rho(t) \quad (4)$$

and α is a force constant associated with the restraints. In this way, as the simulation proceeds, the ensemble of structures is progressively biased toward structures that, on average, satisfy the restraints.

The replica-averaged MD simulations were conducted using the SASA[114,115] implicit solvation model with default cutoff distances for nonbonded and electrostatic interactions and rectangular periodic boundary conditions. Following the protocol developed using synthetic data for αS,[17] 24 replicas were simulated in parallel. The molecules were first heated to 700 K in 50 K increments (10 ps per temperature), and then α was increased from its starting value of 500 kcal mol^{-1} Å^{-2} to its final value of 364500 kcal mol^{-1} Å^{-2} by a factor of 3 every 10 ps. After a brief equilibration (200 ps), the coordinates were collected every 5 ps for 400 ps per replica, giving 1920 structures in total. The temperature, T, was then lowered by 25 K and the system re-equilibrated before 1920 structures were collected at the new T. Q values quantifying the agreement with the experimental data (see below) were calculated at each T so that the agreement with experiment could be monitored constantly. The cooling−equilibration−collection cycle was continued until the various Q values (Table 3) were simultaneously minimized. An additional 5760 structures (1.2 ns per replica) were collected at this optimal T for further analysis.

Analysis. *Calculation of R_g and R_h.* The geometric radius of gyration, R_g, was calculated from the heavy atoms of each structure using CHARMM analysis facilities. For comparison with experimental data, the hydrodynamic radius, R_h, of each ensemble was computed. For each protein, the R_h of 200 randomly selected structures of varying degrees of compactness was computed using HYDROPRO[116] with default settings and six sizes of minibeads ranging from 0.18 to 0.28 nm. The molecular weight and partial specific volume were evaluated from the amino acid sequence. Relationships between R_g^{-1} and R_h^{-1} were then determined by linear regression (uncertainty represents standard error) (Table 1).

Table 1

protein	relationship	correlation coefficient
αS	$R_h^{-1} = 0.0148(\pm0.0003) + 0.4882(\pm0.0038)R_g^{-1}$	0.994
βS	$R_h^{-1} = 0.0163(\pm0.0002) + 0.4537(\pm0.0042)R_g^{-1}$	0.991
βS_{HC}	$R_h^{-1} = 0.0151(\pm0.0002) + 0.4943(\pm0.0044)R_g^{-1}$	0.990

These equations were used to convert the calculated R_g of each structure into an R_h. The overall $\langle R_h^{-1}\rangle^{-1}$ was then computed according to

$$\langle R_h^{-1}\rangle^{-1} = \left(\frac{1}{N^{struct}} \sum_{k=1}^{N^{struct}} R_{h,k}^{-1} \right)^{-1} \quad (5)$$

where N^{struct} is the number of structures in the ensemble, to reflect the averaging inherent in the experimental measurement.

Compaction Factors. Compaction factors, C_f, quantifying the degree of compaction relative to that of a fully unfolded (random coil) and natively folded state were calculated according to[27]

$$C_f = \frac{R_h^U - R_h^{exp}}{R_h^U - R_h^F} \quad (6)$$

where R_h^{exp} is the experimental R_h and R_h^F and R_h^U are the R_h values expected if the protein is natively folded (F) and fully unfolded (U), respectively:

$$R_h^F = 4.75N^{0.29}$$
$$R_h^U = 2.21N^{0.57} \quad (7)$$

Q Values. The agreement between the synthetic or experimental observables and those calculated from a calculated ensemble was quantified with a "quality factor":[117]

$$Q = \frac{\sqrt{\sum_{k=1}^{N^{obs}} (f_k^{calc} - f_k^{exp})^2}}{\sqrt{\sum_{k=1}^{N^{obs}} (f_k^{exp})^2}} \quad (8)$$

where N^{obs} is the number of observables of that type (e.g., working or free PRE distances) and the f_k^{calc} values are the averages over the pooled ensemble.

dx.doi.org/10.1021/bi5009326 | Biochemistry 2014, 53, 7170−7183

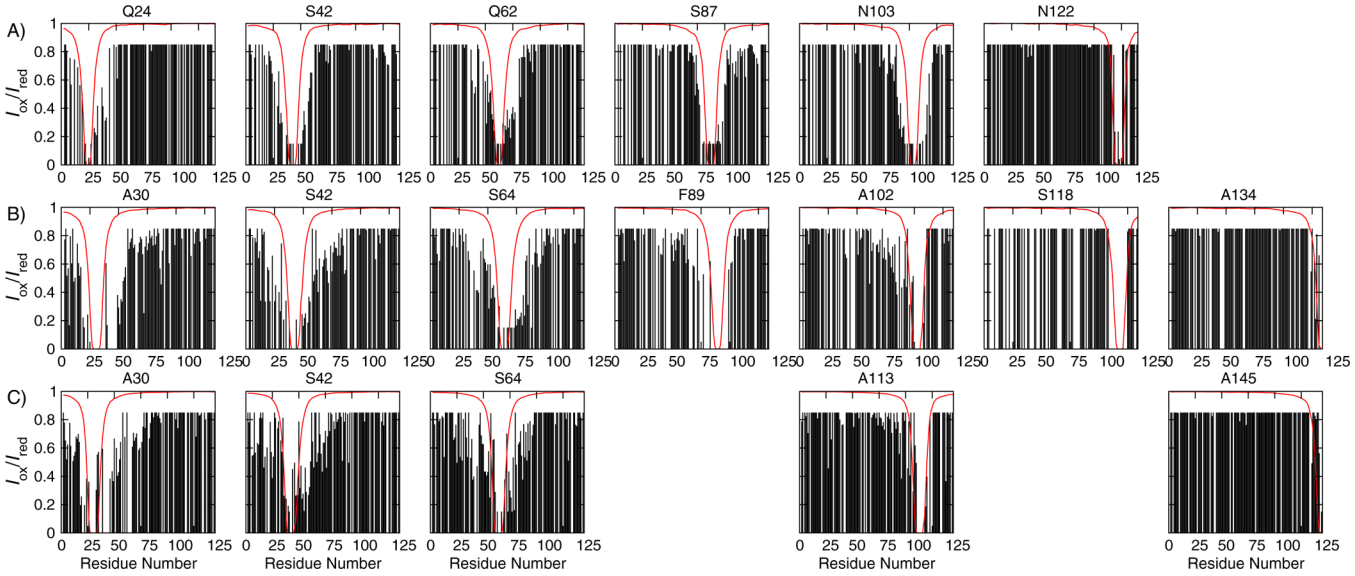


Figure 2. Intensity ratios I_{ox}/I_{red} for each spin-label position for (A) αS,[12,17] (B) βS, and (C) βS_{HC}. The experimental data are shown as black bars, and the I_{ox}/I_{red} values calculated from the random coil ensemble are plotted as thick red lines. PRE−NMR experiments were conducted on 100 μM uniformly ^{15}N-labeled protein with MTSL attached in 10 mM sodium phosphate (pH 7.4), 100 mM NaCl, and 10% D_2O at 10 °C. The experimental I_{ox}/I_{red} values are those processed for use in the simulations (see Methods); thus, any I_{ox}/I_{red} of <0.15 or >0.85 has been set to 0.15 or 0.85, respectively. If no bar is present, then either I_{ox}/I_{red} was not measured for this residue or it was discarded because of an uncertainty of >10%.

Distance Comparison Maps. Distance comparison (DC) maps were created by plotting the root-mean-square (rms) distance between two residues, i and j, normalized by the rms distance predicted for a purely random coil:

$$\frac{\langle d_{ij}^{\mathrm{calc}^2}\rangle^{1/2}}{\langle d_{ij}^{\mathrm{rc}^2}\rangle^{1/2}} \tag{9}$$

The rms inter-residue distances for the calculated ensemble were calculated as

$$\langle d_{ij}^{\mathrm{calc}^2}\rangle^{1/2} = \frac{1}{N^{\mathrm{struct}}}\sum_{k=1}^{N^{\mathrm{struct}}} (d_{ij,k}{}^2)^{1/2} \tag{10}$$

where N_{struct} is the number of structures in the calculated ensemble. The rms inter-residue distances for a random coil were calculated according to

$$\langle (d_{ij}^{\mathrm{rc}})^2\rangle^{1/2} = 5.31(N^{\mathrm{sep}})^{0.6} \tag{11}$$

which predicts the rms distance between two residues with sequence separation N^{sep} for a random flight chain with excluded volume and dihedral angles taken from a Protein Data Bank coil library.[118] Similar results were obtained if $\langle (d_{ij}^{\mathrm{rc}})^2\rangle^{1/2}$ was calculated from the random coil model of the protein in question. The normalization by $\langle (d_{ij}^{\mathrm{rc}})^2\rangle^{1/2}$ is important because it removes the dependence of the inter-residue distance on the sequence separation, allowing pairs of residues with different sequence separations and also proteins of different lengths to be compared.

^3J Couplings. The $^3J_{\mathrm{HNH}\alpha}$ couplings were calculated for each structure using the GROMACS[119] program g_chi with default settings, values for the Karplus relation parameters of A = 6.4, B = −1.4, and C = 1.9,[120] and an offset of −60° and then averaged over all structures in an ensemble.

Residual Dipolar Couplings. Residual dipolar couplings were calculated for each structure using steric PALES[121] with default settings and then averaged over all structures in an ensemble.

