

# Effect of an amyloidogenic SARS-COV-2 protein fragment on $\alpha$ -synuclein monomers and fibrils

*Asis K. Jana\**, *Chance W. Lander*, *Andrew D. Chesney* and *Ulrich H. E. Hansmann\**

Department of Chemistry & Biochemistry, University of Oklahoma, Norman, OK 73019, USA.

**KEYWORDS:** Molecular Dynamics Simulations, SARS-COV-2,  $\alpha$ -synuclein

## ABSTRACT

Using molecular dynamic simulations we study whether amyloidogenic regions in viral proteins can initiate and modulate formation of  $\alpha$ -synuclein aggregates, thought to be the disease-causing agent in Parkinson's Disease. As an example we choose the nine-residue fragment SFYVYSRVK (SK9), located on the C-terminal of the Envelope protein of SARS-COV-2. We probe how the presence of SK9 affects the conformational ensemble of  $\alpha$ -synuclein monomers and the stability of two resolved fibril polymorphs. We find that the viral protein fragment SK9 may alter  $\alpha$ -synuclein amyloid formation by shifting the ensemble toward aggregation-prone and preferentially rod-like fibril seeding conformations. However, SK9 has only little effect of the stability of pre-existing or newly-formed fibrils.

## Introduction

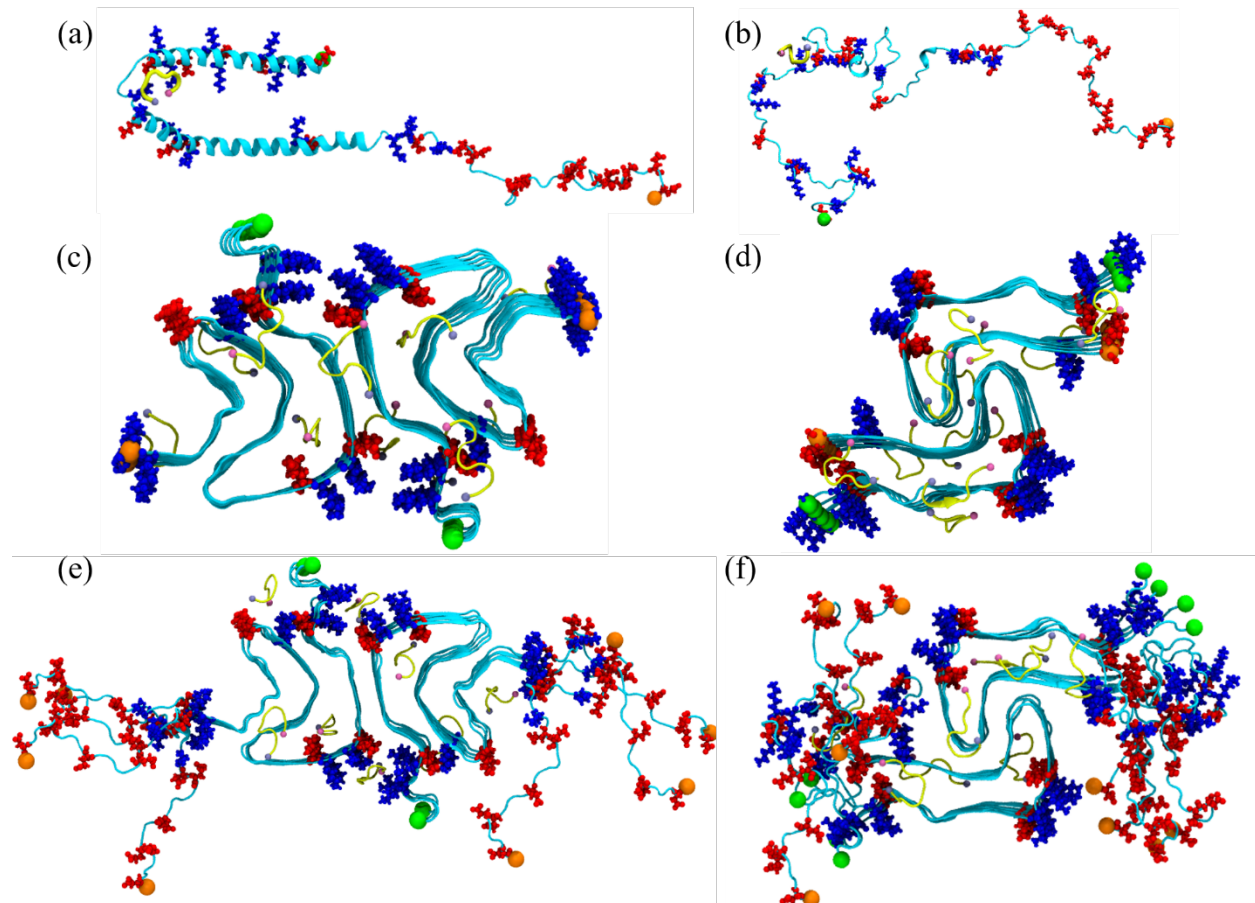
While most COVID-19 survivors seem to recover fully, not much is known about potential long-term or delayed neurological effects of SARS-COV-2 infections. Acute infection-associated symptoms include loss of smell and other neurological impairments<sup>1</sup>; and a number of case studies suggests a role of SARS-CoV-2 in the onset of Parkinson's Disease and related neurodegenerative diseases<sup>2-5</sup>. A possible mechanism for triggering Parkinson's Disease could be virus-initiated amyloid-formation as aggregates of  $\alpha$ -synuclein (aS) are the cell-toxic agent in Parkinson's Disease and related pathologies<sup>6</sup>. SARS-COV-2 accelerated aS amyloid formation, caused by interaction with amyloidogenic regions on the Spike, Envelope or Nucleocapsid protein, has been observed *in vitro*<sup>4</sup>, and may not be accidental. It has been speculated in the context of Alzheimer's Disease that amyloid fibrils form as an immune response to an infection, entrapping and neutralizing pathogens<sup>7-9</sup>; and a similar mechanism may also play a role in Parkinson's Disease. However, the link between exposure to pathogens, appearance of fibrils, and subsequently of disease symptoms is, even in the case of Alzheimer's Disease, only indirect. In order to evaluate whether such a microbial protection hypothesis<sup>7,8</sup> can explain the correlations between COVID-19 and outbreak of Parkinson's Disease, one needs to understand the mechanism by which amyloidogenic regions on SARS-COV-2 proteins initiate and modulate aggregation of  $\alpha$ -synuclein.

