HHS Public Access

Author manuscript

J Phys Chem B. Author manuscript; available in PMC 2023 September 15.

Exploring Structural Flexibility and Stability of α-Synuclein by the Landau-Ginzburg-Wilson Approach

Anatolii Korneev[1], **Alexander Begun**[1], **Sergei Liubimov**[1], **Khatuna Kachlishvili**[2], **Alexander Molochkov**[1], **Antti J. Niemi**[1,3,4,5], **Gia G. Maisuradze**[2,*]

[1]Pacific Quantum Center, Far Eastern Federal University, 690922, Vladivostok, Russia

[2]Baker Laboratory of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853-1301, U. S. A.

[3]Nordita, Stockholm University and Uppsala University, SE-106 91 Stockholm, Sweden

[4]Laboratoire de Mathematiques et Physique Theorique CNRS UMR 6083, Fédération Denis Poisson, Université de Tours, F37200, Tours, France

[5]School of Physics, Beijing Institute of Technology, Beijing 100081, P.R. China

Abstract

α-Synuclein (αS) is the principal protein component of the Lewy body and Lewy neurite deposits that are found in the brains of the victims of one of the most prevalent neurodegenerative disorders, Parkinson's disease. The αS can be qualified as a chameleon protein because of the large number of different conformations that it is able to adopt: it is disordered under physiological conditions in solution, in equilibrium with a minor α-helical tetrameric form in the cytoplasm, and is α-helical when bounded to a cell membrane. Also, *in vitro*, αS forms polymorphic amyloid fibrils with unique arrangements of cross-β-sheet motifs. Therefore, it is of interest to elucidate the origins of the structural flexibility of αS, and what makes αS stable in different conformations. We address these questions here by analyzing the experimental structures of the micelle-bound, tetrameric and fibrillar αS in terms of a kink (heteroclinic standing wave solution) of a generalized discrete nonlinear Schrödinger equation. It is illustrated that without molecular dynamics simulations the kinks are capable to identify the key residues causing structural flexibility of αS. Also, the stability of the experimental structures of αS is investigated by simulating heating/cooling trajectories using the Glauber algorithm. The findings are consistent with experiments.

[*]To whom correspondence may be addressed: gm56@cornell.edu, Telephone: 1-607-255-4399.

Author contributions: G.G.M. designed research; A.K., A.B., and S.L. performed research; A.K., A.B., S.L., K.K., A.M., A.J.N., and G.G.M. analysed data; G.G.M. wrote the paper.

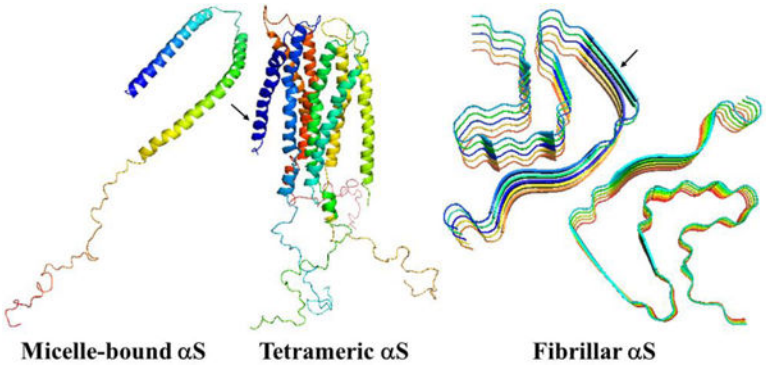
The authors declare no competing financial interest.

Associated Content

Supporting Information

The Supporting Information is available free of charge at the https://pubs.acs.org Protein backbone geometry. The experimental and three calculated representative structures for the micelle-bound, tetrameric and fibrillar αS. The radius of gyrations and helicity (in %) as functions of Glauber temperature for the micelle-bound and tetrameric αS. The radius of gyrations and β-strand contents (in %) as functions of Glauber temperature for the fibrillar αS.

Graphical Abstract



Introduction

α-Synuclein (αS) is a small (140 residue) intrinsically disordered protein (IDP) expressed in neurons and presynaptic nerve terminals.[1] It is abundant in the neuronal cytosol of a healthy brain, and its function is thought to be associated with axonal transport.[2,3] The assembly of monomeric αS into amyloid fibrils that form Lewy bodies (LBs) and Lewy neurites is a hallmark of Parkinson's disease (PD),[4] the second most common neurodegenerative disorder (after Alzheimer's disease) nowadays with a significant rise in the number of affected patients and cost of care.[5] αS is characterized by three domains: (i) an amphipathic N-terminal region dominated by four 11- residue repeats including the consensus sequence KTKEGV (residues 1−60); (ii) a central hydrophobic region, which includes the nonamyloid-β component (NAC) region and two additional repeats, involved in protein aggregation (residues 61−95); and (iii) a highly acidic and proline-rich region which has no distinct structural propensity (residues 96−140).

From a structural point of view, αS is a chameleon protein;[6,7] i.e., it is a disordered monomer in solution, which undergoes a large-scale disorder-to-helix transition upon binding to vesicles of different types,[8,9,10] and also can adopt a folded tetrameric state that has a relatively high helical content under physiological conditions.[11,12] When incubated under physiological conditions *in vitro*, αS aggregates into fibrils with polymorphic cross-β-sheet conformations, in which a core of β-strands is aligned perpendicular to the fibril axis forming extended regular β-sheets with different arrangements.[13,14,15,16] In addition to cylindrical fibrils, ribbon aggregates have also been observed.[17] Therefore, it is of interest to know (i) what governs the structural flexibility of αS; (ii) the actual stability of these structures; (iii) what underlies the aforementioned structural stability.

It should be noted that αS has been the subject of not only extensive experimental[8–17] but also theoretical[6,18,19,20,21] studies, in which molecular dynamics (MD) simulations were used to complement ensemble-averaged experiments. Although MD simulations have proven to be a very powerful tool for examining protein folding dynamics, the correctness of such simulations depends on the force field employed to describe physical interactions within and between peptide units. The present force fields, both all-atom and coarse-grained, are far from being accurate. Even small errors in the description of protein energy surfaces

can accumulate over a polypeptide chain to distort the correct fold. The inaccuracy problem becomes more significant when the folding of weakly stable proteins is studied, in which barriers between folded and unfolded states are very low. Therefore, in order to answer the abovementioned question, in this study, we investigate the experimental structures of the micelle-bound,[8] tetrameric,[12] and fibrillar[14] αS from another, complementary point of view. In particular, instead of analyzing individual interactions that contribute to the formation of the folded structure, model-independent principles that are based on symmetry are explored. We suggest that all the physical forces, both strong and weak, combine together to give rise to a particular type of protein dynamics, expressed by a generalized version of the discrete nonlinear Schrödinger (DNLS) equation, and propose that various parametrizations of a kink, or heteroclinic standing wave solution, of a generalized version of the DNLS equation to be utilized as the basic modular building blocks of folded proteins.[22,23,24,25,26,27,28,29,30,31,32] The DNLS equation is the paradigm nonlinear equation that describes kinks as soliton solutions. The specific form that we use, is selected as it is the simplest possible form that both supports kinks as stable soliton solutions and introduces chirality of the C^{α} backbone. The main limitation would come from the complexity of loop structure, when the loops become so long that it is no longer practical to decompose them in terms of individual kinks. But that is rarely the case of globular proteins, with regular secondary structures such as helices and strands between loops. Proteins are, of course, examples of linear polymers, but a generic polymer often does not have identifiable super-secondary structures, like globular proteins do in their collapsed phase. Thus, it can be more difficult to reliably identify the kink structures.

Note that, unlike some recent studies,[22–28,30] in which the epithet topological (dark) soliton was used, here we prefer to use kink. This choice highlights that the potential in the DNLS equation displays a spontaneous breakdown of a discrete symmetry, and the kink describes the ensuing domain wall;[33,34,35,36,37] in the case of a folded protein, an individual kink corresponds to a supersecondary structure such as a helix−loop−helix or strand-loop-strand motif. Also, the kink considered here has no direct relationship with the concept of Davydov's soliton.[38] The kink in a protein is a purely topological (geometric) structure, and the Davydov's soliton, in turn, is associated with collective excitations traveling along a chain (perturbations of electron density). In the end, we should mention that proposed approach has been successfully used for studying the dynamics of IDPs, as well.[27,30,32]

Methods

In this work, we are interested in describing the geometry of a protein backbone in terms of a piecewise linear polygonal framed chain, as seen from the point of view of an imaginary observer who traverses the backbone by moving between the C^{α} atoms. For this, we employ the purely geometrically determined discrete of the Frenet framing,[25,29] described in the Supporting Information. We emphasize that these frames depend only on the C^{α} backbone geometry, and neither the side-chain structure nor the individual peptide plane structure are involved in their definition.

Kink of the DNLS Equation.

According to Anfinsen's dogma,[39] the protein in the native state locates at the global minimum of Helmholtz free energy

$$F = U - TS \tag{1}$$

where U is the internal energy, S is the entropy and T is the temperature. The angles κ_i and γ_i are adopted as the structural variables of the free energy because of their intrinsic connection with the coordinates of C^{α} atoms. Considering the $\mathbb{Z}_2$ gauge symmetry shown in Eq. (S8), the Landau-Ginzburg-Wilson free energy F is naturally introduced to describe the backbone C^{α} chain of protein.[22] When the deformations of protein around the energy minimum keep slow and small, the free energy can be expended as follows:

$$F = -\sum_{i=1}^{N-1} 2\kappa_{i+1}\kappa_i + \sum_{i=1}^{N}\left\{2\kappa_i^2 + \lambda\left(\kappa_i^2 - m^2\right)^2 + \frac{q}{2}\kappa_i^2\gamma_i^2 - p\gamma_i + \frac{r}{2}\gamma_i^2\right\} \tag{2}$$

where λ, q, p, r, and m are parameters. The detailed derivation of Eq. (2) can be found in Refs. [22,23,29]; here it suffices to state that this free energy can be shown to relate to the long-distance limit that describes the full microscopic energy of a folded protein in the universal sense of Refs. [40,41,42,43]. As such, it does not explain the details of the (sub)atomic level mechanisms that give rise to protein folding.

In order to determine the kink content of the backbones, at first, we should define positions of the inflection points (kink centers). We can find the inflection points by analyzing the three-dimensional structure (protein topology) and the experimental values of the angles (κ_i, γ_i). Inflection points correspond to the centers of kinks. Then, the angular spectrum should be reconstructed using the $\mathbb{Z}_2$ transformation at the kink centers (it should be noted that the three-dimensional shape of the protein is invariant under $\mathbb{Z}_2$ transformation). After that, for a given kink structure, we look for a minimum of global energy using a combination of simulated annealing and gradient descent methods. For this proposal a special toolkit for protein structure visualization and analysis was developed [https://protoin.ru/propro/index.php].[44]