Solvent Accessible Surface Area. The solvent accessible surface area of each structure was calculated using the algorithm of Lee and Richards,[122] as implemented in CHARMM, using default settings, including a probe radius of 0.16 nm.

Aggregation Propensity. Aggregation propensity profiles ($Z_{\mathrm{agg}}^{\mathrm{prof}}$) of αS, βS, and βS_{HC} were computed using an updated version of the Zyggregator algorithm,[123] which predicts the aggregation propensity of peptides and proteins in aqueous solution from the physicochemical properties of their constituent amino acids and compares this to the aggregation propensity of a set of randomly generated amino acid sequences of the same length.[124] $Z_{\mathrm{agg}}^{\mathrm{prof}}$ indicates the regions that are most prone to aggregation.

RESULTS AND DISCUSSION

Detection of Nonrandom Structure. PRE−NMR experiments combined with the calculation of ensembles of structures consistent with the NMR data have already been conducted for αS.[12,17] The ensemble of structures obtained was validated by comparison with independent experimental data. Here we describe similar experiments and calculations for βS, and for the artificial construct βS_{HC}. We note that in all cases, the nonacetylated form of the protein was studied, as this is the form of the heterologously expressed protein studied experimentally, and the simulations aimed to match the experiments as closely as possible.

In discussing these three proteins, we define the N-terminal and central regions as those regions that can form α-helical lipid-bound structure:[125] residues 1−98 in αS, residues 1−65 in βS, and, assuming the additional 11 residues from αS extend the helical region of βS, residues 1−76 in βS_{HC}, with the remainder of each protein being designated as the C-terminal region.

The backbone assignments of αS^{75} and βS^{71} were previously described. For the βS_{HC} construct, the ^{1}H−^{15}N HSQC

dx.doi.org/10.1021/bi5009326 | Biochemistry 2014, 53, 7170−7183

spectrum overlaps with that of βS, and the additional 11 residues from αS exhibit chemical shifts in βS_{HC} that correspond closely to their resonances in αS (Figure S3 of the Supporting Information). This allowed a backbone assignment strategy in which the identification of individual amino acid resonances was confirmed by a combination of CBCA(CO)NH and HNCO triple-resonance experiments. The chemical shifts for βS_{HC} have been deposited in the BioMagResBank.

Seven distinct single-residue cysteine mutations were introduced into βS, and five into βS_{HC}, to which the MTSL spin-label was subsequently attached. The positions of the cysteine mutations were kept as consistent as possible among the three proteins (Table 3 and Figure 2) to facilitate comparisons. For each cysteine mutant of each protein, $^{1}H-^{15}N$ HSQC spectra were collected with the spin-label in both its oxidized state and its reduced state. The heights of individual NMR resonances were then used to calculate the intensity ratios shown in Figure 2.

A decrease in the intensity ratio is expected for residues proximal in sequence to the spin-label attachment site. The predicted pattern of intensity ratios stemming from this effect is illustrated by the red lines in Figure 2, which show the intensity ratios calculated from random coil representations of each protein. Additional regions with intensity ratios lower than these values correspond to internuclear distances that are significantly shorter than in a random coil ensemble. All three proteins exhibit such long-range contact formation, indicative of nonrandom structure, suggesting that they are more compact than a random coil of the same sequence. Control experiments in which the HSQC spectra were collected for a mixture of isotopically labeled protein and spin-labeled protein confirmed that the observed intensity decreases were due to intramolecular contact formation, rather than from intermolecular contact formation due to aggregation (data not shown).

Examination of the intensity ratios in more detail reveals that the majority of the contact formation is between residues of intermediate (up to 30 residues apart) sequence separation, indicative of local structural collapse. In particular, the decreases in the intensity ratios for residues around spin-label positions Q24, A30, and A30 and Q62, S64, and S64 for αS, βS, and βS_{HC}, respectively, and S42 in βS and βS_{HC} extend further from the spin-label in both directions than what is predicted by the random coil model. However, when the spin-label is attached at position N103, A102, or A113 in αS, βS, or βS_{HC}, respectively, only residues located on the N-terminal side of the spin-label show decreased intensity ratios, suggestive of an extended C-terminus in all three proteins. In αS, attaching the spin-label at position S42 results in lower intensity ratios for C-terminal residues from position 110 onward, and some evidence of the reciprocal interaction can be seen for spin-label position N122. Neither of these effects is observed for βS or βS_{HC}, suggestive of fewer, or at least different, patterns of long-range contact formation for these proteins. Overall, the intensity ratios suggest some local compaction in the N-terminal and central regions of all three proteins, and more extended structure in the C-terminal regions, particularly for βS and βS_{HC}.

Generation of Ensembles of Structures. To determine the molecular details of the structures giving rise to the PRE–NMR data, ensembles of structures compatible with the PRE-derived distances were determined using replica-averaged restrained MD (PRE-RMD) simulations.[12,17,48] To account for the averaging inherent in the experimental data, the PRE-derived distance restraints were applied to multiple (24) independent replicas simulated in parallel. At each point in time, a restraint coordinate, ρ, was obtained by comparing the r^{-6} average of each distance across all replicas to the experimental value (eq 1). An energy penalty, the magnitude of which depends on the magnitude of ρ, was applied only if the value of ρ at that time point was greater than the previous minimum (eq 3). In this way, the simulations were progressively biased toward structures that, on average, satisfy the restraints. The majority of the simulation parameters, including the number of replicas, were optimized previously so they would be suitable for reproducing disordered state ensembles.[17] In particular, an asymmetric flat bottom harmonic potential was adopted to ensure that the structures generated are not overly compact, as can be the case for r^{-6}-averaged distance restraints like those used here. The only parameter that was changed in this work is the simulation temperature, which is used to tune the average dimensions of the structures that make up the ensemble, as quantified by the harmonic average of the hydrodynamic radius, $\langle R_h^{-1}\rangle^{-1}$, so that it matches the experimentally determined value. This tuning was shown greatly to improve the accuracy of the ensemble, measured in terms of the reproduction of distributions of structural properties.[17] To provide further evidence that the ensembles of structures produced here are valid representations of the experimental ensembles, cross-validation, in which only 80% of the PRE-derived distance restraints were used in the PRE-RMD calculations ("working") and the remaining 20% provide a "free" data set whose satisfaction is not preordained by their inclusion as restraints, was conducted. For all three proteins, the agreement with the "free" set of PRE-derived distances is almost as good as that of the "working" PRE-derived distances (Table 3).

Molecular Dimensions. Two different ensembles of αS restrained with PRE-derived distances have been obtained previously, one with an average R_h close to 2.7 nm,[12] consistent with the experimental R_h of 2.66 nm measured in unbuffered D_2O at 298 K (2.66 nm),[76] and one with an average R_h of 3.2 nm,[17] to match the experimental R_h values of 3.20 and 3.19 nm measured in subsequent PFG-NMR experiments at 288 K in unbuffered D_2O[126] and in 20 mM phosphate buffer (pH 6.5) with 100 mM NaCl,[13] respectively. The latter ensemble also made use of an additional 118 distance restraints obtained after determination of the first ensemble as well as the original 478 distance restraints. The R_h values of the βS and βS_{HC} ensembles were matched to experimental values measured under conditions as close as possible to those of the PRE–NMR experiments [pH 6.5 for αS and pH 7.4 for βS and βS_{HC}, 100 mM NaCl, 288 K (Table 2)] by tuning the simulation temperature.

Comparison of the R_h values of αS, βS, and βS_{HC} must take into account their different sequence lengths. The compaction factor[27] (see Methods), C_f, allows for this difference by comparing the experimental R_h to that expected if the polypeptide were to exist in a compact folded state or to be fully unfolded (i.e., random coil-like). A C_f of 1.0 indicates compaction typical of a natively folded protein, while a C_f of zero indicates random coil-like dimensions. According to this measure, the dimensions of βS (C_f = 0.45) and βS_{HC} (C_f = 0.23) resemble those of partially unfolded proteins that retain some nonlocal interactions, such as reduced hen egg white lysozyme at pH 2.0 or BPTI at pH 4.5 (C_f = 0.35),[27] with βS slightly more compact and βS_{HC} more unfolded. In contrast, αS

dx.doi.org/10.1021/bi5009326 | Biochemistry 2014, 53, 7170–7183

Table 2. Predicted[a] and Experimental[b,c] R_h Values (nanometers) and Compaction Factors[d] (C_f) for αS, βS, and βS_{HC} in Various States

		U	F	NaCl[b]	NaCl and SDS[c]
αS	R_h	3.70	1.99	3.19	2.46
	C_f		0.61	0.30	0.725
βS	R_h	3.60	1.97	3.24	3.22
	C_f		–	0.22	0.23
βS_{HC}	R_h	3.77	2.01	–	2.97
	C_f		–	–	0.46

[a]U and F refer to the R_h values predicted according to eq 7[27] for a fully unfolded and a compact folded polypeptide, respectively. [b]Measured by PFG-NMR on 200 μM protein in D_2O and 20 mM phosphate buffer (pH 6.5) with 100 mM NaCl at 288 K.[13] Note that the R_h measured by PFG-NMR for 100 μM αS in unbuffered D_2O at 288 K is almost identical (3.20 nm).[126] [c]Measured by PFG-NMR on 70 μM protein in 10 mM phosphate buffer (pH 7.7) with 100 mM NaCl and 0.5 mM SDS at 298 K.[129] [d]Calculated according to eq 6.[27]

(C_f = 0.72) is significantly more compact, exhibiting a degree of expansion similar to that of the low-pH molten globule state of myoglobin.[27]

An alternative to the R_h value for quantifying the size of a molecule is the radius of gyration (R_g). It should be pointed out, however, that because the R_h is defined as the radius of a hard sphere with the observed diffusion rate, this parameter reflects only approximately the apparent size adopted by the solvated, tumbling molecule. R_g is defined as the mass-weighted average distance of each atom from the center of mass of the molecule, and therefore, calculating its value from the sets of coordinates obtained from an MD simulation is simple and fast. The R_h and R_g values are related to each other and can be interconverted by the approach described in Methods.