In the present paper we focus on the nine-residue fragment SFYVYSRVK (SK9) on the C-terminal of the Envelope protein. We are familiar with SK9 from other work,<sup>10,11</sup> and study now its effect on  $\alpha$ -synuclein amyloid formation. The 140-residue long  $\alpha$ -synuclein monomer consists of three domains: (i) an amphipathic N-terminal region (residues 1-60), where most disease-

causing familial mutations are located and which binds to membranes; (ii) a hydrophobic non-amyloid- $\beta$  component (NAC) region (residues 61-95) that is thought to be the core region for aggregation; and (iii) a disordered negatively charged C-terminal region (residues 96-140).<sup>12-15</sup> A helix-rich model of the  $\alpha$ -synuclein monomer structure has been resolved by solution NMR in the micellar environment and is deposited in the Protein Data Bank (PDB) under PDB ID: 1XQ8.<sup>12</sup> Using molecular dynamics simulations we study how SK9 interacts with  $\alpha$ -synuclein and whether SK9 changes the conformational ensemble of the monomer, potentially modifying oligomerization and hastening the formation of a critical nucleus for seeding fibrils.

In a complementary set of simulations we explore whether and how SK9 differentially alters the stability of experimentally resolved  $\alpha$ -synuclein fibril models, comparing the two predominant forms, the “rod” or “twister” polymorphs.<sup>16</sup> While in both structures the individual chains share a bent  $\beta$ -arch architecture, they have different inter-protofilament interfaces. In the twister polymorph is the interface formed by the hydrophobic aggregation-inducing non-amyloid- $\beta$  component (NAC) region (G68-A78), while in the rod polymorph the interface is made of the preNAC region (E46-A56).<sup>16,17</sup> The residues at the C-terminus in the rod polymorph are more ordered, suggesting higher stability than found for the twister polymorph.<sup>16</sup> Note, however, that the six common mutations E46K, H50Q, G51D, A53E, A53T, and A53V, destabilize the preNAC interface in the rod structure while not disrupting the twister polymorph, i.e., likely shifting the population from rod to twister.<sup>16,18</sup> As these mutations are associated with hereditary forms of Parkinson’s Disease, it appears as if a shift in frequency between rod and twister fibril forms will alter the probability for acquiring Parkinson’s Disease.<sup>16,18</sup>

Our goal in the present paper is therefore to study if and how the viral segment SK9 accelerates aggregation of  $\alpha$ -synuclein, and whether this process is selective, preferentially enhancing the formation of potentially more toxic amyloid species.



**Figure 1:** Initial conformation of the  $\alpha$ -synuclein monomer (a) as resolved by solution NMR (PDB ID: 1XQ8), and (b) after heating at 500 K to obtain a randomized stretched conformation. Initial conformation for the fibril as derived by cryo-EM structures are shown in (c) for the rod (PDB ID: 6CU7) and in (d) for the twister (PDB ID: 6CU8) polymorph. In (e) and (f) are the corresponding structures shown for the fibrils where the individual chains are extended to residues 38-120. Acidic residues are colored in red and basic ones in blue, while the SK9-segments are shown yellow. The N- and C-termini are represented by green and orange spheres, respectively.

## Materials and Methods

## *System Preparation*

In order to understand the effect of SARS-COV-2 proteins on  $\alpha$ -synuclein aggregation we use molecular dynamics simulations to study how the presence of the nine-residue segment S<sup>55</sup>FYVYSRVK<sup>63</sup> (SK9) of the Envelope protein changes the conformational ensemble and stability of  $\alpha$ -synuclein monomers and fibrils. For the initial configuration for the monomer, we chose the model resolved by solution NMR and deposited in the Protein Data Bank (PDB) under PDB ID: 1XQ8. As this structure is characterized by a largely helical N-terminus,<sup>12</sup> and to avoid any bias by this choice of an initial configuration, we have also considered a fully extended  $\alpha$ -synuclein monomer as an alternative start configuration. It is obtained by heating the experimentally resolved conformation up to 500 K, and simulating it for 5 ns. On the other hand, for our investigations into the stability of  $\alpha$ -synuclein fibrils, we have considered the two predominant forms, the rod and a twister polymorphs. For this purpose, we have generated decamers, made of two folds (protofilaments) and five layers, from the cryo-EM structures deposited in the PDB under identifier 6CU7 and 6CU8, respectively.<sup>16</sup> Note that for the rod model only 60 residues (L38-K97) out of the 140 amino acids in  $\alpha$ -synuclein have been resolved, and only 41 residues (K43-E83) for the twister model. For this reason, we have also generated models for the rod and the twister fibril where all of the experimentally unsolved regions are added as predicted by homology modeling. However, since subsequent docking experiments did not indicate binding of the SK9 segment to the N-terminal residues 1-37 or the C-terminal residues 121-140, we decided to reduce computational costs by only considering extended rod and twister fibril models made of residues 38-120. For both monomer and fibril models, the N- and C-terminus of the chains are capped by NH<sub>3</sub><sup>+</sup> and COO<sup>-</sup> groups, respectively.

Simulations starting from the two monomer and four fibril models serve as controls against simulations in which SK9 is present. The initial configuration for the SK9 fragment was derived by removing the residues 1-54 and 64-75 from the model of the SARS-COV-2 E-protein as deposited in the MOLSSI COVID-19 hub<sup>19</sup>, with the N- and C-termini of the peptide capped by a NH<sub>3</sub><sup>+</sup> and -CONH<sub>2</sub> group, respectively. Using the AutoDock Vina software<sup>20</sup>, we have generated start configurations for our simulations by docking SK9 segments in a ratio of 1:1 with the  $\alpha$ -synuclein chains in the monomer and fibril systems described above. The so-generated six structures are shown in **Figure 1**, but note that SK9 segment is not constrained to these positions and can move freely (and even detach) in the simulations.

The twelve models of  $\alpha$ -synuclein monomers and fibrils, either with SK9 binding to the  $\alpha$ -synuclein chain or with SK9 absent, are each put in a rectangular box with periodic boundary conditions. The box is chosen such that there is a minimum distance of 15 Å between any SAA or SK9 atom and all box sides. This large box size ensures that the  $\alpha$ -synuclein chains do not interact with their periodic images. Each box is filled with water molecules, and counterions are added to neutralize the system. Additional Na<sup>+</sup> and Cl<sup>-</sup> ions were added to arrive at a physiological ion concentration of 0.15 M NaCl. **Table 1** lists for all systems the number of water molecules and the total number of atoms.