In short, the search for the minimum energy occurs in the following order. The virtual-bond angles κ are first extended to negative values, using the $\mathbb{Z}_2$ gauge symmetry [Eq. (S8)]. The virtual-bond-dihedral angles γ are then expressed as functions of the virtual-bond angles κ

$$\gamma_i[\kappa] = \frac{p}{r + q\kappa_i^2} \equiv \frac{u}{1 + v\kappa_i^2} \tag{3}$$

with $u = p/r$ and $v = q/r$. By inserting Eq. (3) into Eq. (2), the virtual-bond-dihedral angles γ are eliminated and the following system of equations for the motion of the virtual-bond angles κ is obtained:

$$\kappa_{i+1} = 2\kappa_i - \kappa_{i-1} + \frac{dV[\kappa]}{d\kappa_i^2}\kappa_i \quad (i = 1, \ldots, N), \tag{4}$$

where $\kappa_0 = \kappa_{N+1} = 0$ and

$$V[\kappa] = \frac{p}{r + q\kappa^2} + 2\left(1 - \lambda m^2\right)\kappa^2 + \lambda\kappa^4, \tag{5}$$

where the familiar structure of the generalized DNLS equation is recognized.[22–32,44] The kink solution to Eq. (4) can be constructed numerically by following the iterative procedure of Ref. [23]. But its explicit form has not so far been found in terms of elementary functions. However, an excellent approximation is obtained by naively discretizing the heteroclinic standing wave solution to the continuum nonlinear Schrödinger equation[22–32,44]

$$\kappa_i = \frac{b \quad exp[\sigma_1(i-s)] + a \quad exp[-\sigma_2(i-s)]}{exp[\sigma_1(i-s)] + exp[-\sigma_2(i-s)]} \tag{6}$$

and with $\gamma_i[\kappa]$ evaluated from Eq. (3). Here, s is a parameter that determines the center of the kink. The $a, b \in [0,\pi]$ mod(2π) are parameters which determine the amplitude of the variation of κ and the asymmetry of the inflection regions; they correspond to the minima of the potential energy contribution V $[\kappa]$ in Eq. (5). The parameters σ_1 and σ_2 are related to the inverse of the range of the kink. It should be noted that, in the case of proteins, the values of a, b are determined entirely by the adjacent helices and strands. Far away from the center of the kink (see Figure 6 in Ref. [29])

$$\kappa_i \to \begin{cases} b & mod(2\pi) \quad i > s \\ a & mod(2\pi) \quad i < s \end{cases} \tag{7}$$

and according to Eqs. (S9) and (S10), the asymptotic values

$$\kappa_i \approx \pi/2 \quad \text{or} \quad -\pi/2 \quad \text{and} \quad \kappa_i \approx 1 \quad \text{or} \quad -1 \tag{8}$$

correspond to the α-helix or β-strand, respectively. In order to satisfy the monotonic character of the profile of Eq. (6), the experimentally measured values of κ_i have to vary monotonically along the amino-acid sequence. Otherwise, a multiple of 2π is added to the experimental values. This does not affect the backbone geometry because κ_i's are defined mod (2π). The σ_1 and σ_2 are intrinsically specific parameters for a given loop. But they specify only the length of the loop, not its shape which is defined by the functional form of Eq. (6), and in the case of a and b, they are combinations of the parameters in Eq. (5).

In Eq. (3) for the virtual-torsion angles, γ_i, there are only two independent parameters u and v. As a consequence, the profile of γ_i is determined entirely by the profile of κ_i, and on the structure of the adjacent regular secondary structures.

Finally, we introduce the concept of folding index of a protein backbone. The formation, evolution and structure of a loop along a folding protein can be monitored in terms of topologically determined folding indices. Folding index[28,29] is defined by following equation:

$$Ind_f = \left[\frac{\Gamma}{\pi}\right] \tag{9}$$

$$\Gamma = \sum_{i=n_1+2}^{n_2-2} \begin{cases} \gamma_{i,i+1} - \gamma_{i-1,i} - 2\pi & if \quad \gamma_{i,i+1} - \gamma_{i-1,i} > \pi \\ \gamma_{i,i+1} - \gamma_{i-1,i} + 2\pi & if \quad \gamma_{i,i+1} - \gamma_{i-1,i} < -\pi \\ \gamma_{i,i+1} - \gamma_{i-1,i} & otherwise \end{cases} \tag{10}$$

where $[x]$ denotes the integer part of x, and Γ is the total rotation angle (in radians) that the projections of the C^{α} atoms of the consecutive loop residues make around the north pole. The folding index is a positive integer when the rotation is counterclockwise, and a negative integer when the rotation is clockwise. The folding index classifies loop structures and entire folded proteins in terms of its values.

Results and Discussion

Scrutiny of the Experimental Structures of the Micelle-Bound, Tetrameric, and Fibrillar α-Synuclein in Terms of Kinks.

The study of kinks is more easily understood from a differential representation of the protein backbone. Assuming constant virtual bond distances between the C^{α} atoms, the free-energy of the polymer can be written in terms of (κ, γ) internal coordinates using a Landau-Ginzburg-Wilson free-energy model[22–32] [Eq. (2)]. Minimizing this free-energy relative to κ and γ lead to a nonlinear equation of motion for the curvature similar to a generalized DNLS which admits kinks as particular solutions.

In earlier studies,[22–32] the modular building blocks of folded proteins, i.e. super-secondary structure such as a helix-loop-helix or strand-loop-strand motif, have been described in terms of the kink of a generalized version of the DNLS equation. It was found that the formation of a kink is initiated by an abrupt change in the orientation of a pair of consecutive side chains in the loop region.[29] Moreover, kink analysis enabled us to realize the importance of local interactions, specifically the bimodal character of the potential of mean force in virtual-bond angles κ, as the driving force of folding.[29] It should be noted that the only long-range interaction present in the Landau-Ginzburg-Wilson Hamiltonian[22–32] is a step-wise Pauli exclusion that introduces self-avoidance and prevents chain crossing. The effects of the long-range Coulomb and Van der Waals interactions are accounted for by the global multi-kink profile resulting from the minimization of free energy.[45] The multi-kink structure describes both local and collective motions. Previously, it was shown that by using kink parameters that are derived from the experimental folded protein structure, its folding can be simulated,[27,44,46,47,48] as well as without MD simulations the kinks are able (i) to capture the sites of the protein that govern the system to change the folding scenario[31] and (ii) to identify the phosphorylated sites of protein, the key players in phosphorylation-induced folding.[32] Also, phase diagram of protein with temperature (T) and acidity (pH) as the thermodynamic variables can be constructed as well.[46] Therefore, it is of interest to investigate whether a kink analysis of the experimental structures[8,12,14] can uncover the

mechanisms of folding, misfolding and aggregation, as well as the reasons for the structural flexibility and stability of α-synuclein.

Kink structures of the micelle-bound, tetrameric and fibrillar α-synuclein.—In the present work, we have carried out a detailed analysis of experimental structures of the micelle-bound αS (PDB ID: 1XQ8),[8] the first chain of tetrameric αS (a structure generously provided by Dr. Thomas Pochapsky),[12] and A-chain of fibrillar αS (PDB ID: 6CU7)[14] in terms of kinks. It should be noted that, of the 140 residues present in the experimental structures of the micelle-bound and tetrameric αS, 95-residue fragments (Met1 – Val95) were selected (C terminal, which has no distinct structural propensity, was removed, because it is extremely difficult to select kinks that give meaningful results for this region of the system), whereas entire experimental structure (Leu38 – Lys97) was selected in the fibrillar αS. For the sake of simplicity below these truncated and in two cases incomplete (only one chain is examined) systems will be called as the micelle-bound, tetrameric and fibrillar αS. Using experimental values of the κ and γ angles, we started resolving the kink structures of these proteins. The multi-kink is a configuration that minimizes free energy [Eq. (2)], with an acceptable small deviation, the root-mean-square deviation (RMSD), from the three-dimensional experimental structure. Minimum energy, as well as RMSD fitting, is achieved using a combination of simulated annealing and gradient descent methods. Multi-kink structures that are combinations of 5, 8 and 9 individual kinks, differing in the arrangement along the protein chain, were obtained for the micelle-bound, tetrameric, and fibrillar α-synuclein, respectively. With RMSD < 1.0 Å, a total of 31 structures were initially obtained for the micelle-bound αS, 16 structures for the tetrameric αS, and 8 structures for the fibrillar αS. For each system, three representative structures were then selected for a detailed analysis. In particular, the structures with (i) RMSD = 0.79Å, 0.81Å, 0.87Å for the micelle-bound αS, (ii) RMSD = 0.94Å, 0.95Å, 0.96Å for the tetrameric αS, and (iii) RMSD = 0.76Å, 0.76Å, 0.78Å for the fibrillar αS. Figure 1 illustrates comparison of the κ and γ angles, experimental and calculated, along the chains of three selected structures of the micelle-bound, tetrameric, and fibrillar α-synuclein. In addition, for each protein, Figure 2 shows experimental values of the κ and γ angles in the first panels, and most representative kink structures in the three remaining panels. Data on this figure allow us to scrutinize the correlations between the location of kinks and folding/misfolding paths. [The experimental and three computed representative structures for the micelle-bound, tetrameric, and fibrillar αS are shown in Fig. S3 (Supporting Information)].

Figure 2 illustrates that not only the number of the kinks, but also the kink locations along the sequences of the αS are changing, which indicate the structural flexibility and weak stability of the system. In our recent work,[19] by analyzing the all-atom MD trajectories of the monomeric (including the micelle-bound αS) and tetrameric αS in terms of the local and global motions, correlations between the main-chain and the side-chain motions and steric parameters along the amino-acid sequence, we were able to elucidate a fundamental relationship between monomers and tetramers. It was found that the sites with high values of the correlation coefficient and with low values of the steric parameter are responsible for high-flexibility and disorder of the αS.[19] The locations of the kink centers for the micelle-bound and tetrameric αS [see Fig. 2 (A, B)] coincide with the peaks on the correlation

coefficient curves and with the minima on the steric parameter curves. These correlations indicate that each chain separately in the tetramer is more unstable than the micelle-bound monomer, which is more proof that the αS in the monomeric form is an IDP. Moreover, if we extrapolate these correlations on the fibrillar αS, then the here presented results (see Fig. 2 C) indicate an instability of each chain separately in the fibrillar αS, which is in harmony with experiment stating that fibrils are repositories of soluble intermediates and a source of harmful oligomers.[49] It should be noted that all 5 kinks obtained in the micelle-bound αS are present in the tetrameric and fibrillar αS with small shifts of the kink centers caused by different "environments". The three (39/40, 44, 66), six (26, 40, 45, 61/65/66, 75, 85) and eight (40/41, 47, 51/52, 58, 67, 73, 81, 87) kink centers of the micelle-bound, tetrameric and fibrillar αS, respectively, pertain to seven imperfect 11-mer repeats; and one (44), four (26, 45, 61, 85) and three (47, 58, 81) kink centers of the micelle-bound, tetrameric and fibrillar αS, respectively, pertain to the conserved αS "KTKEGV" repeat motifs.[50] Also, the kinks are able to identify the important sites, within an accuracy of one residue, mutations of which (E46K and A53T) decrease tetramer/monomer ratios and αS solubility, and induce neurotoxicity.[51] In particular, the center of the fourth kink (at 45th residue) in the tetrameric αS and centers of the second and third kinks (at 47th and 52nd residues) in the fibrillar αS are shifted from these sites by one residue.

Folding index for the micelle-bound, tetrameric, and fibrillar αS.—Figure 3 illustrates the accumulation of the folding index[28,29] [Eqs. 9 and 10] along the trajectories for the micelle-bound (A), tetrameric (B), and fibrillar (C) αS, together with the folding index dependence on length of the protein (D).

As was expected, two regions corresponding to the right-handed α (A, B) and β (C) structure were explored by the systems. The area explored by the tetrameric αS is larger than areas explored by the micelle-bound and fibrillar αS, which indicates on looseness of the tetrameric structure. Moreover, the "behavior" of the folding index vs residue number for the micelle-bound, tetrameric, and fibrillar αS differ from each other. In particular, one peak at the 38th residue and one significant jump from −0.8 to 1.1 reaching maximum value at the 45th residue can be observed in the folding index graph for the micelle-bound αS (red curve in panel D). In our recent study on dimers of wild type (WT) αS and its mutants,[52] we found that the A53T and E46K mutants exhibit high propensities (the first and second highest peaks, respectively) to aggregate at residue L38. The second significant residue K45 is "a neighbor" of abovementioned important site E46. The shift by one residue might be due to the influence of micelle. One major peak at the 49th residue can be seen in the folding index graph for the tetrameric αS (green curve in panel D), that is one residue away from the H50 residue, the mutation of which (H50Q) enhances the aggregation, secretion, and toxicity of αS.[53] It should be noted that there are some similarities in the behavior of these two folding indices. Three main regions can be picked out: the first and third regions exhibit plateaus with negative and positive values of the folding indices, respectively; the middle region is the most important, acting as a transition region, and contains residues playing a crucial role in the stability of the system. Since the fibrillar αS represents a misfolded structure, it is not surprising that the behavior of the folding index of the fibrillar αS significantly differs from the previous two folded structures of αS. In particular, the folding index of the fibrillar αS

increases gradually along the sequence and exhibits several peaks, which are far from the dominant missense mutation sites. On the contrary, it exhibits a minimum at residue 46 with a zero value for the folding index (blue curve in panel D).