The broad distributions of the R_g values of the ensembles (Figure 3) reflect the wide variety of structures populated at least transiently by IDPs. Comparison with the R_g distributions of the random coil models of each protein, however, reveals that the range of structures accessible to each protein is restricted to conformations that are significantly more compact than those expected for a random coil, reflecting the non-zero C_f values. Also consistent with the trends observed for the compaction factors, the difference distribution of the R_g (Figure 3D) of βS_{HC} differs from those of αS and βS in a manner that indicates that the shift toward structures with smaller R_g values in the PRE-RMD simulations, compared to those expected for a random coil, is more pronounced for βS_{HC} than for the other two proteins. As with the R_h values, however, it is not appropriate to compare directly the R_g probability distributions of the three different synucleins because of their different sequence lengths.

Comparison with Experimental Data Not Used as Restraints. The most stringent test of how well a simulation reproduces the actual ensemble of structures is a quantitative comparison with independent experimental data. As noted above, the agreement between the experimentally derived and calculated "free" PRE distances is almost as good as for the "working" PRE distances (Table 3), allowing a high level of confidence that the ensembles of structures provide a good representation of the long-range structural properties of αS, βS, and βS_{HC}.

Table 3. Summary of the Experimental Restraints and How Well They Were Satisfied during the PRE-RMD Simulations[a]

	data				Q values		
protein	N_{PRE}	N_{wPRE}	N_{fPRE}	% discarded	Q_{R_h}	Q_{wPRE}	Q_{fPRE}
αS	595	476	119	17	0.006	0.19	0.20
βS	635	508	127	17	0.005	0.20	0.19
βS_{HC}	578	462	116	3	0.020	0.20	0.28

[a]N_{PRE} is the total number of distances derived from the PRE–NMR experiment, and N_{wPRE} and N_{fPRE} are the numbers of distances in the working and free data sets, comprising 80 and 20% of the total data, respectively. The percentage of the experimental data that was discarded due to uncertainties of >10% is also shown. The Q values (eq 8) quantify how well the experimental $\langle R_h^{-1}\rangle^{-1}$ (Q_{R_h}) and the working (Q_{wPRE}) and free (Q_{fPRE}) distances were satisfied by the ensemble of structures obtained using PRE-RMD.

While the primary aim of this work was to reproduce the long-range structure of the three proteins, NMR data reporting on more local structural properties, namely $^3J_{HNH\alpha}$ couplings and amide N–H RDCs, were calculated for αS and βS for comparison with experimental values.[11,13] Similar data are not available for βS_{HC}. The $^3J_{HNH\alpha}$ couplings calculated from the PRE-RMD ensembles of αS and βS structures are slightly greater than 5 Hz throughout the sequence, and those calculated from the random coil ensembles are slightly less than 5 Hz (Figure 4A,B). Neither set of calculated $^3J_{HNHa}$ couplings for either protein bears a close resemblance to the experimentally measured values, which in general are larger and fluctuate more dramatically along the sequence. Although no experimental data are available for βS_{HC}, $^3J_{HNHa}$ couplings were calculated from the PRE-RMD and random coil ensembles for

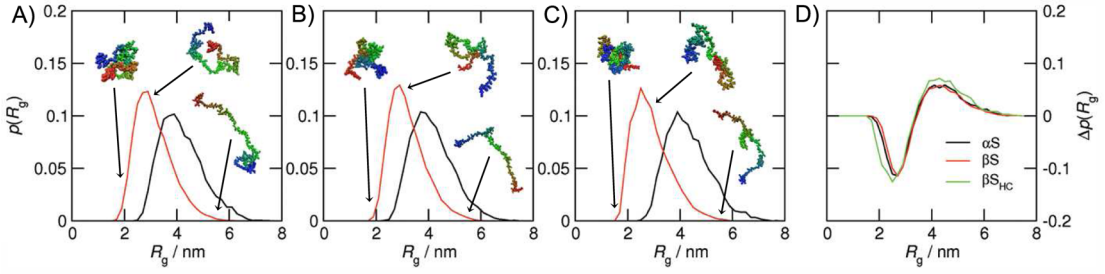


Figure 3. R_g probability distributions for (A) αS, (B) βS, and (C) βS_{HC}. The random coil ensembles (see the text for a definition) are colored black, and the ensembles calculated using PRE-RMD are colored red. Representative structures are shown for various values of R_g. The R_g distributions are shown rather than the R_h distributions because the former are faster to calculate, but the R_h distributions are similar. (D) Distributions of the difference between the random coil and PRE-RMD ensemble R_g probabilities [$\Delta p(R_g) = p(R_g^{\text{random coil}}) - p(R_g^{\text{PRE-RMD}})$].

dx.doi.org/10.1021/bi5009326 | Biochemistry 2014, 53, 7170–7183

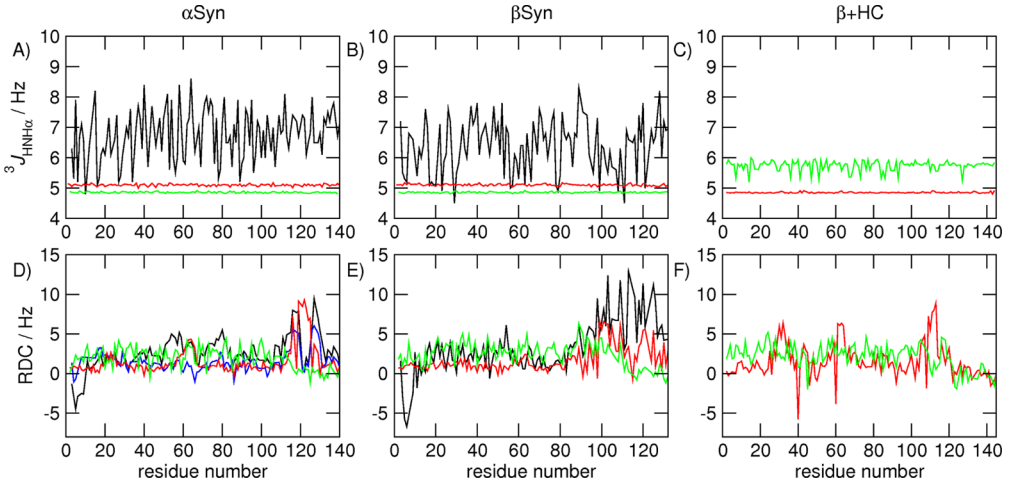


Figure 4. Comparison of experimentally measured[11,13] and calculated NMR observables for (A and D) αS, (B and E) βS, and (C and F) βS_{HC}. (A–C) $^3J_{HNHa}$ couplings (black) measured experimentally, (red) calculated from the PRE-RMD ensembles, and (green) calculated from random coil ensembles. (D–F) Amide N–H RDCs measured experimentally in (black) C8E5/octanol or (blue) Pf1 bacteriophage, (red) calculated from the PRE-RMD ensembles, and (green) calculated from random coil ensembles.

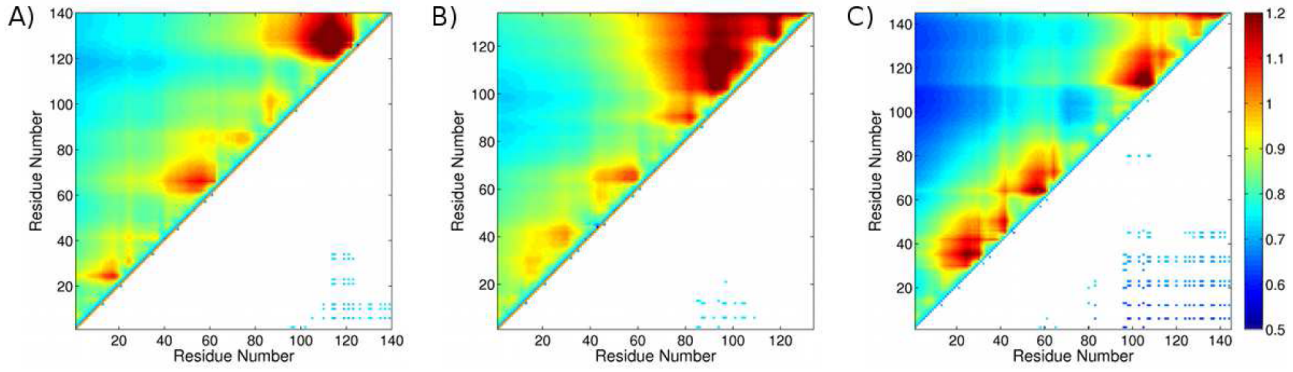


Figure 5. Distance comparison (DC) maps for the (A) αS, (B) βS, and (C) βS_{HC} ensembles determined by PRE-RMD. The top half shows the full DC map, whereas the bottom half shows only the scaled distances that are less than 75% of that expected for a random coil polymer and occur between pairs of oppositely charged residues. The same color scale is used for all the DC maps to aid comparisons.

comparison with those calculated from the αS and βS ensembles. The couplings calculated from the PRE-RMD ensemble are again close to 5 Hz, whereas those calculated from the random coil ensemble lie between 5 and 6 Hz and fluctuate somewhat throughout the sequence, an observation likely to be due to the more compact nature of the βS_{HC} structures inducing more local structure formation.