### ***General Simulation Protocol***

All simulations were carried out using the GROMACS 2018 simulation package.<sup>21</sup> Intramolecular interactions are described for the fibril by the CHARMM 36m all-atom force field<sup>22</sup> and the TIP3P water model<sup>23</sup>, which we found in earlier work well-suited for simulations of fibril and

oligomers.<sup>24,25</sup> However, the monomer simulations rely on the a99SB-disp all-atom force field and the corresponding a99SB-disp water model. This is because CHARMM and AMBER force fields have been shown to lead to conformational ensembles of  $\alpha$ -synuclein chains that are too compact<sup>26</sup>, while the a99SB-disp all-atom force field and a99SB-disp water model lead to ensembles that do agree with the experimental measurements<sup>27</sup>. Prior to the production run, each system is first energy minimized for up to 5 000 000 steps using steepest descent, and equilibrated by 200 ps molecular dynamics in a NVT ensemble of 310 K, and 200 ps in a NPT ensemble at 310 K and 1 atm, with the heavy atoms of  $\alpha$ -synuclein chains and SK9 chains constraint with a force constant of 1000 kJ mol<sup>-1</sup> nm<sup>-2</sup>.

Production simulations are started from the above generated initial conformations and run at 310 K and 1 atm. Temperature is controlled by a v-rescale thermostat with a coupling constant of 0.1 ps, and pressure kept constant by using the Parrinello-Rahman barostat with a pressure relaxation time of 2 ps. Water molecules are kept rigid using the SETTLE algorithm<sup>28</sup>, and non-water bonds involving hydrogen atoms are restrained to their equilibrium lengths using the LINCS<sup>29</sup> algorithm, allowing us to use a time step of 2 fs for integrating the equations of motions. Long-range electrostatic interactions are, because of the periodic boundary conditions, computed with the particle-mesh Ewald (PME) technique, using a real-space cutoff of 12 Å and a Fourier grid spacing of 1.6 Å. Short-range van der Waal interactions are truncated at 12 Å, with smoothing starting at 10.5 Å. For each setup, we follow three independent trajectories starting from different initial velocity distributions. The position of the SK9 segment is not restrained to the original docking sites. Instead, it can move and is even allowed to detach over the course of the simulations. The length of the various trajectories is also listed in **Table 1**.



**Table 1:** Simulated systems.

system description		atoms	water molecules	Independent trajectories	simulation length	total sampling
$\alpha$ -synuclein1-140 monomer, (PDB-ID:1XQ8 model)	w/o SK9	287,045	71,155	3	3.0 $\mu$ s	9.0 $\mu$ s
	w SK9	288,809	71,508	3	3.0 $\mu$ s	9.0 $\mu$ s
$\alpha$ -synuclein1-140 monomer, (fully extended conformation)	w/o SK9	284,469	70,512	2	3.0 $\mu$ s	6.0 $\mu$ s
	w SK9	329,376	81,683	2	3.0 $\mu$ s	6.0 $\mu$ s
Rod fibril model (PDB ID: 6CU7, resolved for residues 38-97)	w/o SK9	183,222	58,098	3	0.2 $\mu$ s	0.6 $\mu$ s
	w SK9	183,383	57,585	3	0.2 $\mu$ s	0.6 $\mu$ s
Extended rod fibril (residue 38-120)	w/o SK9	516,415	168,653	3	0.2 $\mu$ s	0.6 $\mu$ s
	w SK9	516,587	168,273	3	0.2 $\mu$ s	0.6 $\mu$ s
Twister fibril model (PDB ID: 6CU8, resolved for residues 43-83)	w/o SK9	177,275	56,889	3	0.2 $\mu$ s	0.6 $\mu$ s
	w SK9	177,583	56,525	3	0.2 $\mu$ s	0.6 $\mu$ s
Extended twister fibril (residue 38-120)	w/o SK9	324,814	104,900	3	0.2 $\mu$ s	0.6 $\mu$ s
	w SK9	325,049	104,541	3	0.2 $\mu$ s	0.6 $\mu$ s
All trajectories						34.8 $\mu$ s

### *Trajectory Analysis*

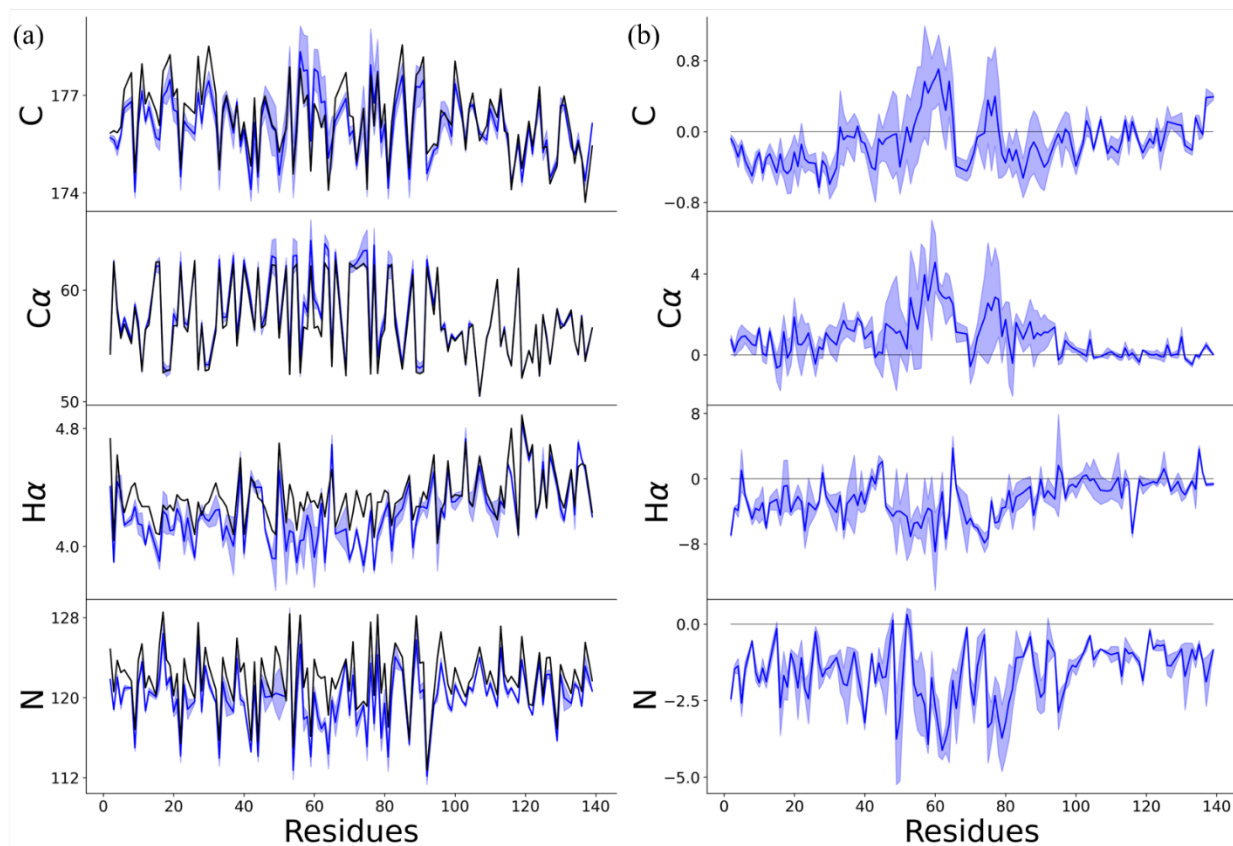
For most of our analyses we rely on the GROMACS tools and VMD software. For calculating root mean square deviation (RMSD) and root mean square fluctuation (RMSF) with respect to the experimentally solved structure we use the GROMACS tools `gmx_rms` and `gmx_rmsf`, respectively. The radius of gyration is calculated using the `gmx_gyrate` utility, while the residue-wise secondary structural propensity is computed with the Dictionary of Secondary Structure in Proteins (DSSP) as implemented in the GROMACS `do_dssp` tool. The solvent accessible surface area (SASA) and residue-residue contact maps are computed with the VMD software, using a spherical probe of 1.4 Å radius for SASA calculations, and defining contacts by a cutoff of 7.0 Å in the closest distance between heavy atoms in a residue pair.