In order to explain the importance of these findings, we should mention that the peak at certain site in the folding index graph indicates that the site plays an essential role in the stability of the protein,[29, 31,32] unless the folding index value at peak is zero. When the folding index vanishes (i.e. it obtains zero value) at a certain site, this is an indication of instability.[29] Based on this, we can conclude that the folding index captures successfully important sites of the system, which are responsible for the stability of αS. These findings are in agreement with previous experimental[50,51,53] and theoretical studies,[52] in which the authors illustrated that the mutations of the sites identified by the folding index decrease the stability, and consequently enhance the aggregation and toxicity of the system. It is not surprising that the folding index value at the 38th residue (one of the peaks in red curve in panel D) is almost zero (~ 0.1), because when αS folds the role of this site in aggregation vanishes. Also, zero value of the folding index at the 46th residue in the fibrillar αS is a good manifestation of how aggregation and fibrillation can change the property of a site in the system.

Glauber algorithm for protein stability.—We examined the stability of the micelle-bound, tetrameric and fibrillar αS by simulating heating and cooling trajectories using the Glauber algorithm.[46] In other words, we combine the Landau-Ginzburg-Wilson approach with Glauber dynamics to study in detail, how these systems unfold and fold. Glauber dynamics is a Markov chain Monte Carlo method, that is widely used to describe near-equilibrium relaxation dynamics of a statistical system towards equilibrium Gibbs state. It should be noted that the Glauber algorithm manages the dynamics of simple spin chain systems coming close to the thermal equilibrium, which also obeys the Arrhenius law.[54] Since proteins can be regarded as spin chains, it is natural to simulate the protein folding in terms of Glauber algorithm.

The Figures 4–6 (panels A-C) illustrate the heating/cooling maps for three calculated structures of the micelle-bound, tetrameric and fibrillar αS, respectively. In particular, they show the systems' disordering/ordering during a heating and cooling simulation cycle as a function of temperature in terms of the average value of torsion angles γ. The first part of the maps (from the beginning up to $\log_{10}T_G \approx -8$) corresponds to the thermal unfolding, the middle part of the maps (from $\log_{10}T_G \approx -8$ to $\log_{10}T_G \approx -8$) is the random walk phase, and the last part of the maps (from $\log_{10}T_G \approx -8$ to the end) corresponds to folding (misfolding for the fibrillar αS) due to protein cooling. The radius of gyrations (R_G), with representative structures of different temperatures, of three calculated structures as functions of Glauber temperature for the micelle-bound, tetrameric and fibrillar αS are plotted in panel D of Figures 4–6.

The maps of the micelle-bound αS [Fig. 4 (A-C)] show that all three calculated structures exhibit more or less similar behavior. Both helices of these three structures correctly react to the heating/cooling process, i.e., γ angles pertaining to α-helices change during thermal unfolding, remain stable in the random walk phase, and return to initial values with cooling.

Among all helices the first helix of the first calculated structure is the most sensitive to the heating process by starting to unfold at $\log_{10}T_G \approx -9$, while others start unfolding at $\log_{10}T_G \approx -8$. Half of the γ angles pertaining to the loops have the same reaction to the heating/cooling process as α-helices, and the other half of the γ angles remain unchanged during the entire heating/cooling process. Only the loop of the first structure exhibits more sensitivity to the heating process by starting to unfold at $\log_{10}T_G \approx -12$, while others start unfolding at $\log_{10}T_G \approx -10$. The R_G vs Glauber temperature plots (panel D in Fig. 4) reflect the differences found in the heating/cooling maps. In particular, the first structure starts unfolding at $\log_{10}T_G \approx -13$ and remains unfolded until $\log_{10}T_G \approx -13$, while the other two start unfolding at $\log_{10}T_G \approx -12$ and remain unfolded until $\log_{10}T_G \approx -12$; consequently, they are folded in the $\log_{10}T_G \approx -17$ - $\log_{10}T_G \approx -13$ and $\log_{10}T_G \approx -13$ - $\log_{10}T_G \approx -17$ temperature regions (first structure), and in the $\log_{10}T_G \approx -17$ - $\log_{10}T_G \approx -12$ and $\log_{10}T_G \approx -12$ - $\log_{10}T_G \approx -17$ temperature regions (second and third structures). The representative structures at six different temperatures illustrate the unfolding/folding pathways of the micelle-bound αS. We also computed helical contents (in %) as functions of temperature for three calculated structures of the micelle-bound αS (Fig. S4), which show how helicity changes with the increase/decrease of temperature.

A reaction of all three calculated structures of the tetrameric αS on the heating/cooling process is similar to the reaction by the micelle-bound αS, however, there are some differences. In particular, loops of all calculated structures of the tetrameric αS are more sensitive to the heating process by starting to unfold at lower temperatures $\log_{10}T_G \approx -14$ (first structure), $\log_{10}T_G \approx -17$ (second structure) and $\log_{10}T_G \approx -12$ (third structure) [Fig. 5 (A-C)], whereas the behavior of α-helices remains almost the same, they start unfolding at $\log_{10}T_G \approx -8.5$. The differences in sensitivity of loops on heating expressed in the heating/cooling maps are not reflected in the R_G vs Glauber temperature plots (panel D in Fig. 5). All three structures start unfolding at $\log_{10}T_G \approx -13$ and remain unfolded until $\log_{10}T_G \approx -13$; hence, they are folded in the $\log_{10}T_G \approx -17$ - $\log_{10}T_G \approx -13$ and $\log_{10}T_G \approx -13$ - $\log_{10}T_G \approx -17$ temperature regions. As in the micelle-bound αS, the representative structures at six different temperatures show the unfolding/folding pathways of the tetrameric αS (panel D in Fig. 5), and helical contents (in %) as functions of temperature for three calculated structures of the tetrameric αS (Fig. S5) illustrate how helicity changes with the increase/decrease of temperature.

The heating/cooling maps for three calculated structures of the fibrillar αS [Fig. 6 (A-C)] are quite versatile compared to the micelle-bound and tetrameric αS. Unlike previous cases, in which helical contents almost disappear (~ 1%) in the random walk phase (Figs. S4 and S5), β-strands and loops in the fibrillar αS partially react to the heating/cooling process, i.e., only part of the γ angles pertaining to β-strands and loops change during thermal unfolding. Plus, these changes are not consistent, manifestation of which are the spectra representing different γ angles in these maps. The parts of β-strands and loops reacting to the heating process start unfolding at $\log_{10}T_G \approx -12$. All three calculated structures exhibit quite similar behavior, which is reflected in the R_G vs Glauber temperature plots (panel D in Fig. 6). In particular, all three structures start unfolding at $\log_{10}T_G \approx -12$ and remain unfolded until $\log_{10}T_G \approx -12$; consequently, they are misfolded in the $\log_{10}T_G \approx -17$ - $\log_{10}T_G \approx -12$ and $\log_{10}T_G \approx -12$ - $\log_{10}T_G \approx -17$ temperature regions. The representative structures at

six different temperatures show the unfolding/misfolding pathways of the fibrillar αS (panel D in Fig. 6), and β-strand contents (in %) as functions of temperature for three calculated structures of the fibrillar αS (Fig. S6) illustrate how β-strand content changes with the increase/decrease of temperature.

Based on the representative structures at different temperatures (panel D in Figs. 4–6), we can conclude that the unfolding/folding pathways and hence scenarios for the micelle-bound and tetrameric αS are more or less similar, but differ significantly from ones for the fibrillar αS. This is another manifestation of the flexibility of αS.

Two-dimensional (2-D) histograms of the micelle-bound, tetrameric and fibrillar αS.—To obtain more insight into the unfolding/(mis)folding kinetics of these systems, we analyzed the distribution of the conformational states in terms of 2-D histograms along the RMSD from the (mis)folded structures and the R_G as order parameters. Figures 7–9 illustrate 2-D histograms along the RMSD and R_G at six different temperatures for the micelle-bound, tetrameric and fibrillar αS, respectively. As was expected, the center of the distribution evolves with the increase of temperature in each system. The white dash line in each panel illustrates the trace of evolvement of the distribution centers of each system. Obtained results indicate on the phase transition, which can be an onset of the (mis)folding-unfolding transitions, at the temperature slightly higher than 45°C for the micelle-bound and tetrameric αS (Figs. 7 and 8), and at ~ 55°C for the fibrillar αS (Fig. 9).

Conclusions

The experimental structures of the micelle-bound, tetrameric and fibrillar αS have been analyzed in terms of kinks. Without MD simulations, kinks were able to capture the key residues playing a crucial role in structural flexibility of these three structures of αS, which makes kinks a very effective method for better understanding IDPs. Moreover, the stability of the micelle-bound, tetrameric and fibrillar αS was studied by simulating heating and cooling trajectories using the Glauber algorithm. The differences in stability of the experimental structures were shown; and the reasons for differences were explained.

Finally, it should be mentioned that kinks can provide (i) a conceptual advantage in addressing the formation of structure in protein collapse; and (ii) a simulation-free prediction of local collective motions, and can therefore be used to describe protein dynamics. This provides a new approach to the understanding of IDPs. This work is the first step in development of an innovative approach, that we are planning to do, to the elucidation of the oligomeric structures of the biomedically important IDPs from experimental fibril structures. In this approach each monomer will be modeled as a multi-kink, but interaction between different monomers will be modeled differently. A combination of electrostatic and Lennard-Jones potentials between side chains of monomers will be introduced. The side chains will be described as point-like. The evolution of the system will be monitored using Glauber dynamics by varying a temperature.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

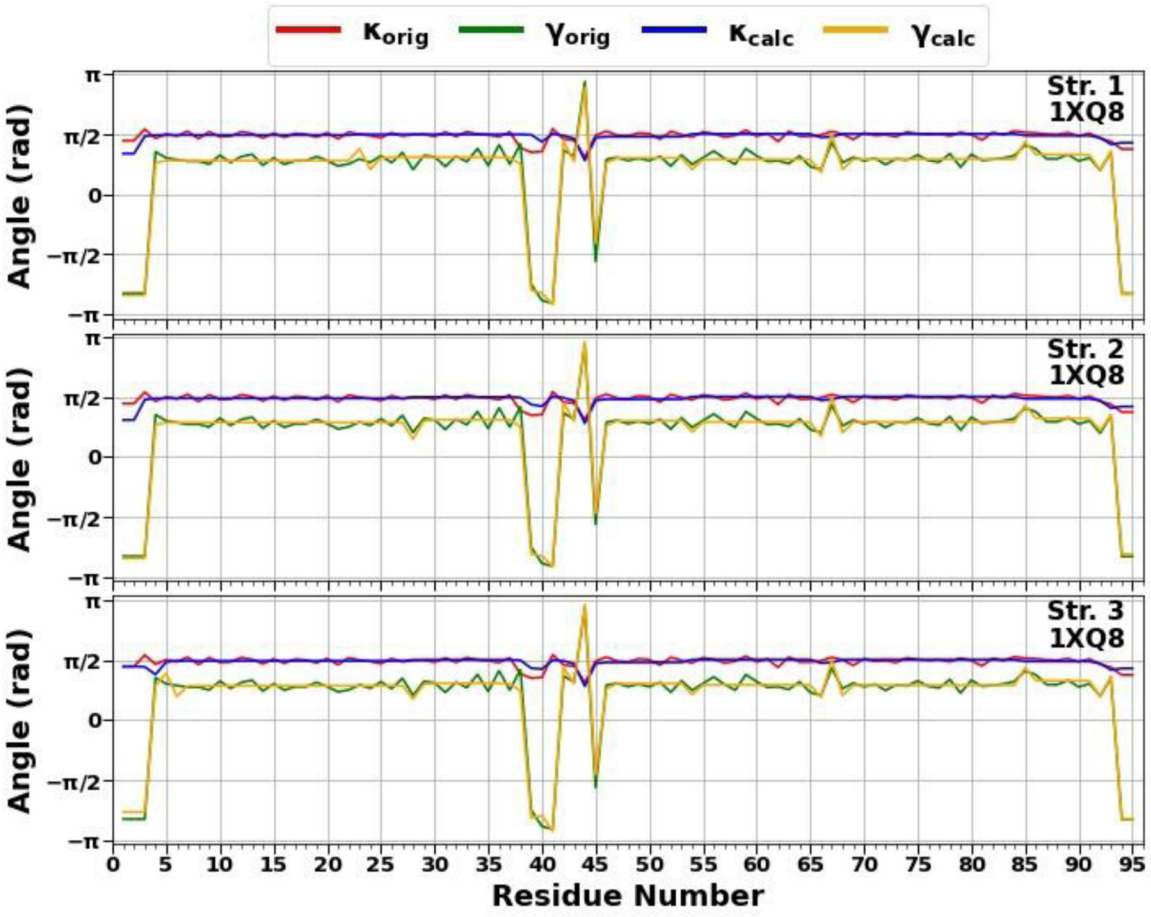
This work was (i) supported by grant from the National Institutes of Health (GM-14312) and (ii) carried out within the state assignment of the Ministry of Science and Higher Education of Russia (Grant No. 0657–2020-0015); (iii) The work by AJN has been supported by the Carl Trygger Foundation Grant CTS 18:276, by the Swedish Research Council under Contract No. 2018–04411, and by COST Action CA17139. This research was conducted by using the resources of (i) our 588-processor Beowulf cluster at the Baker Laboratory of Chemistry and Chemical Biology, Cornell University; and (ii) computing cluster Vostok-1 of Far Eastern Federal University.