The magnitudes of the RDCs calculated from both the PRE-RMD and random coil ensembles of αS and βS are more similar to those of the experimental RDCs (Figure 4D,E). The larger RDC values observed experimentally for the C-termini are not found in the values obtained from the random coil ensemble but are detectable in the RDCs calculated from the PRE-RMD ensemble. Again, however, the residue-specific variations in the experimental data are for the most part not accurately reproduced in either of the calculated ensembles.

Overall, the lack of agreement between the experimental $^3J_{HNHa}$ couplings and those calculated from the PRE-RMD ensembles, coupled to the similarity between those calculated from the PRE-RMD and random coil ensembles, suggests that local residue-specific conformational preferences are not well reproduced in the PRE-RMD ensembles. This result is not surprising, given that the type of restraints used provides information about the long-range residual structures of the proteins under investigation, but not about their local conformations. Rather, it should serve as a warning that reproducing experimental data describing one structural aspect of a protein, particularly a disordered protein, does not imply that other structural properties will be accurately described. RDCs report on both local and global structure, so the improved agreement of the PRE-RMD ensemble with the experimental data for the C-termini is likely to reflect the fact that the long-range structure, in the form of the replica-averaged PRE distances, of this ensemble is in good agreement with that observed experimentally. However, the remaining discrepancies, as with the $^3J_{HNHa}$ couplings, most likely result from the fact that the local structure is not well replicated in the calculated ensembles. A more accurate representation of such local conformations should be obtained by using additional restraints, such as ^3J couplings and chemical shifts. These calculations were not performed here because the aspect of primary interest was the comparison of the long-range conformational behaviors of αS, βS, and βS_{HC} and whether any differences observed might be linked to their differing aggregation propensities.

Residual Structure Propensities. The nature of the structures comprising each ensemble is summarized in the distance comparison (DC) maps[17] (Figure 5). In contrast to

dx.doi.org/10.1021/bi5009326 | *Biochemistry* 2014, 53, 7170–7183

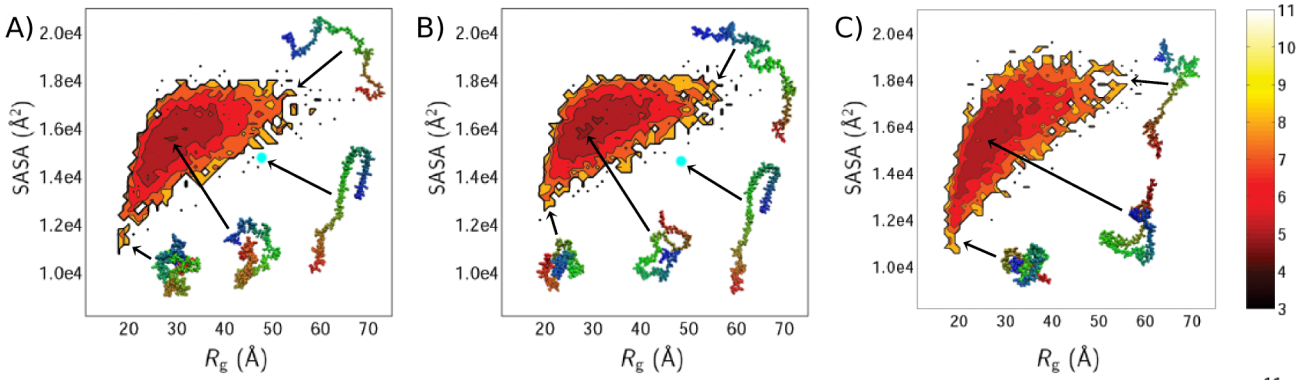


Figure 6. Free energy landscapes of structural ensembles determined for (A) αS, (B) βS, and (C) βS_{HC} ensembles. The free energy is defined as $F(R_g,\mathrm{SASA}) = -\ln p(R_g,\mathrm{SASA})$. Examples of structures found at various points on each landscape are given, and the position of the experimental micelle-bound structure of αS[137] and a homology model of βS based on the αS structure are indicated by filled cyan circles.

the residual contact probability (RCP) maps used previously to characterize disordered state ensembles,[12,48,49] which report on inter-residue distances of <0.85 nm, DC maps reflect the position of the center of the distance distribution relative to that of a random coil. DC values of <1.0 indicate compaction, and those >1.0 represent expansion relative to the random coil. DC maps were used here because unlike RCP maps, they report on aspects of the distribution not accessible experimentally.

For βS_{HC}, the scaled long-range distances are shorter than for either αS or βS, reflecting its larger compaction factor (Table 2). The DC maps for αS, βS, and βS_{HC} (Figure 5), however, all contain distinct regions of inter-residue distances that differ from those expected for a random coil, as indicated by DC values significantly less than or greater than 1.0, suggesting the presence of nonrandom residual structure. In both βS and βS_{HC}, the C-terminal residues exhibiting the shortest distances to residues 1−40 of the N-termini are broadened and include residues located closer to the N-termini compared to that in αS. In βS, the shortest distances are to residues 70−110, and in βS_{HC}, they are to residues 80−145; in αS, they are to residues 100−140. This observation may reflect additional shielding from intermolecular interaction of the central region in βS and βS_{HC}, in keeping with the lower aggregation propensity of both of these polypeptides.

Within the N-termini of all three proteins are clusters of residues close together in sequence separated by distances that are, on average, similar to those observed in a random coil. Such DC values could result from random coil or α-helical structure or some combination of the two, as the expected inter-residue distances are effectively the same for short sequence separations.[127] Additionally, all three proteins, and in particular βS, exhibit distances between residues within the C-terminal regions (residues 100−140 for αS, 100−134 for βS, and 110−145 for βS_{HC}) that are larger on average than in a random coil. This result could be indicative of either extended β-strand-like or PPII structure, each of which is characterized by rms inter-residue distances longer than those of a random flight chain.[127] For βS, PPII structure is most likely to be present, as the C-terminus of βS contains eight proline residues, which are known to disrupt β-sheet formation, and indeed, PPII structure has been observed experimentally.[13] The experimental data for αS, in contrast, suggest a much lower PPII propensity,[13,75] indicating that DC values of >1.0 in the C-terminus of this protein are more likely to correspond to extended β-strand-like structure. Greater β-strand content in αS than in βS is in keeping with the recent observation that αS variants that populate β-strand structure more highly also aggregate faster.[128] There are fewer experimental data available for βS_{HC}, but the cross-peaks in the NMR HSQC spectra overlay with those of βS for a majority of the sequence (Figure S3 of the Supporting Information), indicating that the secondary structure preferences of the C-terminal region of βS_{HC} are likely to be similar to those of βS. Interestingly, the C-terminal region of βS_{HC} does not contain as many DC values greater than 1.0 as βS, suggesting that the insertion of the αS hydrophobic core may have an indirect effect on the structural propensities of this region of the protein.

Free Energy Landscapes. A more global perspective on the nature of the structures sampled by each of the three proteins can be gained by examining the free energy landscapes (Figure 6), which show the probability of the occurrence of different combinations of R_g and solvent accessible surface area (SASA). βS exhibits the narrowest range of SASA and βS_{HC} the widest; this pattern reflects the relationship between the C_f values of the three proteins (Table 2). In all cases, the structures with the lowest R_g values encompass a wide range of SASA values; similarly, there is a large range of R_g values corresponding to the largest SASA values. Thus, having a small R_g poses few restrictions on the fraction of the surface area that is exposed. This may facilitate the role of αS as a hub protein,[130] as a larger surface area allows for a diverse range of binding partners.[2] The greater similarity between the $F(R_g,\mathrm{SASA})$ landscapes of αS and βS suggests that the insertion of the central NAC region into βS_{HC} causes it to behave more like αS in this respect.