## Results

### 1) MONOMER SIMULATIONS

The purpose of this paper is to explore whether and how SARS-COV-2 proteins can change the amyloid formation of  $\alpha$ -synuclein. We first look into the effect of the SK9 segment of the Envelope protein on early steps in the oligomerization of  $\alpha$ -synuclein. Comparing molecular dynamics simulations of the protein in the presence of SK9 with such where SK9 is absent (our control), we study how the ensemble of  $\alpha$ -synuclein monomer conformations is altered by presence of SK9, increasing or decreasing the likelihood of amyloid formation. The trustworthiness of our results depends on a faithful description of the  $\alpha$ -synuclein monomer (i.e., sufficiently accurate force fields), and sufficient sampling of the various systems. Hence, in order to ensure that our data are reliable and not biased by our set-up, we first compare our simulations of the control with experimental data. While  $\alpha$ -synuclein is an intrinsically disordered protein, some structural information has been obtained from NMR experiments<sup>30</sup>. These chemical shift data can be compared with such generated from our simulations using the ShiftX software<sup>31</sup>. We show in **Figure 2a** both the experimentally measured values and the ones calculated from our simulations. The corresponding percentage changes for each residue are shown in **Figure 2b**, where the black line marks zero change. While these figures show a systematic deviation, the percentual change is small and the regression plots in the **Supplemental Figure S1** lead to large Pearson correlation coefficients. We remark that the deviations between numerical and NMR measurements are consistent with previously reported coarse-grained molecular dynamic studies<sup>32,33</sup>, and may arise from either systematic errors in the NMR measurements due to reference quantification or a dilution effect<sup>34</sup>, or from use of non-polarizable force fields, leading to systematic errors in the calculations due to the absence of polarization in the  $H_{\alpha}$  atoms<sup>35</sup>. Note that, the chemical shift

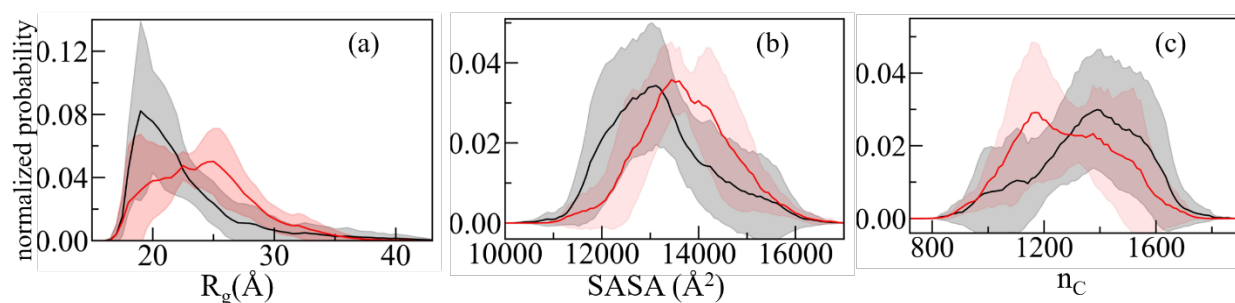
values do not differ between the various trajectories of the isolated  $\alpha$ -synuclein monomer suggesting sufficient sampling to achieve equilibrium.



**Figure 2:** (a) Chemical shifts of carbonyl carbons (C),  $\alpha$ -carbons ( $C\alpha$ ), hydrogen atoms attached to  $\alpha$ -carbons ( $H\alpha$ ), and amide nitrogens (N) for  $\alpha$ -synuclein monomer, as taken from NMR<sup>30</sup> (in black) and simulations (in blue). Percentage change in chemical shift values with respect to experimentally measured values are shown in (b) with the black line marking zero deviation.

Our above analysis shows that our simulation set-up describes faithfully the ensemble of  $\alpha$ -synuclein monomers and gives us confidence in the reliability of the data measured in these simulations. First, we look into the overall behavior of the chains. Visual inspections indicates that in the presence of SK9, the  $\alpha$ -synuclein monomers are more extended and strand-like. This observation is supported by **Figure 3** where we show in **Figure 3a** the distribution of the radius of gyration ( $R_g$ ), in **Figure 3b** the distribution of the solvent accessible surface area (SASA) and in

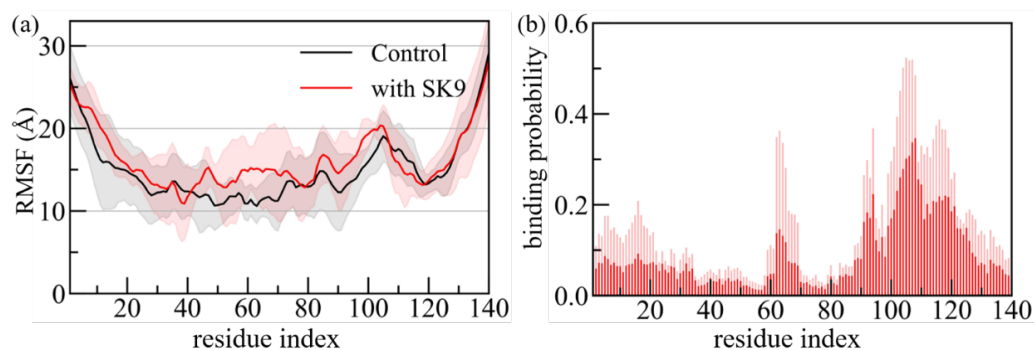
**Figure 3c** that of the number of contacts ( $n_C$ ). We observe in all cases the same trend: the ensemble is in the presence of SK9 shifted toward larger, more solvent-exposed and looser-packed conformations. For the solvent accessible surface area, the effect does not result from increased exposure of hydrophilic residues as it is also seen when the calculation is restricted to hydrophobic residues (in which case the SASA value changes from  $(4404 \pm 522) \text{ \AA}^2$  in the control to  $(4651 \pm 494) \text{ \AA}^2$ ). Hence, the presence of SK9 shifts the distribution of  $\alpha$ -synuclein conformations toward an ensemble with conformations that are more extended than seen in the control, and that expose more hydrophobic residues, making the conformations more aggregation prone.