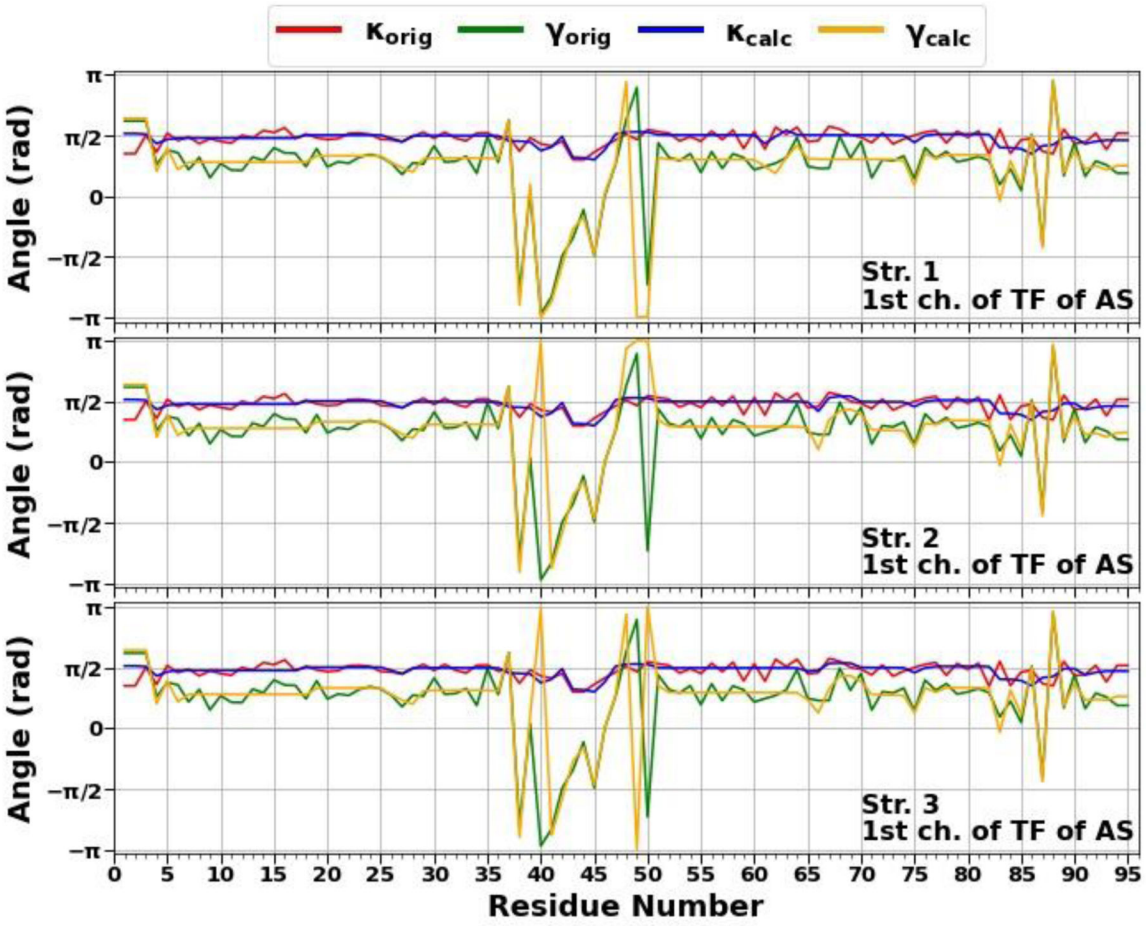
References

1. George JM The synucleins. Genome Biol 2002, 3(1), S3002.
2. Bonini NM; Giasson BI Snaring the function of alpha-synuclein. Cell 2005, 123(3), 359–361. [PubMed: 16269324]
3. Iwai A; Masliah E; Yoshimoto M; Ge N; Flanagan L; de Silva HA; Kittel A; Saitoh T The precursor protein of non-A beta component of Alzheimer's disease amyloid is a presynaptic protein of the central nervous system. Neuron 1995, 14(2), 467–475. [PubMed: 7857654]
4. Spillantini MG; Crowther RA; Jakes R; Cairns NJ; Lantos PL; Goedert M Filamentous alpha-synuclein inclusions link multiple system atrophy with Parkinson's disease and dementia with Lewy bodies. Neurosci. Lett 1998, 251(3), 205–208. [PubMed: 9726379]
5. Bertram L; Tanzi RE The genetic epidemiology of neurodegenerative disease. J. Clin. Invest 2005, 115, 1449–1457. [PubMed: 15931380]
6. Ullman O; Fisher CK; Stultz CM Explaining the structural plasticity of α-synuclein. J. Am. Chem. Soc 2011, 133, 19536–19546. [PubMed: 22029383]
7. Mor DE; Ugras SE; Daniels MJ; Ischiropoulos H Dynamic structural flexibility of α-synuclein. Neurobiol. Dis 2016, 88, 66–74. [PubMed: 26747212]
8. Ulmer TS; Bax A; Cole NB; Nussbaum RL Structure and dynamics of micelle-bound human α-synuclein. J. Biol. Chem 2005, 280, 9595–9603. [PubMed: 15615727]
9. Fauvet B; Mbefo MK; Fares MB; Desobry C; Michael S; Ardah MT; Tsika E; Coune P; Prudent M; Lion N; et al. α-Synuclein in central nervous system and from erythrocytes, mammalian cells, and escherichia coli exists predominantly as disordered monomer. J. Biol. Chem 2012, 287, 15345–15364. [PubMed: 22315227]
10. Burre J; Vivona S; Diao J; Sharma M; Brunger AT; Südhof TC Properties of native brain α-synuclein. Nature 2013, 498, E4–E6. [PubMed: 23765500]
11. Bartels T; Choi JG; Selkoe DJ α-synuclein occurs physiologically as a helically folded tetramer that resists aggregation. Nature 2011, 477, 107–110. [PubMed: 21841800]
12. Wang W; Perovic I; Chittuluru J; Kaganovich A; Nguyen LTT; Liao J; Auclair JR; Johnson D; Landeru A; Simorellis AK; et al. A Soluble α-synuclein construct forms a dynamic tetramer. Proc. Natl. Acad. Sci. U.S.A 2011, 108, 17797–17802. [PubMed: 22006323]
13. Tuttle MD; Comellas G; Nieuwkoop AJ; Covell DJ; Berthold DA; Kloepper KD; Courtney JM; Kim JK; Barclay AM; Kendall A; et al. Solid-state NMR structure of a pathogenic fibril of full-length human α-synuclein. Nat. Struct. Mol. Biol 2016, 23, 409–415. [PubMed: 27018801]
14. Li B; Ge P; Murray KA; Sheth P; Zhang M; Nair G; Sawaya MR; Shin WS; Boyer DR; Ye S; et al. Cryo-EM of full-length α-synuclein reveals fibril polymorphs with a common structural kernel. Nat. Commun 2018, 9, 3609. [PubMed: 30190461]
15. Guerrero-Ferreira R; Taylor NM; Arteni AA; Kumari P; Mona D; Ringler P; Britschgi M; Lauer ME; Makky A; Verasdonck J; et al. Two new polymorphic structures of human full-length alpha-synuclein fibrils solved by cryo-electron microscopy. eLife 2019, 8, e48907. [PubMed: 31815671]

16. Guerrero-Ferreira R; Kovacik L; Ni D; Stahlberg H New insights on the structure of alpha-synuclein fibrils using cryo-electron microscopy. Curr. Opin. Neurobiol 2020, 61, 89–95. [PubMed: 32112991]
17. Bousset L; Pieri L; Ruiz-Arlandis G; Gath J; Jensen PH; Habenstein B; Madiona K; Olieric V; Bockmann A; Meier BH; et al. Structural and functional characterization of two alpha-synuclein strains. Nat. Commun 2013, 4, 2575. [PubMed: 24108358]
18. Gurry T; Ullman O; Fisher CK; Perovic I; Pochapsky T; Stultz CM The dynamic structure of α-synuclein multimers. J. Am. Chem. Soc 2013, 135, 3865–3872. [PubMed: 23398399]
19. Cote Y; Delarue P; Scheraga HA; Senet P; Maisuradze GG From highly-disordered to metastable state: Uncovering insights of α-synuclein. ACS Chem. Neurosci 2018, 9, 1051–1065. [PubMed: 29451381]
20. Brodie NI; Popov KI; Petrotchenko EV; Dokholyan NV; Borchers CH Conformational ensemble of native α-synuclein in solution as determined by short-distance crosslinking constraint-guided discrete molecular dynamics simulations. PLoS Comput. Biol 2019, 15(3), e1006859. [PubMed: 30917118]
21. Guzzo A; Delarue P; Rojas A; Nicolaï A; Maisuradze GG; Senet P, Missense P mutations modify the conformational ensemble of the α-synuclein monomer which exhibits a two-phase characteristic. Front. Mol. Biosci 2021, 8, 786123. [PubMed: 34912851]
22. Chernodub M; Hu S; Niemi AJ Topological solitons and folded proteins. Phys. Rev. E 2010, 82, 011916.
23. Molkenthin N; Hu S; Niemi AJ Discrete nonlinear Schrödinger equation and polygonal solitons with applications to collapsed proteins. Phys. Rev. Lett 2011, 106, 078102. [PubMed: 21405545]
24. Hu S; Krokhotin A; Niemi AJ; Peng X Towards quantitative classification of folded proteins in terms of elementary functions. Phys. Rev. E 2011, 83, 041907.
25. Hu S; Lundgren M; Niemi AJ Discrete Frenet frame, inflection point solitons, and curve visualization with applications to folded proteins. Phys. Rev. E 2011, 83, 061908.
26. Krokhotin A; Niemi AJ; Peng X Soliton concepts and protein structure. Phys. Rev. E 2012, 85, 031906.
27. Krokhotin A; Liwo A; Niemi AJ; Scheraga HA Coexistence of phases in a protein heterodimer. J. Chem. Phys 2012, 137, 035101. [PubMed: 22830730]
28. Lundgren M; Krokhotin A; Niemi AJ Topology and structural self-organization in folded proteins. Phys. Rev. E 2013, 88, 042709.
29. Krokhotin A; Liwo A; Maisuradze GG; Niemi AJ; Scheraga HA Kinks, loops, and protein folding with protein A as an example. J. Chem. Phys 2014, 140, 025101. [PubMed: 24437917]
30. Niemi A; Liu J; Dai J; He J; Ilieva N Towards multistage algorithm to model intrinsically unstructured proteins. APS March Meeting Abstracts 2017, M1, 248.
31. Kachlishvili K; Korneev A; Maisuradze L; Liu J; Scheraga HA; Molochkov A; Senet P; Niemi AJ; Maisuradze GG New insights into folding, misfolding and nonfolding dynamics of a WW domain. J. Phys. Chem. B 2020, 124, 3855–3872. [PubMed: 32271570]
32. Sieradzan AK; Korneev A; Begun A; Kachlishvili K; Scheraga HA; Molochkov A; Senet P; Niemi AJ Maisuradze GG Investigation of phosphorylation-induced folding of an intrinsically disordered protein by coarse-grained molecular dynamics. J. Chem. Theory & Comput 2021, 17, 3203–3220. [PubMed: 33909430]
33. Faddeev LD; Takhtajan L Hamiltonian methods in the theory of solitons; Springer-Verlag: Berlin, Germany, 1987.
34. Ablowitz MJ; Prinardi B; Trubatch A Discrete and continuous nonlinear Schrödinger systems; Cambridge University Press: Cambridge, UK, 2004.
35. Kevrekidis P The Discrete nonlinear Schrödinger equation: mathematical analysis, numerical computations and physical perspectives; Springer-Verlag: Berlin, Germany, 2009.
36. Manton N; Sutcliffe P Topological solitons; Cambridge University Press: Cambridge, UK, 2004.
37. Weinberg S The quantum theory of fields; Cambridge University Press: Cambridge, UK, 1995; Vol. 2.