Implications for Aggregation. The construction and study of βS_{HC} was initiated with the aim of understanding whether investigation of the transient long-range interactions can provide insight about why the aggregation rate of βS is lower than that of αS.[71] It was originally thought that the fundamental cause of the different aggregation propensities of αS and βS was simply the absence of 11 residues (73−83) from the NAC region of βS[72,131] (Figure 1). Contrary to this expectation, however, βS_{HC}, which contains residues 73−83 of αS within the βS sequence following residue 72, was found to have aggregation properties similar to those of βS.[129] Further investigations, including analysis of the aggregation properties of two deletion mutants, $\alpha\Delta$73−83 and $\alpha\Delta$71−82, showed that the most likely reason for the similar aggregation behavior of βS

dx.doi.org/10.1021/bi5009326 | Biochemistry 2014, 53, 7170−7183

and βS_{HC} is the inclusion of E83 in the βS_{HC} construct.[71] This negatively charged residue is thought to disrupt the intermolecular interactions of the hydrophobic core and may therefore act as an aggregation "gatekeeper".[72,73] It has also been shown that the aggregation properties of αS and βS can be effectively interchanged by swapping six residues among them (63−66, 71 and 72).[74] Further evidence of the role of residue E83 in αS aggregation comes from a study that showed that the interaction of dopamine and related derivatives with residues 125−129 of αS is mediated by electrostatic interactions between the ligand and E83, with replacement of glutamine by alanine preventing dopamine from inhibiting αS aggregation.[132] Additionally, the incorporation of charged residues into the hydrophobic core of full-length αS decreases the rate of fibril formation, suggesting that the lower experimental and theoretical aggregation propensities of βS and βS_{HC}, both of which have a net charge greater than that of αS, may be due to intermolecular repulsion between charged residues.[72,73]

The role of charge in preventing aggregation is not confined to the intermolecular interactions. While any contacts made by the C-terminus with the NAC region are thought to be hydrophobic in nature, interactions with the N-terminus are most likely to be electrostatic.[10] The lower panels of the DC maps (Figure 5) indeed reveal that many of the inter-residue distances that are on average considerably shorter than would be expected for a random coil polymer occur between oppositely charged residues. The increased negative charge of the C-terminal regions of βS and βS_{HC} may therefore enhance these intramolecular electrostatic interactions. Correspondingly, comparison of the DC maps shows that the scaled distances between the N- and C-terminal regions of βS and βS_{HC} are shorter than those of αS (Figure 5), and the bottom panels show that many of the shortest scaled average inter-residue distances in βS_{HC} occur between oppositely charged residues. Moreover, the predicted aggregation propensities of the C-terminal regions of βS and βS_{HC} are even lower than that of αS (Figure 7). In addition to the effects of the long-range

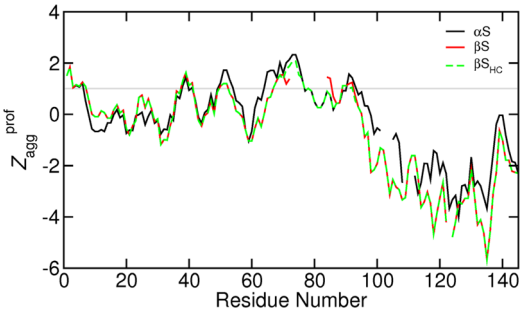


Figure 7. Aggregation propensity, Z_{agg}^{prof}, predicted using the Zyggregator algorithm[123] for (black, solid) αS, (red, solid) βS, and (green, dashed) βS_{HC}. The residue numbers and gaps correspond to the sequence alignment shown in Figure 1. The gray line at $Z_{agg}^{prof} = 1$ indicates the threshold for classifying a sequence as being aggregation prone; regions exhibiting Z_{agg}^{prof} values greater than this are considered to be aggregation prone.

conformational properties on aggregation, contributions from the local secondary structure propensities can be expected.[128] The importance of electrostatic interactions between the N- and C-terminal regions in determining the aggregation properties is also supported by experimental data. C-Terminal truncation mutants of αS aggregate faster than the wild type only if the truncation removes the majority of the charged residues from the C-terminal region.[133] Additionally, the binding of positively charged polyamines such as spermine to the C-terminal region increases the aggregation rates of βS in SDS and αS in the absence of the addition of SDS.[129,133−136] Thus, features apparent in the PRE-RMD ensembles correlate with the experimental data and provide further support for the suggestion that charge plays a key role in controlling the aggregation propensities of the synucleins.

CONCLUSIONS

We have exploited the opportunities offered by the use of PRE-derived distances as replica-averaged structural restraints in MD simulations to increase the amount of information available from experimental measurements by providing atomic-level structural detail. Analysis of the transient long-range intramolecular interactions shows that the distances between the N- and C-terminal regions of all three proteins are shorter than expected for random coil structures, indicative of interactions between the two regions that may be electrostatic in nature. The resemblance between the structural propensities of βS_{HC} and βS echoes their similar aggregation propensities, with the main difference likely to be related to aggregation between these two proteins and being that αS and βS_{HC} exhibit a greater number of inter-residue distances between the N- and C-terminal regions that are shorter than expected for a random coil. As interactions between the N- and C-terminal regions are expected to be electrostatic in nature, this factor strengthens the case for charge playing a key role in modulating the aggregation properties of these polypeptides.

ASSOCIATED CONTENT

Supporting Information

Overlays of the $^{1}H-^{15}N$ HSCQ spectra measured for S64C βS and βS_{HC} with the attached MTSL spin-label in its oxidized and reduced states, plots showing the change in the ^{15}N and ^{1}H chemical shifts measured for each residue of S64C βS and βS_{HC} when the attached MTSL spin-label is in its oxidized and reduced forms, and the $^{1}H-^{15}N$ HSCQ spectrum of βS_{HC} alone and overlaid with that of βS or of αS. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: j.allison@massey.ac.nz.
*E-mail: mv245@cam.ac.uk.
*E-mail: cmd44@cam.ac.uk.

Present Addresses

[†]J.R.A.: Centre for Theoretical Chemistry and Physics, Institute of Natural and Mathematical Sciences, Massey University, Albany Highway, Auckland 0632, New Zealand.
[‡]R.C.R.: Office of Cancer Clinical Proteomics Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892-2580.
[§]J.C.C.: Institute of Structural and Molecular Biology, University College London and Birkbeck College, Gower Street, London WC1E 6BT, U.K.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The provision of the ^3J couplings and residual dipolar couplings by Carlos Bertoncini prior to publication is gratefully acknowledged. We are grateful for financial support from the Woolf Fisher Trust (J.R.A.), the Gates Cambridge Trust (R.C.R), the Royal Society (M.V.) and the Wellcome Trust (C.M.D.).