**Figure 3:** Normalized distribution of (a) radius of gyration ( $R_g$ ), (b) solvent accessible surface area (SASA) and (c) the number of contacts ( $n_C$ ) of  $\alpha$ -synuclein monomer in absence (black) and presence (red) of the SK9-segment. Data are averaged over the final 2.0  $\mu\text{s}$  of each trajectory.

In order to understand how interaction with the viral protein fragment SK9 moves the ensemble of  $\alpha$ -synuclein monomer conformations toward more aggregation prone conformations, we show in **Figure 4a** the root-mean-square-fluctuations (RMSF) of the residues, a quantity that measures the flexibility of a given residue. Comparing values measured in simulations that include SK9 with simulations where SK9 is absent, allow us therefore to identify the regions in the  $\alpha$ -synuclein chain affected by SK9. **Figure 4a** shows that the  $\alpha$ -synuclein monomer is more flexible in the simulations where the viral protein segment is present. An exception is the C-terminal region (residues 100 – 140), which is disorganized in the solution NMR structure of 1XQ8 and which

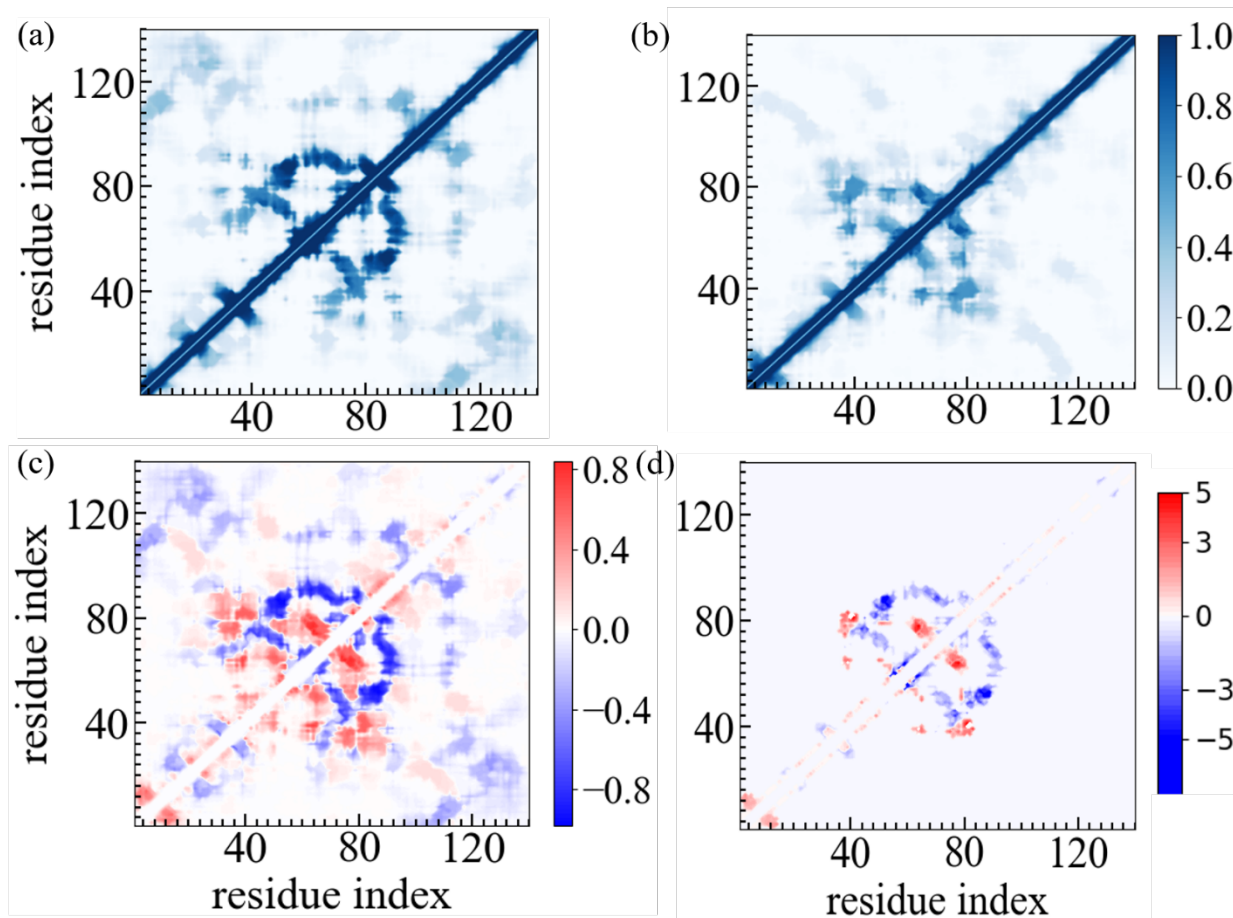
corresponds to the main binding site of SK9, see **Figure 4b**. However, in the other two segments that have enhanced binding probability, residues Q62 to G68 and residues T92 to V95, both located in the NAC region (residues E61 to V95), the binding leads to a higher flexibility and to reduced helicity, see Supplemental **Figure 2a**. As a consequence, the average helicity in the NAC region is reduced from approximately 13% to about 7%. Only a small reduction in helicity (from 5% to 4%) is seen for residues A27-T54, where most of the known familial mutations are located and which is assumed to play a key-role in the early stages of  $\alpha$ -synuclein mis-folding and aggregation.<sup>32,36-38</sup> Similar to the NAC region of residues E61-V95 this region is also characterized by an increased tendency for  $\beta$ -hairpin formation. In the absence of the viral segment SK9, the  $\beta$ -hairpin is transient and is observed with a frequency of less than 11% for residues A27-T54 and for the NAC region. On the other hand, in the presence of SK9, the  $\beta$ -hairpin is formed at an earlier stage and is found in the two regions with a frequency of 15% and 14%, respectively.