38. Davydov AS Solitons and energy transfer along protein molecules. J. Theor. Biol 1977, 66, 379–387. [PubMed: 886872]
39. Anfinsen CB; Scheraga HA Experimental and theoretical aspects of protein folding. Adv. Protein Chem 1975, 29, 205–300. [PubMed: 237413]
40. Widom B Surface tension and molecular correlations near the critical point. J. Chem. Phys 1965, 43, 3892–3897.
41. Kadanoff LP Scaling laws for Ising models near T_c. Physics 1966, 2, 263–272.
42. Wilson K Renormalization group and critical phenomena. I. Renormalization group and the Kadanoff scaling picture. Phys. Rev. B 1971, 4, 3174–3183.
43. Fisher ME The renormalization group in the theory of critical behavior. Rev. Mod. Phys 1974, 46, 597–616.
44. Molochkov A; Begun A; Niemi AJ Gauge symmetries and structure of proteins. EPJ Web Conf 2017, 137, 04004.
45. Niemi AJ WHAT IS LIFE – Sub-cellular physics of live matter. In Topological Aspects of Condensed Matter Physics, Les Houches Summer School 103; Oxford University Press, 2017.
46. Begun A; Molochkov A; Niemi AJ Protein tertiary structure and the myoglobin phase diagram. Sci. Rep 2019, 9, 10819. [PubMed: 31346242]
47. Krokhotin A; Lundgren M; Niemi AJ; Peng X Soliton driven relaxation dynamics and protein collapse in the villin headpiece. J. Phys.: Condens. Matter 2013, 25, 325103. [PubMed: 23838560]
48. Peng X; Sieradzan AK; Niemi AJ Thermal unfolding of myoglobin in the Landau-Ginzburg-Wilson approach. Phys. Rev. E 2016, 94, 062405. [PubMed: 28085346]
49. Cascella R; Chen SW; Bigi A; Camino JD; Xu CK; Dobson CM; Chiti F; Cremades N; Cecchi C The release of toxic oligomers from α-synuclein fibrils induces dysfunction in neuronal cells. Nat. Commun 2021, 12, 1814. [PubMed: 33753734]
50. Dettmer U; Newman AJ; von Saucken VE; Bartels T; Selkoe D KTKEGV repeat motifs are key mediators of normal α-synuclein tetramerization: their mutation causes excess monomers and neurotoxicity. Proc. Natl. Acad. Sci. U. S. A 2015, 112, 9596–9601. [PubMed: 26153422]
51. Dettmer U; Newman AJ; Soldner F; Luth ES; Kim NC; von Saucken VE; Sanderson JB; Jaenisch R; Bartels T; Selkoe D Parkinson-causing α-synuclein missence mutations shift native tetramers to monomers as a mechanism for disease initiation. Nat. Commun 2015, 6, 7314. [PubMed: 26076669]
52. Guzzo A;, Delarue P; Rojas A; Nicolaï A; Maisuradze GG; P. Senet, P. Wild-type α-synuclein and variants occur in different disordered dimers and pre-fibrillar conformations in early stage of aggregation. Front. Mol. Biosci 2022, 9, 910104. [PubMed: 35836937]
53. Khalaf O; Fauvet B; Oueslati A; Dikiy I; Mahul-Mellier A-L; Ruggeri FS; Mbefo MK; Vercruysse F; Dietler G; Lee S-J; et al. The H50Q mutation enhances α-synuclein aggregation, secretion, and Toxicity. J. Biol. Chem 2014, 289, 21856–21876. [PubMed: 24936070]
54. Scalley ML; Baker D Protein folding kinetics exhibit an Arrhenius temperature dependence when corrected for the temperature dependence of protein stability. Proc. Natl. Acad. Sci. U. S. A 1997, 94, 10636–10640. [PubMed: 9380687]





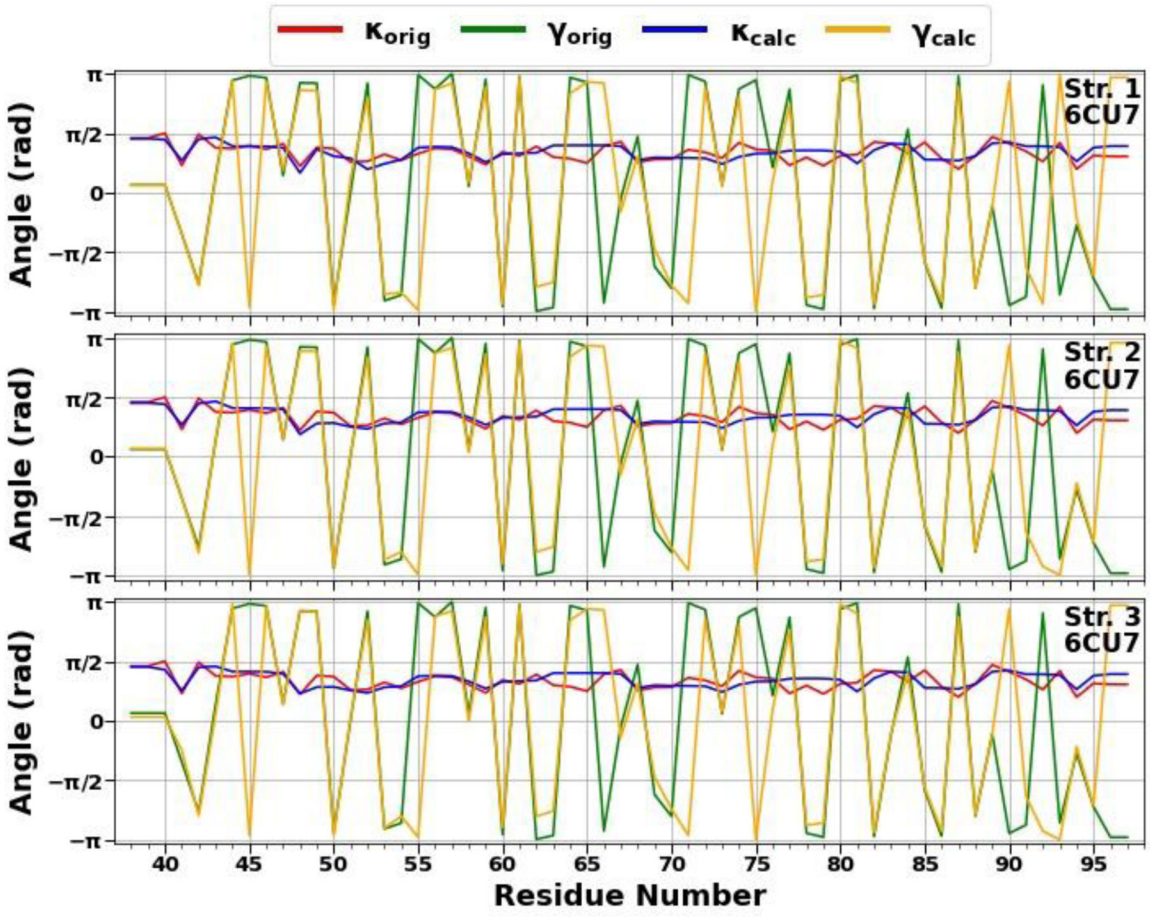
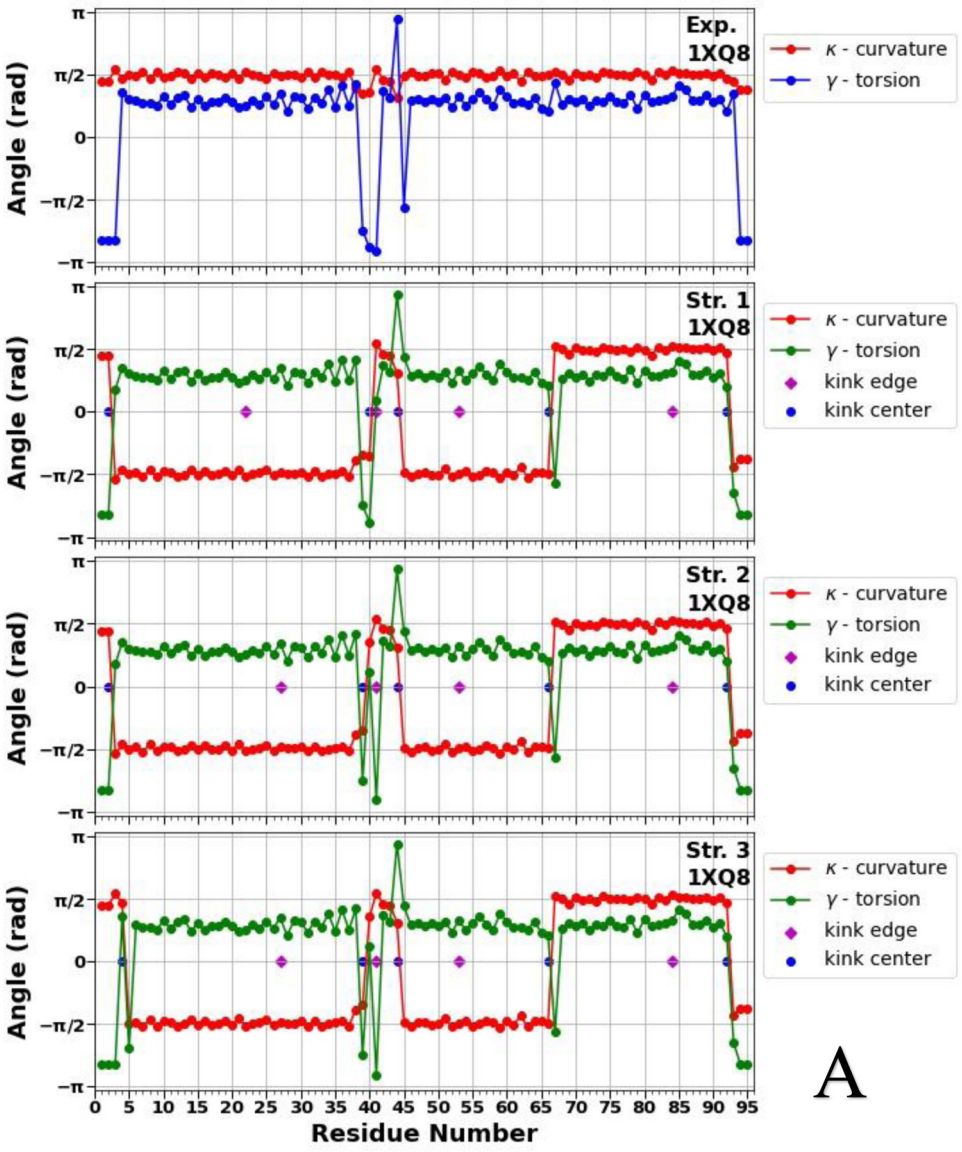
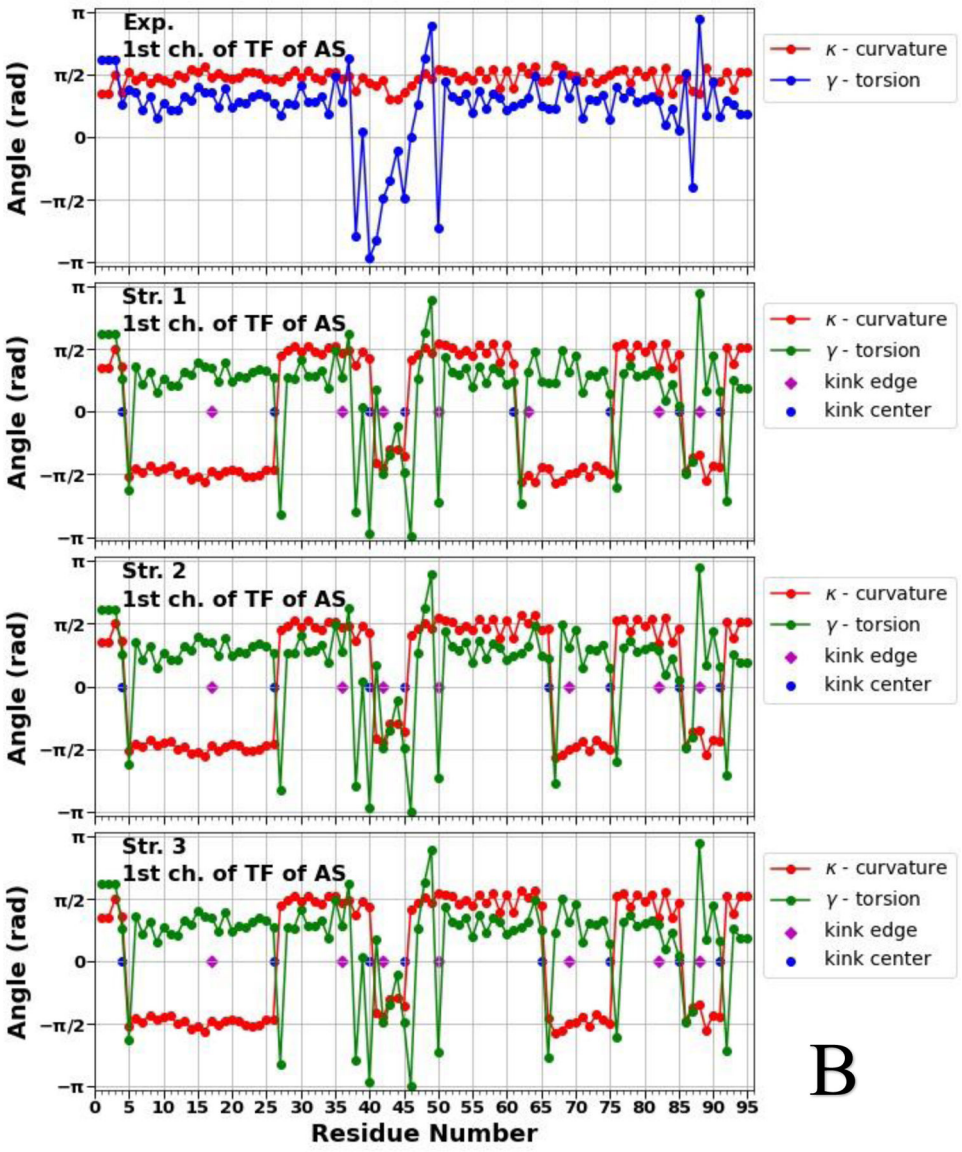