■ REFERENCES

(1) Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol. 6*, 197−208.
(2) Gunasekaran, K., Tsai, C. J., Kumar, S., Zanuy, D., and Nussinov, R. (2003) Extended disordered proteins: Targeting function with less scaffold. *Trends Biochem. Sci. 28*, 81−85.
(3) Uversky, V. N., and Dunker, A. K. (2010) Understanding protein non-folding. *Biochim. Biophys. Acta 1804*, 1231−1264.
(4) Babu, M. M., van der Lee, R., de Groot, N. S., and Gsponer, J. (2011) Intrinsically disordered proteins: Regulation and disease. *Curr. Opin. Struct. Biol. 21*, 432−440.
(5) Tompa, P. (2012) Intrinsically disordered proteins: A 10-year recap. *Trends Biochem. Sci. 37*, 509−516.
(6) Chiti, F., and Dobson, C. M. (2006) Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem. 75*, 333−366.
(7) Dobson, C. M. (2001) The structural basis of protein folding and its links with human disease. *Philos. Trans. R. Soc., B 356*, 133−145.
(8) Irvine, G. B., El-Agnaf, O. M., Shanker, G. M., and Walsh, D. M. (2008) Protein aggregation in the brain: The molecular basis for Alzheimer's and Parkinson's disease. *Mol. Med. 14*, 451−464.
(9) Knowles, T. P. J., Vendruscolo, M., and Dobson, C. M. (2014) The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol. 15*, 384−396.
(10) Bernadó, P., Bertoncini, C. W., Griesinger, C., Zweckstetter, M., and Blackledge, M. (2005) Defining long-range order and local disorder in native α-synuclein using residual dipolar couplings. *J. Am. Chem. Soc. 127*, 17968−17969.
(11) Bertoncini, C. W., Jung, Y. S., Fernandez, C. O., Hoyer, W., Griesinger, C., Jovin, T. M., and Zweckstetter, M. (2005) Release of long-range tertiary interactions potentiates aggregation of natively unstructured α-synuclein. *Proc. Natl. Acad. Sci. U.S.A. 102*, 1430−1435.
(12) Dedmon, M. M., Lindorff-Larsen, K., Christodoulou, J., Vendruscolo, M., and Dobson, C. M. (2005) Mapping long-range interactions in α-synuclein using spin-label NMR and ensemble molecular dynamics simulations. *J. Am. Chem. Soc. 127*, 476−477.
(13) Bertoncini, C. W., Rasia, R. M., Lamberto, G. R., Binolfi, A., Zweckstetter, M., Griesinger, C., and Fernandez, C. O. (2007) Structural characterization of the intrinsically unfolded protein β-synuclein, a natural negative regulator of α-synuclein aggregation. *J. Mol. Biol. 372*, 708−722.
(14) Mukrasch, M. D., Markwick, P., Biernat, J., von Bergen, M., Bernadó, P., Griesinger, C., Mandelkow, E., Zweckstetter, M., and Blackledge, M. (2007) Highly populated turn conformations in natively unfolded tau protein identified from residual dipolar couplings and molecular simulation. *J. Am. Chem. Soc. 129*, 5235−5243.
(15) Sung, Y.-h., and Eliezer, D. (2007) Residual structure, backbone dynamics, and interactions within the synuclein family. *J. Mol. Biol. 372*, 689−707.
(16) Jensen, M. R., Markwick, P. R., Meier, S., Griesinger, C., Zweckstetter, M., Grzesiek, S., Bernadó, P., and Blackledge, M. (2009) Quantitative determination of the conformational properties of partially folded and intrinsically disordered proteins using NMR dipolar couplings. *Structure 17*, 1169−1185.
(17) Allison, J. R., Varnai, P., Dobson, C. M., and Vendruscolo, M. (2009) Determination of the free energy landscape of α-synuclein using spin label nuclear magnetic resonance measurements. *J. Am. Chem. Soc. 131*, 18314−18326.
(18) Fisher, C. K., Huang, A., and Stultz, C. M. (2010) Modeling intrinsically disordered proteins with Bayesian statistics. *J. Am. Chem. Soc. 132*, 14919−14927.
(19) Jensen, M. R., Salmon, L., Nodet, G., and Blackledge, M. (2010) Defining conformational ensembles of intrinsically disordered and partially folded proteins directly from chemical shifts. *J. Am. Chem. Soc. 132*, 1270−1272.
(20) Camilloni, C., De Simone, A., Vranken, W. F., and Vendruscolo, M. (2012) Determination of secondary structure populations in disordered states of proteins using nuclear magnetic resonance chemical shifts. *Biochemistry 51*, 2224−2231.
(21) Ozenne, V., Schneider, R., Yao, M., Huang, J.-r., Salmon, L., Zweckstetter, M., Jensen, M. R., and Blackledge, M. (2012) Mapping the potential energy landscape of intrinsically disordered proteins at amino acid resolution. *J. Am. Chem. Soc. 134*, 15138−15148.
(22) Sibille, N., Huvent, I., Fauquant, C., Verdegem, D., Amniai, L., Leroy, A., Wieruszeski, J.-M., Lippens, G., and Landrieu, I. (2012) Structural characterization by nuclear magnetic resonance of the impact of phosphorylation in the proline-rich region of the disordered Tau protein. *Proteins: Struct., Funct., Bioinf. 80*, 454−462.
(23) Zhu, M., De Simone, A., Schenk, D., Toth, G., Dobson, C. M., and Vendruscolo, M. (2013) Identification of small-molecule binding pockets in the soluble monomeric form of the Aβ42 peptide. *J. Chem. Phys. 139*, 035101.
(24) Tóth, G., et al. (2014) Targeting the intrinsically disordered structural ensemble of α-synuclein by small molecules as a potential therapeutic strategy for Parkinson's disease. *PLoS One 9*, e87133.
(25) Svergun, D. I., and Koch, M. H. J. (2003) Small-angle scattering studies of biological macromolecules in solution. *Rep. Prog. Phys. 66*, 1735−1782.
(26) Bilsel, O., and Matthews, C. R. (2006) Molecular dimensions and their distributions in early folding intermediates. *Curr. Opin. Struct. Biol. 16*, 86−93.
(27) Wilkins, D. K., Grimshaw, S. B., Receveur, V., Dobson, C. M., Jones, J. A., and Smith, L. J. (1999) Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry 38*, 16424−16431.
(28) Mohana-Borges, R., Goto, N. K., Kroon, G. J. A., Dyson, H. J., and Wright, P. E. (2004) Structural characterization of unfolded states of apomyoglobin using residual dipolar couplings. *J. Mol. Biol. 340*, 1131−1142.
(29) Jensen, M. R., Houben, K., Lescop, E., Blanchard, L., Ruigrok, R. W. H., and Blackledge, M. (2008) Quantitative conformational analysis of partially folded proteins from residual dipolar couplings: Application to the molecular recognition element of Sendai virus nucleoprotein. *J. Am. Chem. Soc. 130*, 8055−8061.
(30) Nodet, G., Salmon, L., Ozenne, V., Meier, S., Jensen, M. R., and Blackledge, M. (2009) Quantitative description of backbone conformational sampling of unfolded proteins at amino acid resolution from NMR residual dipolar couplings. *J. Am. Chem. Soc. 131*, 17908−17918.
(31) Blackledge, M., Bernadó, P., and Jensen, M. R. (2010) *Instrumental Analysis of Intrinsically Disordered Proteins*, pp 89−106, John Wiley & Sons, Inc., New York.
(32) Jensen, M. R., Salmon, L., Nodet, G., Markwick, P., Bernadó, P., and Blackledge, M. (2011) *Protein NMR Spectroscopy: Practical Techniques and Applications*, pp 319−345, John Wiley & Sons, Ltd., New York.
(33) Salmon, L., Jensen, M. R., Bernadó, P., and Blackledge, M. (2012) In *Intrinsically Disordered Protein Analysis* (Uversky, V. N., and Dunker, A. K., Eds.) Vol. *895*, Chapter 9, pp 115−125, Methods in Molecular Biology, Humana Press, Totowa, NJ.
(34) Marsh, J. A., and Forman-Kay, J. D. (2009) Structure and disorder in an unfolded state under nondenaturing conditions from ensemble models consistent with a large number of experimental restraints. *J. Mol. Biol. 391*, 359−374.
(35) Kjaergaard, M., and Poulsen, F. M. (2012) Disordered proteins studied by chemical shifts. *Prog. Nucl. Magn. Reson. Spectrosc. 60*, 42−51.

(36) Camilloni, C., Cavalli, A., and Vendruscolo, M. (2013) Replica-averaged metadynamics. *J. Chem. Theory Comput. 9*, 5610−5617.
(37) Camilloni, C., and Vendruscolo, M. (2014) Statistical mechanics of the denatured state of a protein using replica-averaged metadynamics. *J. Am. Chem. Soc. 136*, 8982−8991.
(38) Battiste, J. L., and Wagner, G. (2000) Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear Overhauser effect data. *Biochemistry 39*, 5355−5365.
(39) Gillespie, J. R., and Shortle, D. (1997) Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. Paramagnetic relaxation enhancement by nitroxide spin labels. *J. Mol. Biol. 268*, 158−169.
(40) Gillespie, J. R., and Shortle, D. (1997) Characterization of long-range structure in the denatured state of staphylococcal nuclease. II. Distance restraints from paramagnetic relaxation and calculation of an ensemble of structures. *J. Mol. Biol. 268*, 170−184.
(41) Cino, E. A., Wong-ekkabut, J., Karttunen, M., and Choy, W.-Y. (2011) Microsecond molecular dynamics simulations of intrinsically disordered proteins involved in the oxidative stress response. *PLoS One 6*, e27371.
(42) Higo, J., Nishimura, Y., and Nakamura, H. (2011) A free-energy landscape for coupled folding and binding of an intrinsically disordered protein in explicit solvent from detailed all-atom computations. *J. Am. Chem. Soc. 133*, 10448−10458.
(43) Sgourakis, N. G., Merced-Serrano, M., Boutsidis, C., Drineas, P., Du, Z., Wang, C., and Garcia, A. E. (2011) Atomic-level characterization of the ensemble of the Aβ(1−42) monomer in water using unbiased molecular dynamics simulations and spectral algorithms. *J. Mol. Biol. 405*, 570−583.
(44) Knott, M., and Best, R. B. (2012) A preformed binding interface in the unbound ensemble of an intrinsically disordered protein: Evidence from molecular simulations. *PLoS Comput. Biol. 8*, e1002605.
(45) Lindorff-Larsen, K., Trbovic, N., Maragakis, P., Piana, S., and Shaw, D. E. (2012) Structure and dynamics of an unfolded protein examined by molecular dynamics simulation. *J. Am. Chem. Soc. 134*, 3787−3791.
(46) Mittal, J., Yoo, T. H., Georgiou, G., and Truskett, T. M. (2013) Structural ensemble of an intrinsically disordered polypeptide. *J. Phys. Chem. B 117*, 118−124.
(47) Best, R. B., and Vendruscolo, M. (2004) Determination of protein structures consistent with NMR order parameters. *J. Am. Chem. Soc. 126*, 8090−8091.
(48) Lindorff-Larsen, K., Kristjansdottir, S., Teilum, K., Fieber, W., Dobson, C. M., Poulsen, F. M., and Vendruscolo, M. (2004) Determination of an ensemble of structures representing the denatured state of the bovine acyl-coenzyme A binding protein. *J. Am. Chem. Soc. 126*, 3291−3299.
(49) Kristjansdottir, S., Lindorff-Larsen, K., Fieber, W., Dobson, C. M., Vendruscolo, M., and Poulsen, F. M. (2005) Formation of native and non-native interactions in ensembles of denatured ACBP molecules from paramagnetic relaxation enhancement studies. *J. Mol. Biol. 347*, 1053−1062.
(50) Francis, C., Lindorff-Larsen, K., Best, R., and Vendruscolo, M. (2006) Characterization of the residual structure in the unfolded state of the Δ131Δ fragment of staphylococcal nuclease. *Proteins: Struct., Funct., Bioinf. 65*, 145−152.
(51) Chen, J. (2009) Intrinsically disordered p53 extreme C-terminus binds to S100B($\beta\beta$) through "Fly-Casting. *J. Am. Chem. Soc. 131*, 2088−2089.
(52) Ganguly, D., and Chen, J. (2009) Atomistic details of the disordered states of KID and pKID. Implications in coupled binding and folding. *J. Am. Chem. Soc. 131*, 5214−5223.
(53) Zhang, W., Ganguly, D., and Chen, J. (2012) Residual structures, conformational fluctuations, and electrostatic interactions in the synergistic folding of two intrinsically disordered proteins. *PLoS Comput. Biol. 8*, e1002353.
(54) Vendruscolo, M., and Dobson, C. M. (2005) Towards complete descriptions of the free-energy landscapes of proteins. *Philos. Trans. R. Soc., A 363*, 433−452.
(55) Vendruscolo, M. (2007) Determination of conformationally heterogeneous states of proteins. *Curr. Opin. Struct. Biol. 17*, 15−20.
(56) Torda, A., Scheek, R., and van Gunsteren, W. F. (1989) Time-dependent distance restraints in molecular dynamics simulations. *Chem. Phys. Lett. 157*, 289−294.
(57) Torda, A. E., Scheek, R. M., and van Gunsteren, W. F. (1990) Time-averaged nuclear Overhauser effect distance restraints applied to Tendamistat. *J. Mol. Biol. 214*, 223−235.
(58) Scheek, R. M., Torda, A. E., Kemmink, J., and van Gunsteren, W. F. (1991) *Computational Aspects of the Study of Biological Macromolecules by NMR*, NATO ASI Series A22, pp 209−217, Plenum Press, New York.
(59) Bonvin, A., Boelens, R., and Kaptein, R. (1994) Time-averaged and ensemble averaged direct NOE restraints. *J. Biomol. NMR 4*, 143−149.
(60) Kemmink, J., van Mierlo, C. P. M., Scheek, R. M., and Creighton, T. E. (1993) Local structure due to an aromatic-amide interaction observed by ^{1}H-nuclear magnetic resonance spectroscopy in peptides related to the N terminus of bovine pancreatic trypsin inhibitor. *J. Mol. Biol. 230*, 312−322.
(61) Mierke, D. F., Scheek, R. M., and Kessler, H. (1994) Coupling constants as restraints in ensemble distance driven dynamics. *Biopolymers 34*, 559−563.
(62) Pitera, J. W., and Chodera, J. D. (2012) On the use of experimental observations to bias simulated ensembles. *J. Chem. Theory Comput. 8*, 3445−3451.
(63) Roux, B., and Weare, J. (2013) On the statistical equivalence of restrained-ensemble simulations with the maximum entropy method. *J. Chem. Phys. 138*, 084107.
(64) Cavalli, A., Camilloni, C., and Vendruscolo, M. (2013) Molecular dynamics simulations with replica-averaged structural restraints generate structural ensembles according to the maximum entropy principle. *J. Chem. Phys. 138*, 094112.
(65) Boomsma, W., Ferkinghoff-Borg, J., and Lindorff-Larsen, K. (2014) Combining experiments and simulations using the maximum entropy principle. *PLoS Comput. Biol. 10*, e1003406.
(66) Uversky, V. N. (2002) Natively unfolded proteins: A point where biology waits for physics. *Protein Sci. 11*, 739−756.
(67) George, J. (2001) The synucleins. *Genome Biol. 3*, 3002.1−3002.6.
(68) Jakes, R., Spillantini, M. G., and Goedert, M. (1994) Identification of two distinct synucleins from human brain. *FEBS Lett. 345*, 27−32.
(69) Park, J. Y., and Lansbury, P. T., Jr. (2003) β-Synuclein inhibits formation of α-synuclein protofibrils: A possible therapeutic strategy against Parkinson's disease. *Biochemistry 42*, 3696−3700.
(70) Tsigelny, I. F., Bar-On, P., Sharikov, Y., Crews, L., Hashimoto, M., Miller, M. A., Keller, S. H., Platoshyn, O., Yuan, J. X.-J., and Masliah, E. (2007) Dynamics of α-synuclein aggregation and inhibition of pore-like oligomer development by β-synuclein. *FEBS J. 274*, 1862−1877.
(71) Rivers, R. C., Kumita, J. R., Tartaglia, G. G., Dedmon, M. M., Pawar, A., Vendruscolo, M., Dobson, C. M., and Christodoulou, J. (2008) Molecular determinants of the aggregation behaviour of α- and β-synuclein. *Protein Sci. 17*, 887−898.
(72) Giasson, B. I., Murray, I. V. J., Trojanowski, J. Q., and Lee, V. M.-Y. (2001) A hydrophobic stretch of 12 amino acid residues in the middle of α-synuclein is essential for filament assembly. *J. Biol. Chem. 276*, 2380−2386.
(73) Du, H. N., Tang, L., Luo, X. Y., Li, H. T., Hu, J., Zhou, J. W., and Hu, H. Y. (2003) A peptide motif consisting of glycine, alanine, and valine is required for the fibrillization and cytotoxicity of human α-synuclein. *Biochemistry 42*, 8870−8878.
(74) Roodveldt, C., Andersson, A., De Genst, E. J., Labrador-Garrido, A., Buell, A. K., Dobson, C. M., Tartaglia, G. G., and Vendruscolo, M. (2012) A rationally designed six-residue swap generates comparability