**Figure 4:** (a) Residue-wise mean square fluctuation (RMSF; in Å) obtained from  $\alpha$ -synuclein monomer simulations in the absence (in black) and in the presence (in red) of the SK9-segment. The RMSF values are calculated with reference to the experimentally resolved  $\alpha$ -synuclein monomer (PDB ID: 1XQ8) considering only backbone atoms. The residue-wise normalized binding probability of the SK9 segment for  $\alpha$ -synuclein monomer is shown in (b). Data are averaged over the final 2.00  $\mu$ s of each trajectory. Shaded region mark the standard deviation of the averages.

The changes in secondary structure and in the biophysical properties of the  $\alpha$ -synuclein monomer results from the interaction of the SK9 segment with the above discussed residues. These interactions modify the network of contacts that stabilize the monomer. Hence, we show in **Figure 5a** and **5b** the residue-residue plot of the contact frequencies observed in our simulations for the control, and in presence of SK9. We show in **Figure 5c** the difference of frequencies of contacts between the two cases, and in **Figure 5d** the difference in their life-times. Note that we show only data for contacts appearing with a frequency of at least 50%. In the absence of SK9 most of the contacts are formed in the familial mutation sites and NAC region, with only low probability to form contacts at the N-terminus and disordered C-terminus. Upon SK9-binding,  $\alpha$ -synuclein loses the contacts between residues T44-T54 and A69-V77, located at the familial mutation sites and in the central part of the NAC region, with the average life time of contacts between these regions reduced by about 1.7 ns. A similar loss in contacts is observed between residues K45-K60 and residues K80-G93, where the average lifetime of such contacts is also lowered by about 2.4 ns. Loss of these contacts and such between residues E61-T75 and residues K80-V95 increases the solvent exposure of these regions and also reduces  $\alpha$ -helix propensity of the familial mutation sites and in the NAC region while increasing their flexibility, as also seen in **Figure 3** and **Supplemental Figure S2**. On the other hand, the presence of SK9 increases charged and polar interaction within the NAC region involving the residue pairs K60-Q79, E61-K80, Q62-Q79, T64-T75, N65-Q79 and to a smaller extent hydrophobic interactions involving the residue pairs V63-V77, V66-V74 and V66-V77. Together with newly formed contacts between the residues in the segment T33-S42 and with a subset of the NAC region (T33-T75, V37-V77, V37-V82, L38-V82, Y39-T81, V40-V82 and S42-T81) these new interactions cause the structural

rearrangement of  $\alpha$ -synuclein monomer with its higher  $\beta$ -hairpin frequency and resulting increased probability for aggregation.



**Figure 5:** Residue–residue contact frequencies of  $\alpha$ -synuclein monomer in (a) the absence and (b) in the presence of the SK9 segment. Data were averaged over the final 2.0  $\mu$ s of each trajectory. The differences between the two frequencies are shown in (c) while we plot in (d) the corresponding difference in the average lifetime of contacts, i.e, how much longer on average a contact persists in presence of SK9 compared to its average lifetime in the control simulation.

Note that in presence of SK9 the residues in the monomer that form in the rod fibril polymorph (see **Figure 6**) the inter-protofilament interface (E46-A56) are exposed, thereby increasing the chance to bind with another  $\alpha$ -synuclein chain in this region. This suggests that the presence of SK9 raises preferentially the likelihood for forming the rod polymorph. This assumption is

supported by a comparison of the frequency of contacts that are also observed in either the twister or the rod fibril motif. Two contacts, T54-T75 and E61-T72 are found in both fibril polymorphs, and are seen in the monomer simulations with 7% frequency, with the frequency not affected by the presence of SK9. SK9 also does not noticeably change in the monomer simulations the frequency of two contacts found solely in the rod polymorph: A69-A91 and V77-I88. Both are seen in about 40% of configurations. On the other hand, the presence of SK9 strongly increases the frequency, and doubles the lifetime, of two other stabilizing contacts that are seen in the rod fibril: E46-K80 (from 5% to 36%) and V52-A76 (from 10% to 50%). These two contacts connect residues located in the NAC region thereby fixating the inter-protofilament interface in the rod structure. While we see a similar increase in frequency of contacts in the presence of SK9, from about 16% to 32% between the residue pairs V49-Q79 and V49-V77, which stabilize the inter-protofilament interface in the individual chains in the twister, there is no noticeable change in their average lifetime of about 1 ns. However, the increase in frequency of the two stabilizing contacts is partially compensated by a lower frequency of the contact E46-T81 which, competing with the E46-K80 contact, decreases from 27% to 6%, and a decrease in the frequency of the V63-V70 contact from 57% in the control to 23%. While the average lifetime of the E46-T81 contact decreases from 1.5 ns to 0.3 ns, it stays unchanged at around 1 ns for the V63-V70 contact. Hence, our analysis of the monomer simulations suggests not only that the presence of the SARS-COV-2 protein fragment SK9 raises the probability for  $\alpha$ -synuclein amyloid formation, but also that this effect favors the rod motif.