Figure 1.

The comparison of experimental and calculated angle spectra of three selected structures for the micelle-bound (top 3 panels), tetrameric (middle 3 panels), and fibrillar (bottom 3 panels) αS in terms of virtual-bond κ_i (experimental-red, calculated-blue) and torsion γ_i (experimental-green, calculated-yellow) values.



A



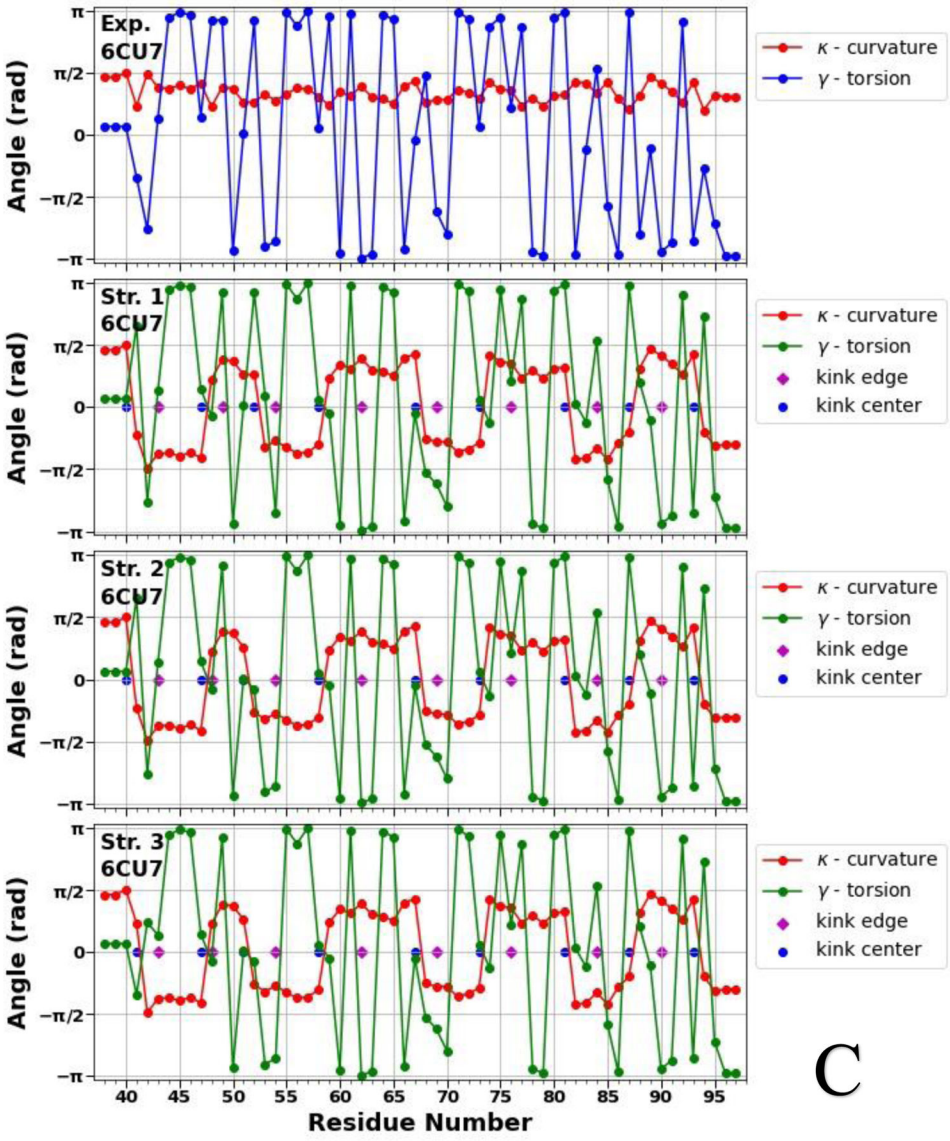


Figure 2.

Top panels of A, B and C represent experimental values of the virtual-bond κ_i (red) and torsion γ_i (blue) angle spectra of the micelle-bound (panel A), tetrameric (panel B) and fibrillar (panel C) αS. Remaining three panels in A, B and C illustrate kink structures, i.e., the virtual-bond κ_i (red) and torsion γ_i (green) angle spectra for three selected structures of the micelle-bound (panel A), tetrameric (panel B) and fibrillar (panel C) αS, after $\mathbb{Z}_2$ gauge transformation [Eq. (S8)] was made. Blue circles represent centers of kinks, purple rhombuses represent the right edges of left kinks.

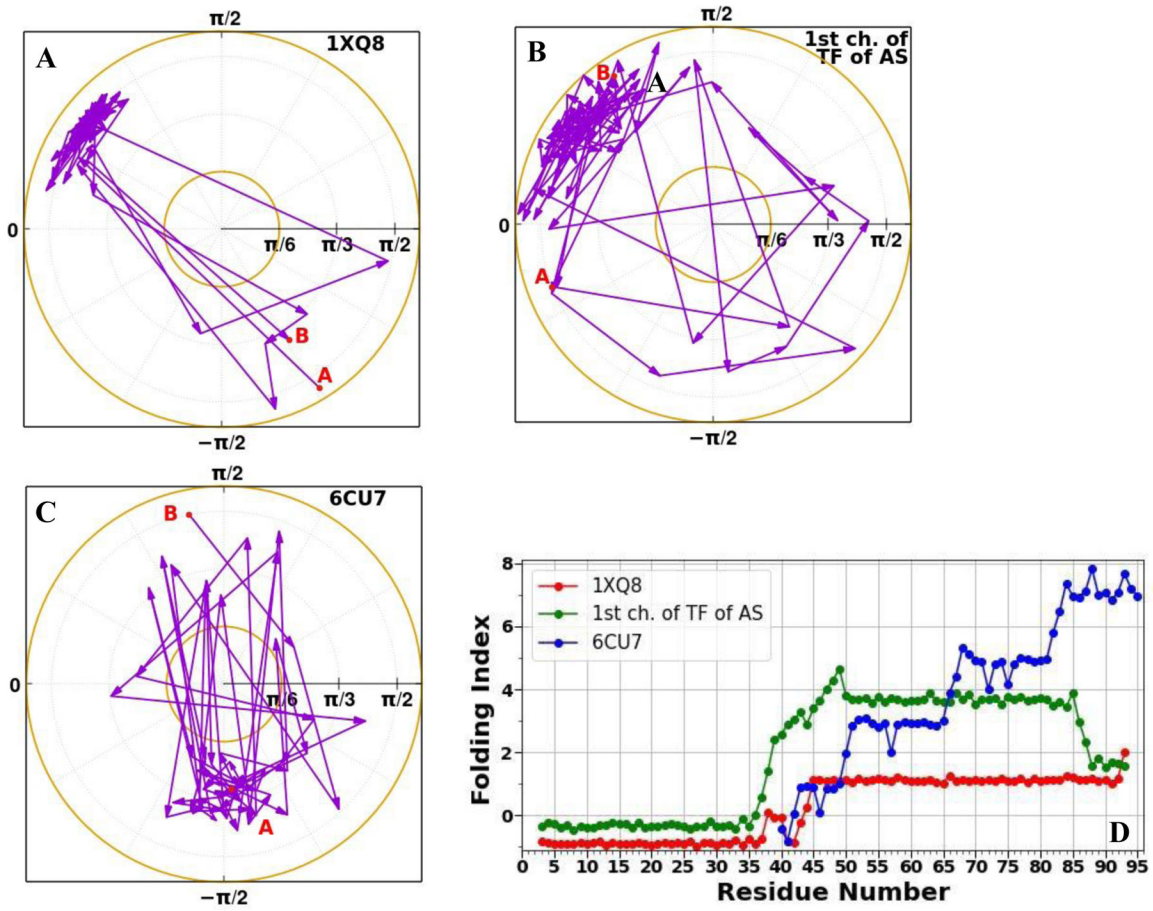
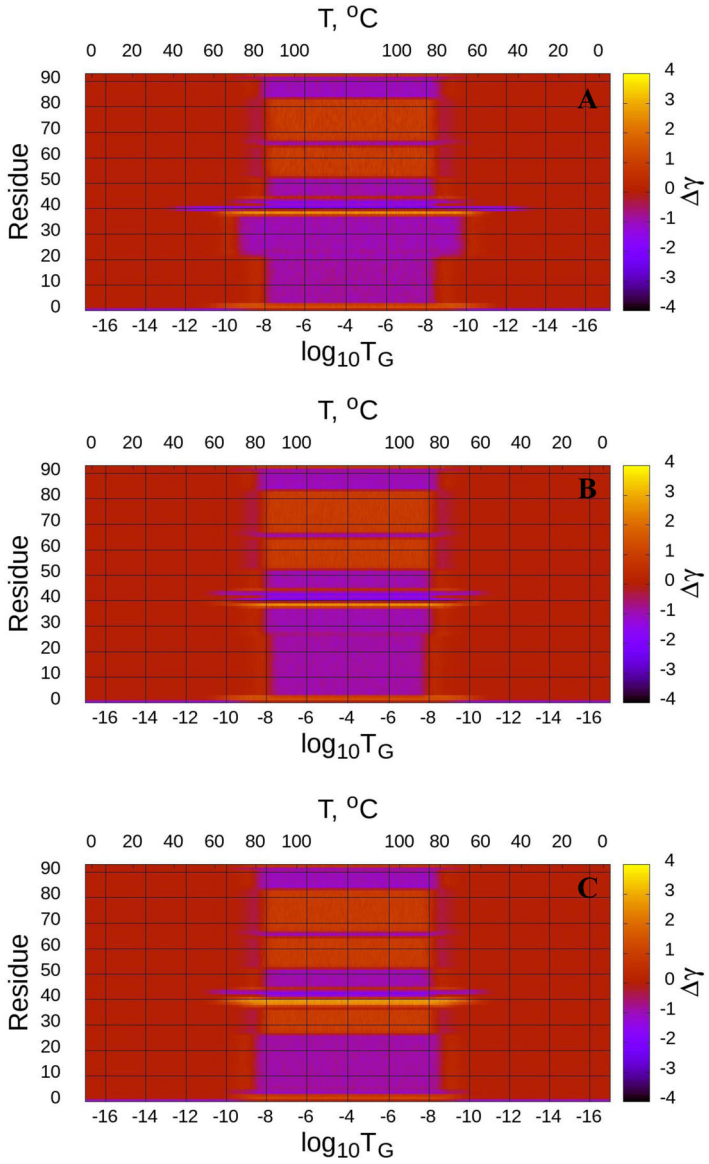


Figure 3.