in the aggregation behavior of α-synuclein and β-synuclein. *Biochemistry 51*, 8771−8778.
(75) Eliezer, D., Kutluay, E., Bussell, R., Jr., and Browne, G. (2001) Conformational properties of α-synuclein in its free and lipid-associated states. *J. Mol. Biol. 307*, 1061−1073.
(76) Morar, A. S., Olteanu, A., Young, G. B., and Pielak, G. J. (2001) Solvent-induced collapse of α-synuclein and acid-denatured cytochrome c. *Protein Sci. 10*, 2195−2199.
(77) Syme, C. D., Blanch, E. W., Holt, C., Jakes, R., Goedert, M., Hecht, L., and Barron, L. D. (2002) A Raman optical activity study of rheomorphism in caseins, synucleins and tau: New insight into the structure and behaviour of natively unfolded proteins. *Eur. J. Biochem. 269*, 148−156.
(78) Bernstein, S. L., Liu, D., Wyttenbach, T., Bowers, M. T., Lee, J. C., Gray, H. B., and Winkler, J. R. (2004) α-Synuclein: Stable compact and extended monomeric structures and pH dependence of dimer formation. *J. Am. Soc. Mass Spectrom. 15*, 1435−1443.
(79) Marsh, J. A., Singh, V. K., Jia, Z., and Forman-Kay, J. D. (2006) Sensitivity of secondary structure propensities to sequence differences between α- and γ-synuclein: Implications for fibrillation. *Protein Sci. 15*, 2795−2804.
(80) McNulty, B. C., Young, G. B., and Pielak, G. J. (2006) Macromolecular crowding in the *Escherichia coli* periplasm maintains α-synuclein disorder. *J. Mol. Biol. 355*, 893−897.
(81) McNulty, B. C., Tripathy, A., Young, G. B., Charlton, L. M., Orans, J., and Pielak, G. J. (2006) Temperature-induced reversible conformational change in the first 100 residues of α-synuclein. *Protein Sci. 15*, 602−608.
(82) Cho, M.-K., Kim, H.-Y., Bernado, P., Fernandez, C. O., Blackledge, M., and Zweckstetter, M. (2007) Amino acid bulkiness defines the local conformations and dynamics of natively unfolded α-synuclein and tau. *J. Am. Chem. Soc. 129*, 3032−3033.
(83) Wu, K.-P., Weinstock, D. S., Narayanan, C., Levy, R. M., and Baum, J. (2009) Structural reorganization of α-synuclein at low pH observed by NMR and REMD simulations. *J. Mol. Biol. 391*, 784−796.
(84) Coelho-Cerqueira, E., Carmo-Gonçalves, P., Sá Pinheiro, A., Cortines, J., and Follmer, C. (2013) α-Synuclein as an intrinsically disordered monomer fact or artefact? *FEBS J. 280*, 4915−4927.
(85) Waudby, C. A., Camilloni, C., Fitzpatrick, A. W. P., Cabrita, L. D., Dobson, C. M., Vendruscolo, M., and Christodoulou, J. (2013) In-cell NMR characterization of the secondary structure populations of a disordered conformation of α-synuclein within *E. coli* cells. *PLoS One 8*, e72286.
(86) Bartels, T., Choi, J. G., and Selkoe, D. J. (2011) α-Synuclein occurs physiologically as a helically folded tetramer that resists aggregation. *Nature 477*, 107−110.
(87) Jónsson, S. A., Mohanty, S., and Irbäck, A. (2012) Distinct phases of free α-synuclein: A Monte Carlo study. *Proteins: Struct., Funct., Bioinf. 80*, 2169−2177.
(88) Salmon, L., Nodet, G., Ozenne, V., Yin, G., Jensen, M. R., Zweckstetter, M., and Blackledge, M. (2010) NMR characterization of long-range order in intrinsically disordered proteins. *J. Am. Chem. Soc. 132*, 8407−8418.
(89) Rao, J. N., Jao, C. C., Hegde, B. G., Langen, R., and Ulmer, T. S. (2010) A combinatorial NMR and EPR approach for evaluating the structural ensemble of partially folded proteins. *J. Am. Chem. Soc. 132*, 8657−8668.
(90) Ullman, O., Fisher, C. K., and Stultz, C. M. (2011) Explaining the structural plasticity of α-synuclein. *J. Am. Chem. Soc. 133*, 19536−19546.
(91) Narayanan, C., Weinstock, D. S., Wu, K.-P., Baum, J., and Levy, R. M. (2012) Investigation of the polymeric properties of α-synuclein and comparison with NMR experiments: A replica exchange molecular dynamics study. *J. Chem. Theory Comput. 8*, 3929−3942.
(92) Mantsyzov, A. B., Maltsev, A. S., Ying, J., Shen, Y., Hummer, G., and Bax, A. (2014) A maximum entropy approach to the study of residue-specific backbone angle distributions in α-synuclein, an intrinsically disordered protein. *Protein Sci. 23*, 1275−1290.
(93) Uversky, V. N., Li, J., Souillac, P., Millett, I. S., Doniach, S., Jakes, R., Goedert, M., and Fink, A. L. (2002) Biophysical properties of the synucleins and their propensities to fibrillate: Inhibition of α-synuclein assembly by β- and γ-synucleins. *J. Biol. Chem. 277*, 11970−11978.
(94) Delaglio, F., Grzesiek, S., Vuister, G., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277−293.
(95) Goddard, T., and Kneller, D. (2008) *SPARKY 3*, University of California, San Francisco.
(96) Allison, J. R. (2008) Computational methods for characterising disordered states of proteins. Ph.D. Thesis, University of Cambridge, Cambridge, U.K.
(97) Brooks, B., Bruccoler, R., Olafson, B., States, D., Swaminathan, S., and Karplus, M. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem. 4*, 187−217.
(98) Swope, W. C., Andersen, H. C., Berens, P. H., and Wilson, K. R. (1982) A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *J. Chem. Phys. 76*, 637−649.
(99) Nose, S. (1984) A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys. 81*, 511−519.
(100) Hoover, W. G. (1985) Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A 31*, 1695−1697.
(101) Reiher, I. W. (1985) Theoretical studies of hydrogen bonding. Ph.D. Thesis, Harvard University, Cambridge, MA.
(102) Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J. Comput. Phys. 23*, 327−341.
(103) Lazaridis, T., and Karplus, M. (1999) Effective energy function for proteins in solution. *Proteins: Struct., Funct., Genet. 35*, 133−152.
(104) Kemmink, J., and Scheek, R. (1995) Dynamic modeling of a helical peptide in solution using NMR data: Multiple conformations and multi-spin effects. *J. Biomol. NMR 6*, 33−40.
(105) Bonvin, A. M., and Brünger, A. T. (1995) Conformational variability of solution nuclear magnetic resonance structures. *J. Mol. Biol. 250*, 80−93.
(106) Vendruscolo, M., Paci, E., Dobson, C. M., and Karplus, M. (2003) Rare fluctuations of native proteins sampled by equilibrium hydrogen exchange. *J. Am. Chem. Soc. 125*, 15686−15687.
(107) Lindorff-Larsen, K., Best, R. B., Depristo, M. A., Dobson, C. M., and Vendruscolo, M. (2005) Simultaneous determination of protein structure and dynamics. *Nature 433*, 128−132.
(108) Clore, G. M., and Schwieters, C. D. (2004) How much backbone motion in ubiquitin is required to account for dipolar coupling data measured in multiple alignment media as assessed by independent cross-validation? *J. Am. Chem. Soc. 126*, 2923−2938.
(109) Clore, G. M., and Schwieters, C. D. (2004) Amplitudes of protein backbone dynamics and correlated motions in a small α/β protein: Correspondence of dipolar coupling and heteronuclear relaxation measurements. *Biochemistry 43*, 10678−10691.
(110) Clore, G. M., and Schwieters, C. D. (2006) Concordance of residual dipolar couplings, backbone order parameters and crystallographic B-factors for a small α/β protein: A unified picture of high probability, fast atomic motions in proteins. *J. Mol. Biol. 355*, 879−886.
(111) Hess, B., and Scheek, R. M. (2003) Orientation restraints in molecular dynamics simulations using time and ensemble averaging. *J. Magn. Reson. 164*, 19−27.
(112) Gsponer, J., Hopearuoho, H., Cavalli, A., Dobson, C., and Vendruscolo, M. (2006) Geometry, energetics, and dynamics of hydrogen bonds in proteins: Structural information derived from NMR scalar couplings. *J. Am. Chem. Soc. 128*, 15127−15135.
(113) Richter, B., Gsponer, J., Varnai, P., Salvatella, X., and Vendruscolo, M. (2007) The MUMO (minimal under-restraining minimal over-restraining) method for the determination of native state ensembles of proteins. *J. Biomol. NMR 37*, 117−135.