## **2) FIBRIL SIMULATIONS**

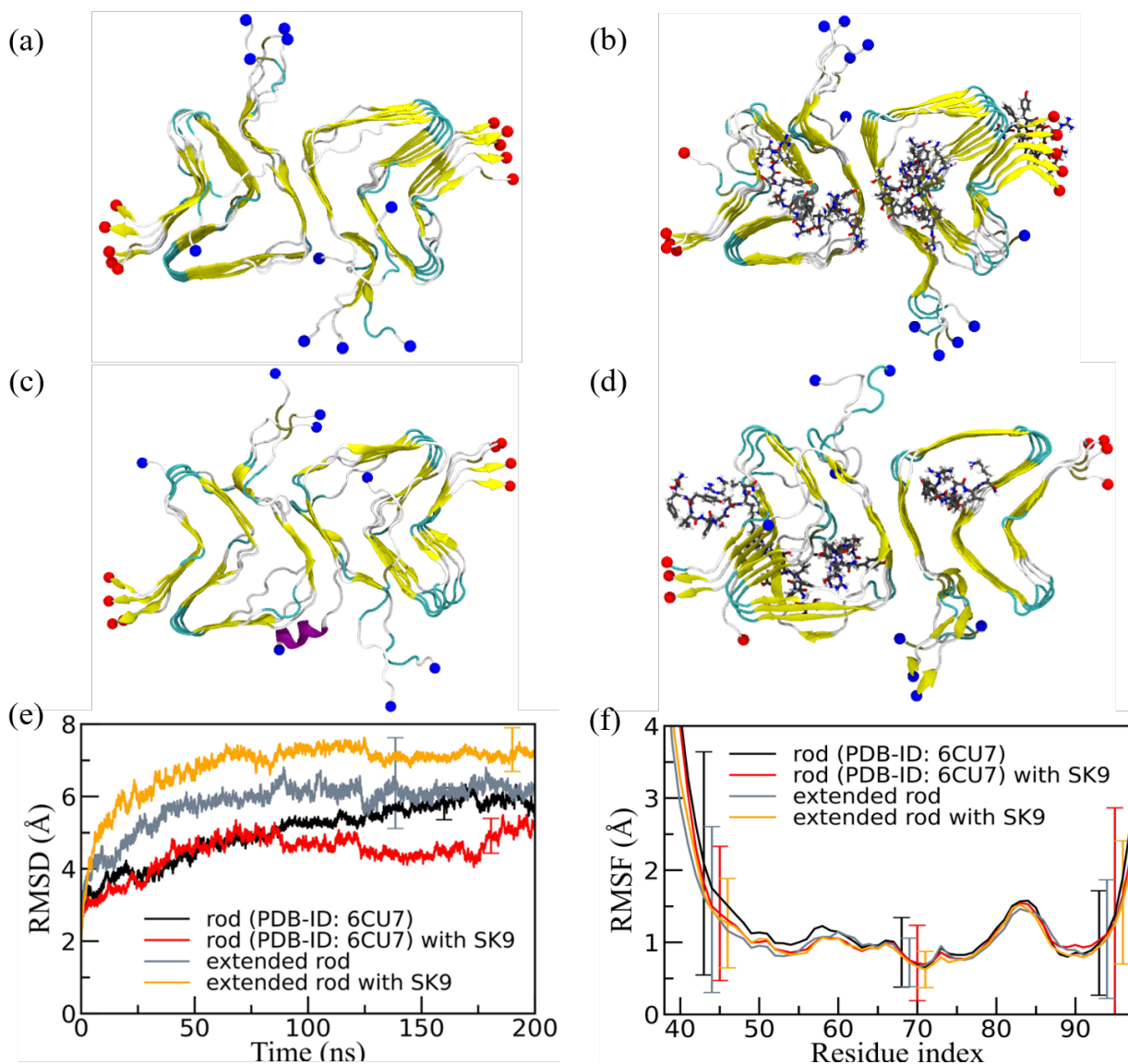
The net-effect of the interaction between SK9 and  $\alpha$ -synuclein monomers is a shift in the conformational ensemble toward monomer conformations that are more prone to aggregate.



Modulating the association kinetics in this way, the SARS-COV-2 protein segment potentially encourages association of the monomers into a narrow distribution of only few oligomer types which may favor seeding of rod-like fibrils. However, SARS-COV-2 proteins can also enhance amyloid formation by a second mechanism. By raising the stability of the fibril polymorphs differentially, the Le Chatelier principle implies that interactions with viral proteins may shift the equilibrium between monomers, oligomers and fibrils toward specific fibril polymorphs. For this reason, we have performed a second set of molecular dynamics simulations studying the effect of SK9 on the stability of the two most common  $\alpha$ -synuclein fibril structures<sup>16</sup>, the rod and the twister polymorphs of **Figure 1**. Our goal is to investigate whether the presence of SK9 stabilizes  $\alpha$ -synuclein fibrils; and if yes, how this effect depends on the fibril geometry. As for the monomer simulations we compare for the two fibril polymorphs trajectories where SK9 fragments bind initially to the  $\alpha$ -synuclein (but can change position or even detach), with our control, the corresponding simulations where the SK9 fragments are absent. Since only parts of the chain conformations have been resolved in the experimental models, and these models differ in length for rod and twister, we have also studied corresponding fibril models where missing residues are added to achieve equal chain length. Note, however, that in order to reduce the computational costs these models are not made of full-size  $\alpha$ -synuclein chains but only of fragments 38-120. For details, see the method section.

A visual inspection of the final conformations in **Figure 6a-6d**, where for a more easy comparison we display only the residues 38-97, shows only marginal stabilization by SK9 for the rod fibril architecture when only the experimentally resolved fragment of residues 38 to 97 is simulated, and even some destabilization for the extended rod fibril (residues 38-120). The visual inspection is quantified by the time evolution in RMSD in **Figure 6e** and the residue-wise RMSF in **Figure 6f**. Both quantities are again calculated only for residues 38 to 97, i.e., ignoring the disordered and unresolved parts of the extended fibril model. Note, however, that the error bars are in both cases large, suggesting that the presence of SK9 has only little effect on the stability of the rod-fibril architecture. This is consistent with the rather small

changes in SASA, packing distance between the folds, number of both the packing contacts (between folds) and stacking contacts (between the layers in each fold), and the frequencies of native inter- and intra-chain contacts, see **Table 2**.