Folding index trajectories for the micelle-bound (A), tetrameric (B), and fibrillar (C) αS, and the folding index (D) for the micelle-bound (red), tetrameric (green), and fibrillar (blue) αS.



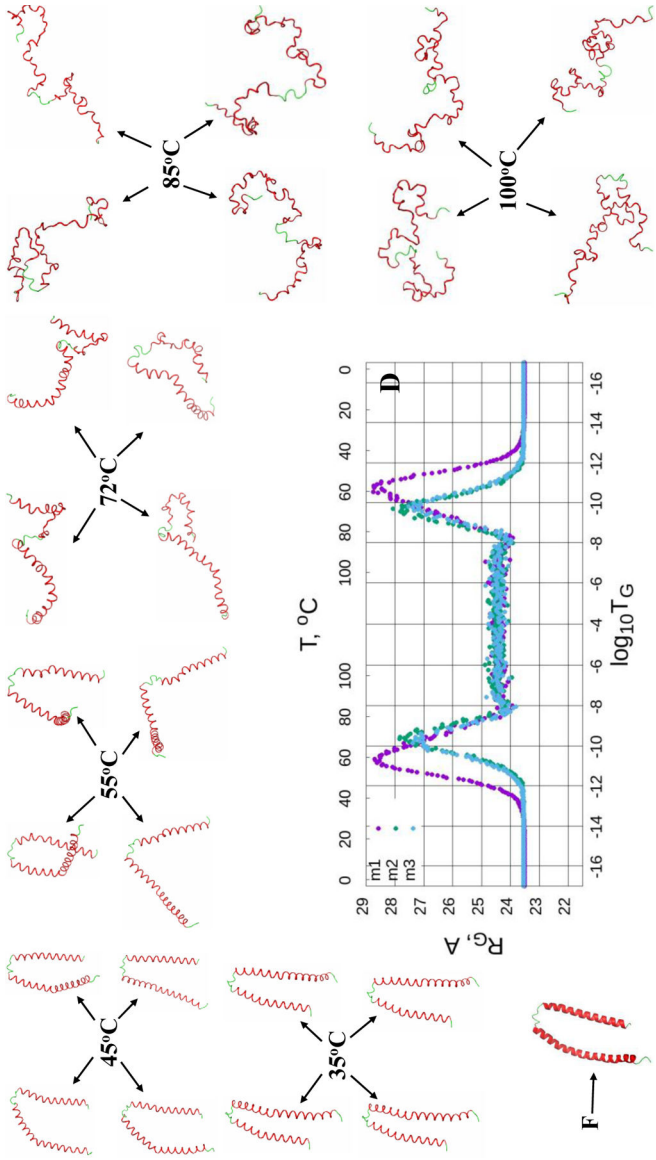
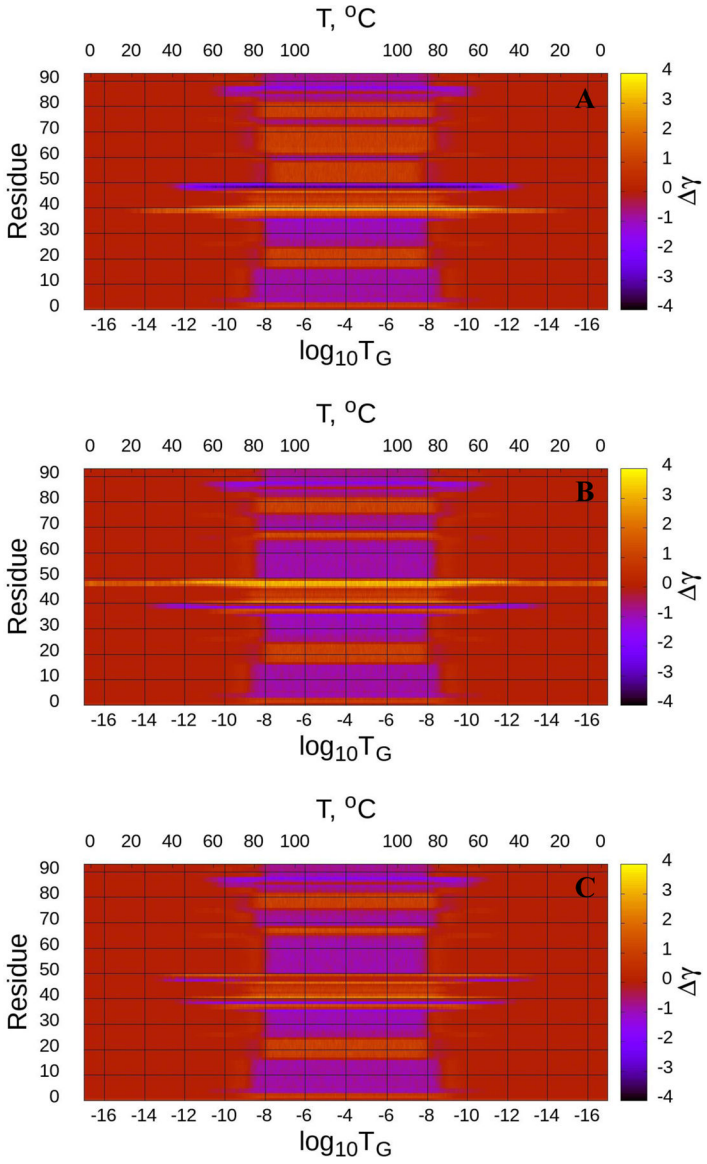


Figure 4.

(Dis)ordering during a heating and cooling simulation cycle as a function of Glauber temperature (T_G) in terms of the average value of torsion angles γ for three selected structures of the micelle-bound αS (A-C), and the radius of gyrations (R_G) of three selected structures as functions of Glauber temperature (D). The representative structures at six different temperatures along with the folded (F) structure show the unfolding/folding pathways of the system (D).



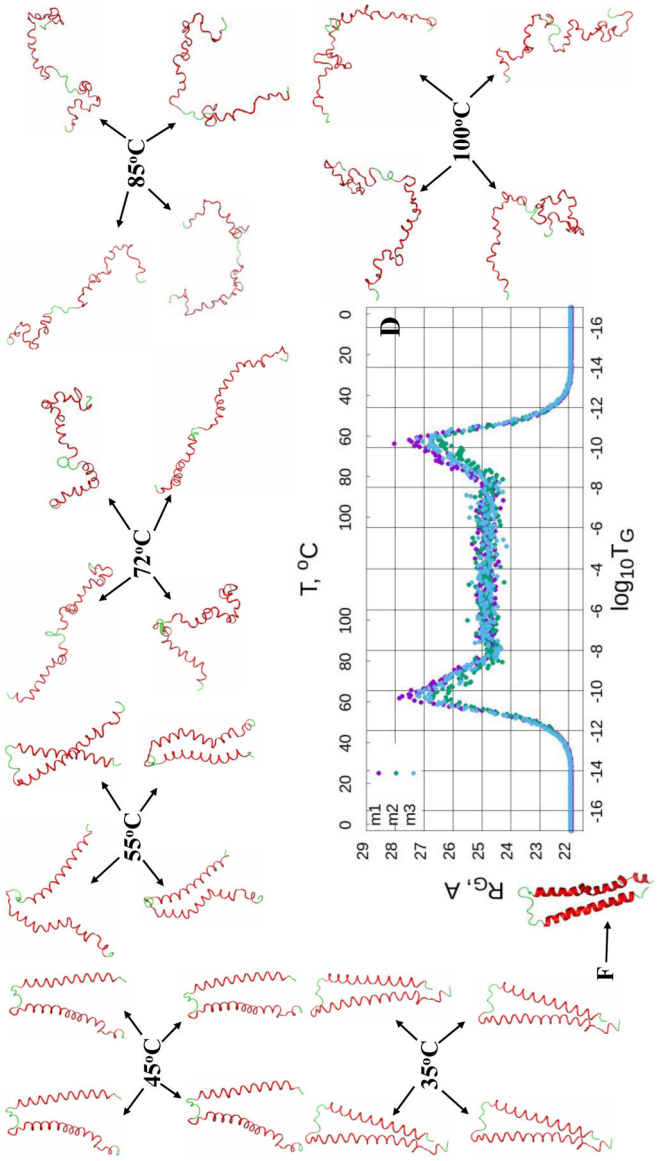
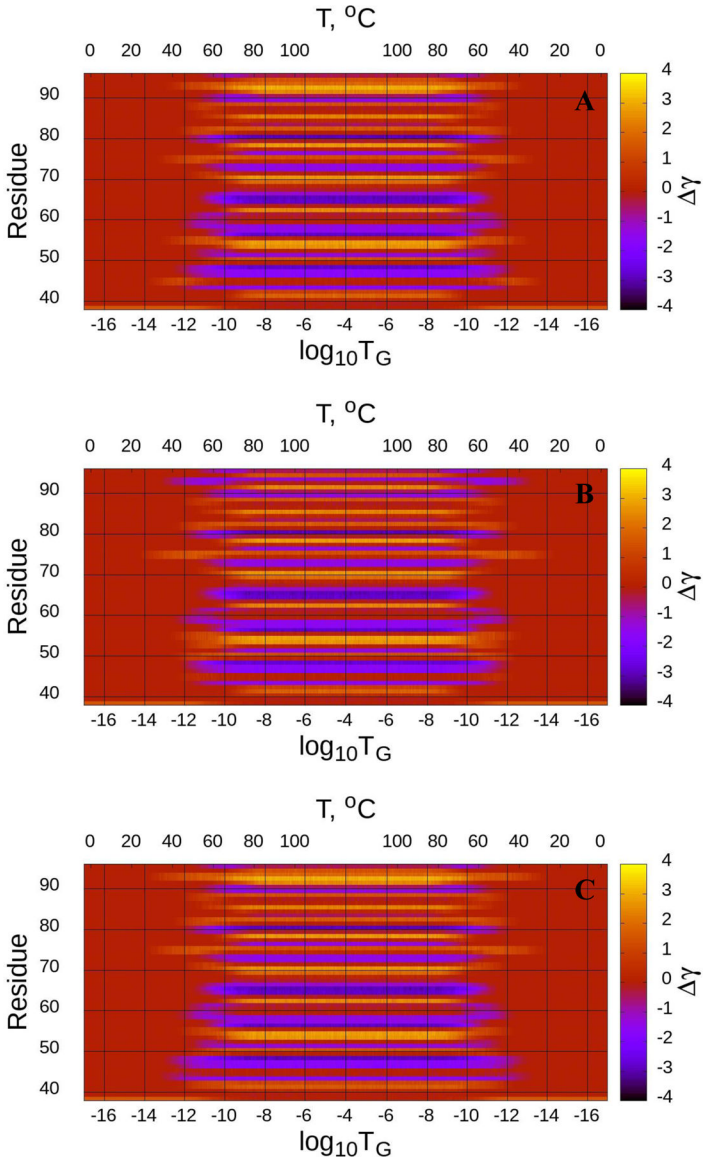


Figure 5.