(114) Ferrara, P., Apostolakis, J., and Caflisch, A. (2002) Evaluation of a fast implicit solvent model for molecular dynamics simulations. *Proteins: Struct., Funct., Genet. 46*, 24−33.
(115) Fraternali, F., and van Gunsteren, W. F. (1996) An efficient mean solvation force model for use in molecular dynamics simulations of proteins in aqueous solution. *J. Mol. Biol. 256*, 939−948.
(116) Garcia de la Torre, J., Huertas, M. L., and Carrasco, B. (2000) Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys. J. 78*, 719−730.
(117) Bax, A. (2003) Weak alignment offers new NMR opportunities to study protein structure and dynamics. *Protein Sci. 12*, 1−16.
(118) Zhou, H.-X. (2002) Dimensions of denatured protein chains from hydrodynamic data. *J. Phys. Chem. B 106*, 5769−5775.
(119) Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. C. (2005) GROMACS: Fast, flexible, and free. *J. Comput. Chem. 26*, 1701−1718.
(120) Pardi, A., Billeter, M., and Wüthrich, K. (1984) Calibration of the angular dependence of the amide proton-C^{α} proton coupling constants, $^{3}J_{HN\alpha}$, in a globular protein: Use of $^{3}J_{HN\alpha}$ for identification of helical secondary structure. *J. Mol. Biol. 180*, 741−751.
(121) Zweckstetter, M., and Bax, A. (2000) Prediction of sterically induced alignment in a dilute liquid crystalline phase: Aid to protein structure determination by NMR. *J. Am. Chem. Soc. 122*, 3791−3792.
(122) Lee, B., and Richards, F. M. (1971) The interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol. 55*, 379−400.
(123) Pawar, A. P., Dubay, K. F., Zurdo, J., Chiti, F., Vendruscolo, M., and Dobson, C. M. (2005) Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases. *J. Mol. Biol. 350*, 379−392.
(124) DuBay, K. F., Pawar, A. P., Chiti, F., Zurdo, J., Dobson, C. M., and Vendruscolo, M. (2004) Prediction of the absolute aggregation rates of amyloidogenic polypeptide chains. *J. Mol. Biol. 341*, 1317−1326.
(125) Fusco, G., De Simone, A., Gopinath, T., Vostrikov, V., Vendruscolo, M., Dobson, C. M., and Veglia, G. (2014) Direct observation of the three regions in α-synuclein that determine its membrane-bound behaviour. *Nat. Commun. 5*, 3827.
(126) Binolfi, A., Rasia, R. M., Bertoncini, C. W., Ceolin, M., Zweckstetter, M., Griesinger, C., Jovin, T. M., and Fernandez, C. O. (2006) Interaction of α-synuclein with divalent metal ions reveals key differences: A link between structure, binding specificity and fibrillation enhancement. *J. Am. Chem. Soc. 128*, 9893−9901.
(127) Zagrovic, B., and Pande, V. S. (2003) Structural correspondence between the α-helix and the random-flight chain resolves how unfolded proteins can have native-like properties. *Nat. Struct. Biol. 10*, 955−961.
(128) Camilloni, C., and Vendruscolo, M. (2013) A relationship between the aggregation rates of α-synuclein variants and the β-sheet populations in their monomeric forms. *J. Phys. Chem. B 117*, 10737−10741.
(129) Rivers, R. C. (2007) Biophysical analysis of the aggregation behaviour and structural properties of α- and β-synuclein. Ph.D. Thesis, University of Cambridge: Cambridge, UK.
(130) Dunker, A. K., Cortese, M. S., Romero, P., Iakoucheva, L. M., and Uversky, V. N. (2005) Flexible nets. The roles of intrinsic disorder in protein interaction networks. *FEBS J. 272*, 5129−5148.
(131) Biere, A. L., Wood, S. J., Wypych, J., Steavenson, S., Jiang, Y., Anafi, D., Jacobsen, F. W., Jarosinski, M. A., Wu, G. M., Louis, J. C., Martin, F., Narhi, L. O., and Citron, M. (2000) Parkinson's disease-associated α-synuclein is more fibrillogenic than β- and γ-synuclein and cannot cross-seed its homologs. *J. Biol. Chem. 275*, 34574−34579.
(132) Herrera, F. E., Chesi, A., Paleologou, K. E., Schmid, A., Munoz, A., Vendruscolo, M., Gustincich, S., Lashuel, H. A., and Carloni, P. (2008) Inhibition of α-synuclein fibrillization by dopamine is mediated by interactions with five C-terminal residues and with E83 in the NAC region. *PLoS One 3*, e3394.
(133) Murray, I. V., Giasson, B. I., Quinn, S. M., Koppaka, V., Axelsen, P. H., Ischiropoulos, H., Trojanowski, J. Q., and Lee, V. M. (2003) Role of α-synuclein carboxy-terminus on fibril formation in vitro. *Biochemistry 42*, 8530−8540.
(134) Antony, T., Hoyer, W., Cherny, D., Heim, G., Jovin, T. M., and Subramaniam, V. (2003) Cellular polyamines promote the aggregation of α-synuclein. *J. Biol. Chem. 278*, 3235−3240.
(135) Fernández, C. O., Hoyer, W., Zweckstetter, M., Jares-Erijman, E., Subramaniam, V., Griesinger, C., and Jovin, T. M. (2004) NMR of α-synuclein-polyamine complexes elucidates the mechanism and kinetics of induced aggregation. *EMBO J. 23*, 2039−2046.
(136) Li, W., West, N., Colla, E., Pletnikova, O., Troncoso, J. C., Marsh, L., Dawson, T. M., Jakala, P., Hartmann, T., Price, D. L., and Lee, M. K. (2005) Aggregation promoting C-terminal truncation of α-synuclein is a normal cellular process and is enhanced by the familial Parkinson's disease-linked mutations. *Proc. Natl. Acad. Sci. U.S.A. 102*, 2162−2167.
(137) Ulmer, T. S., Bax, A., Cole, N. B., and Nussbaum, R. L. (2005) Structure and dynamics of micelle-bound human α-synuclein. *J. Biol. Chem. 280*, 9595−9603.