(Dis)ordering during a heating and cooling simulation cycle as a function of Glauber temperature (T_G) in terms of the average value of torsion angles γ for three selected structures of the tetrameric αS (A-C), and the radius of gyrations (R_G) of three selected structures as functions of Glauber temperature (D). The representative structures at six different temperatures along with the folded (F) structure show the unfolding/folding pathways of the system (D).



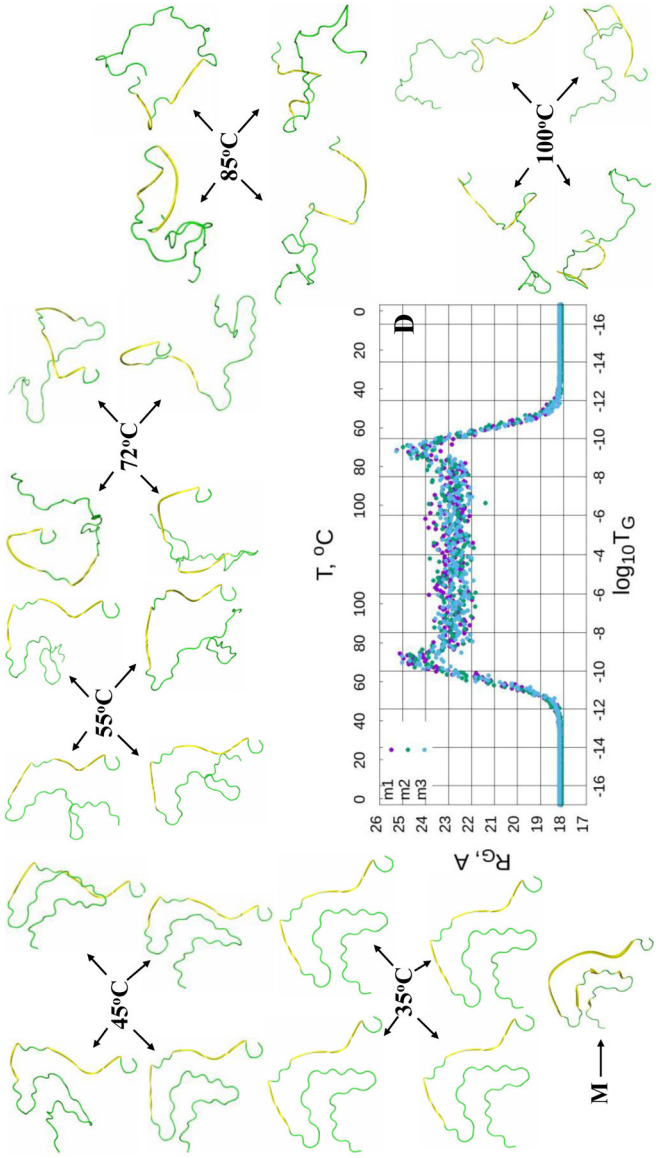


Figure 6.

(Dis)ordering during a heating and cooling simulation cycle as a function of Glauber temperature (T_G) in terms of the average value of torsion angles γ for three selected structures of the fibrillar αS (A-C), and the radius of gyrations (R_G) of three selected structures as functions of Glauber temperature (D). The representative structures at six different temperatures along with the misfolded (M) structure show the unfolding/folding pathways of the system (D).

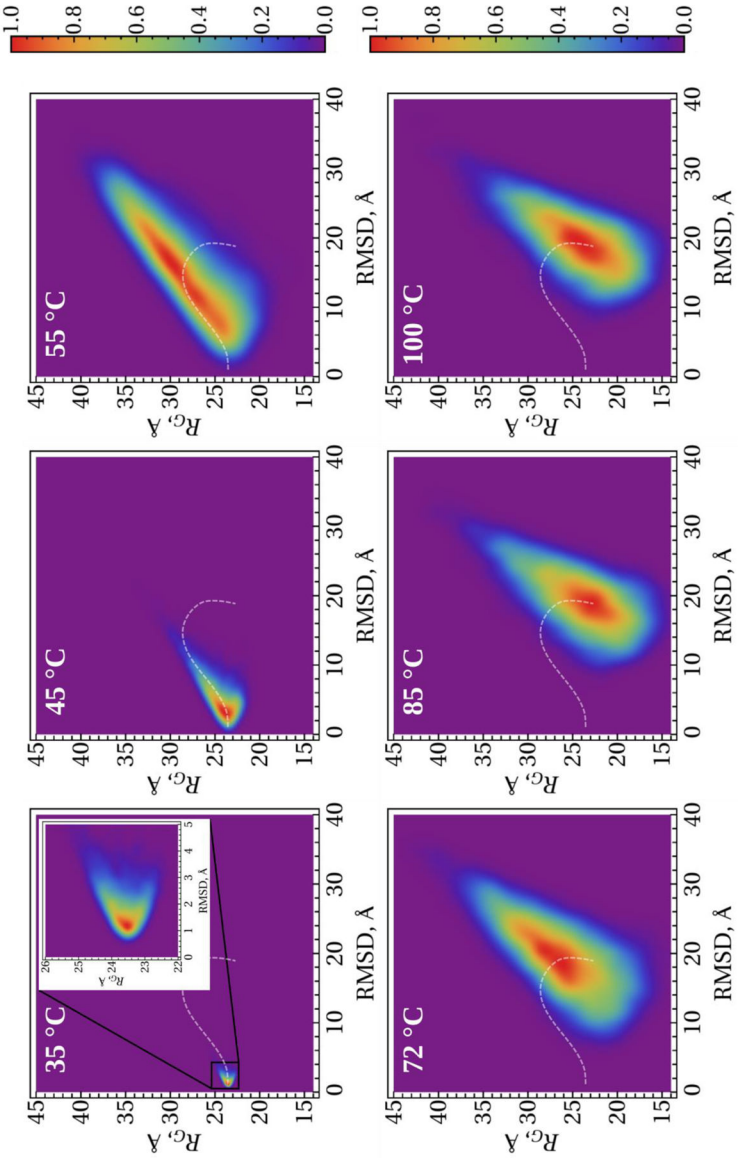


Figure 7.

2-D histograms along the RMSD and R_G order parameters for the micelle-bound αS. The white dash lines show the trace of evolvement of the distribution centers with the temperature.

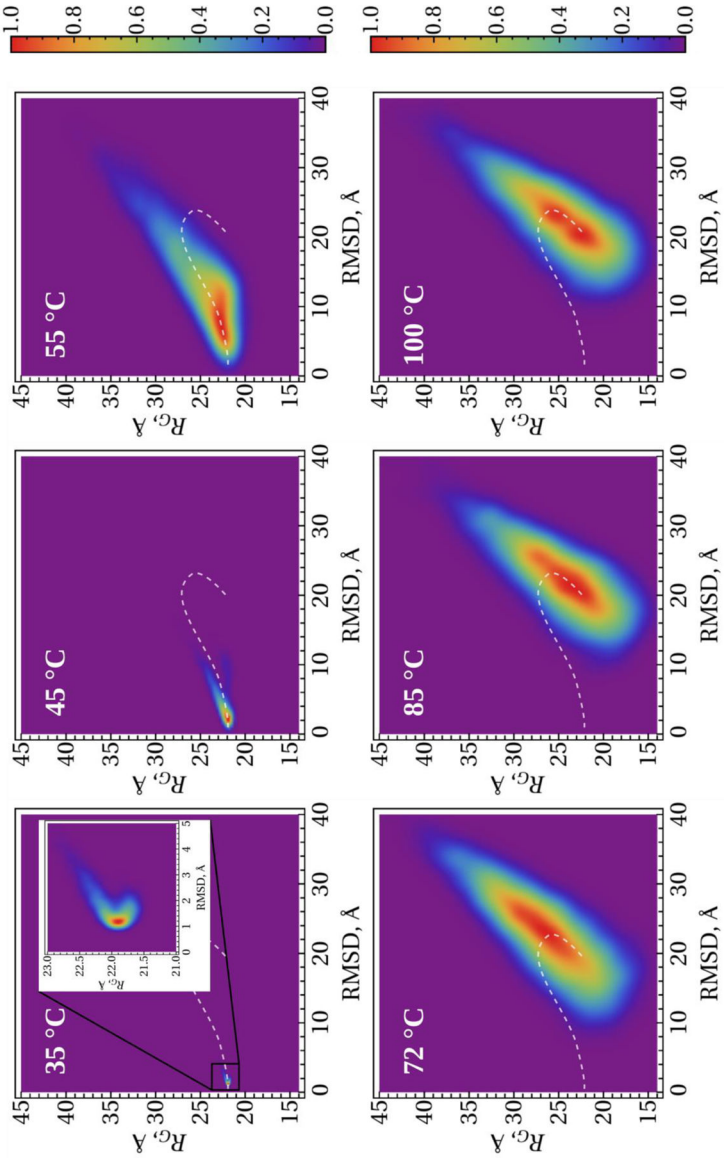


Figure 8.

2-D histograms along the RMSD and R_G order parameters for the tetrameric αS. The white dash lines show the trace of evolvement of the distribution centers with the temperature.

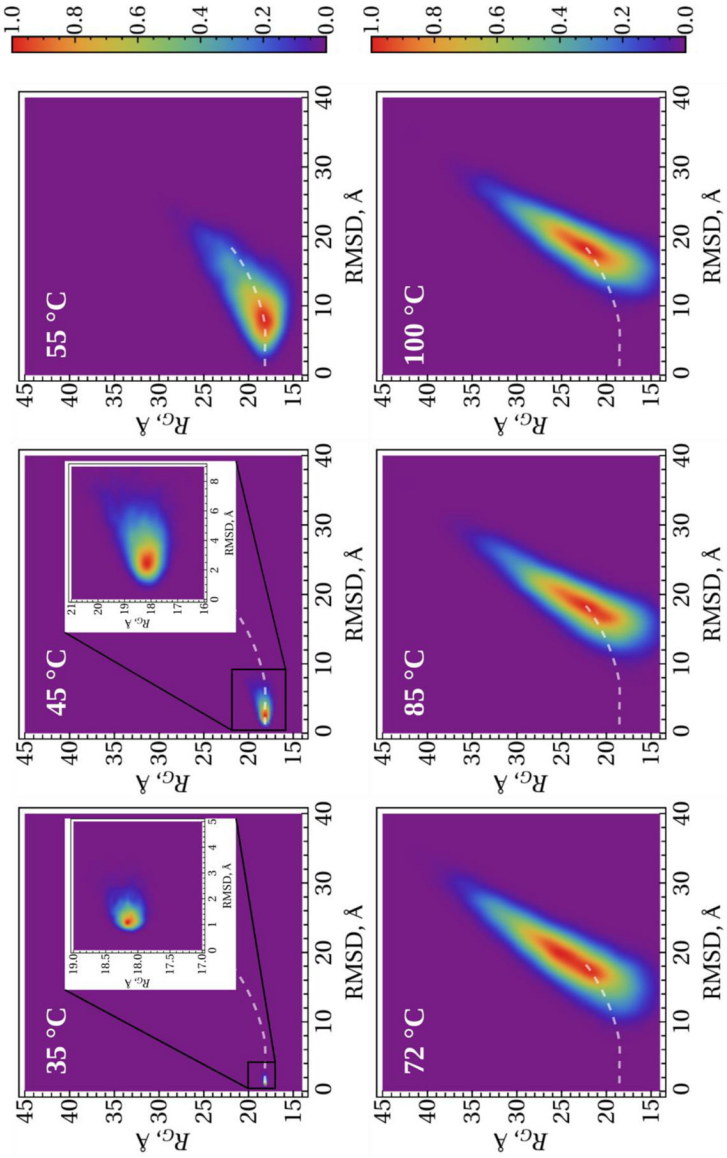


Figure 9.

2-D histograms along the RMSD and R_G order parameters for the fibrillar αS. The white dash lines show the trace of evolvement of the distribution centers with the temperature.