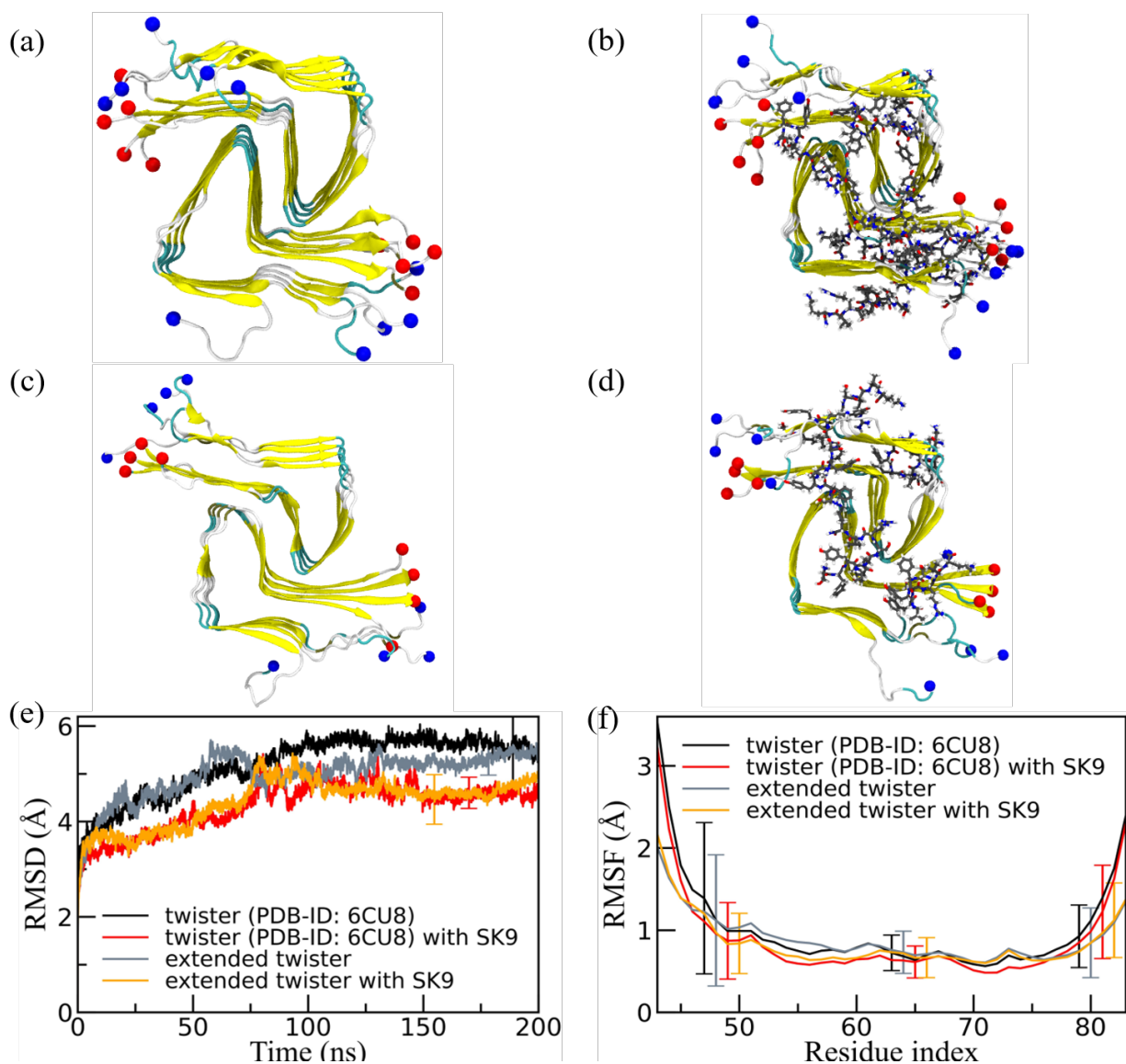


**Figure 6:** Representative final configurations extracted from simulations starting from (a) the experimentally determined rod-like  $\alpha$ -synuclein fibril model (PDB-ID: 6CU7) and (c) the extended model. Corresponding final snapshots extracted from simulations in the presence of SK9-segment are shown in (b), and (d). N- and C-terminus are represented by blue and red spheres, respectively. Only residues 38-97 are shown for the extended model configurations in (c) and (d). The time evolution of the RMSD in the simulation of these systems is shown in (e), and residue-wise RMSF in (f). We calculate RMSD and RMSF again only for the experimentally resolved region 38-97, i.e., ignoring the disordered and unresolved parts of the fibril models, considering all backbone atoms. Only a few typical errorbars are shown to make figures more readable.

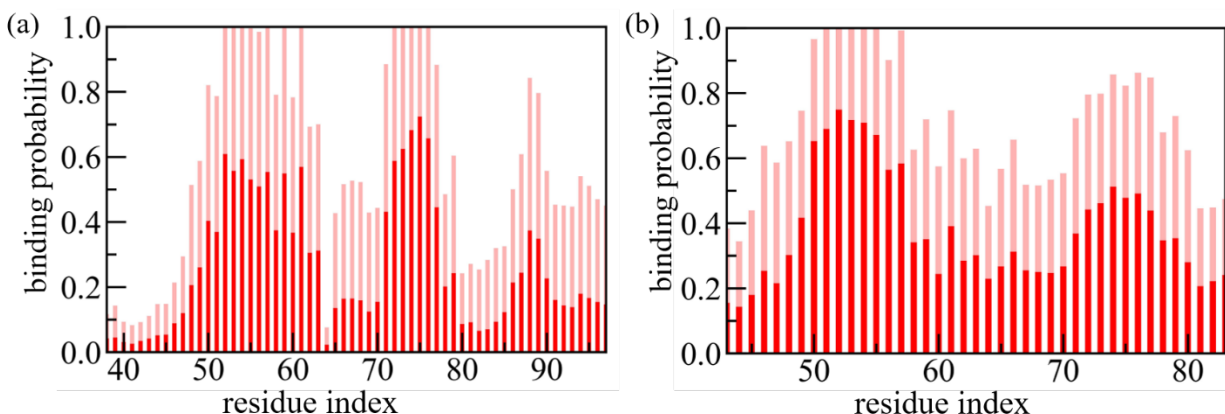
**Table 2:** Mean values of total SASA, packing distance, number of intrachain, stacking and packing contacts, and frequency of native interchain and intrachain contacts. All values are averages over the final 50 ns of all trajectories of the respective systems. Note that we consider only for the experimentally resolved regions in our calculation of these quantities, i.e., ignoring the disordered and unresolved parts of the fibril models.

system		SASA (Å <sup>2</sup> )	packing distance (Å)	intrachain contacts	stacking contacts	packing contacts	% of native interchain contacts	% of native intrachain contacts
Rod fibril model (PDB-ID: 6CU7; residues 38-97)	control	29472 (878)	10.4 (0.7)	56 (4)	359 (4)	27 (2)	84 (6)	60 (5)
	with SK9	30490 (656)	10.6 (0.7)	53 (5)	354 (5)	23 (2)	84 (6)	60 (5)
Extended rod fibril (residues 38-120)	control	22514 (423)	10.0 (0.9)	56 (3)	350 (17)	25 (3)	82 (4)	60 (4)
	with SK9	27874 (685)	10.6 (0.6)	62 (7)	332 (6)	24 (2)	77 (6)	60 (4)
Twister fibril model PDB-ID: 6CU8; residues 43-83)	control	20918 (615)	10.6 (0.4)	19 (2)	221 (6)	42 (2)	80 (6)	46 (4)
	with SK9	21717 (485)	10.3 (0.7)	17 (3)	220 (3)	40 (1)	81 (4)	46 (4)
Extended twister fibril model (residues 38-120)	control	15008 (389)	9.6 (0.8)	21 (3)	215 (5)	38 (2)	80 (6)	44 (5)
	with SK9	19882 (771)	9.8 (0.8)	13 (6)	213 (10)	38 (3)	80 (6)	35 (5)

On the other hand, both visual inspection and analysis of RMSD and RMSF in the corresponding **Figure 7** show a signal for stabilization of the twister  $\alpha$ -synuclein fibril structure in the simulations that started with SK9 bound to the fibrils, but again no clear signal is seen for this stabilization in the quantities listed in **Table 2**. This may be because the effect of SK9 is localized: the RMSF plots in **Figure 7f** suggests a lower flexibility of the region 45-65 in simulations where the viral protein fragment SK9 is present. This region is one of the main binding-partners for SK9, see **Figure 8** where we show the binding probability as measured in simulations starting from fibril models with non-extended chains (the corresponding figures for the extended fibril models are shown as **Supplemental Figure S3**).



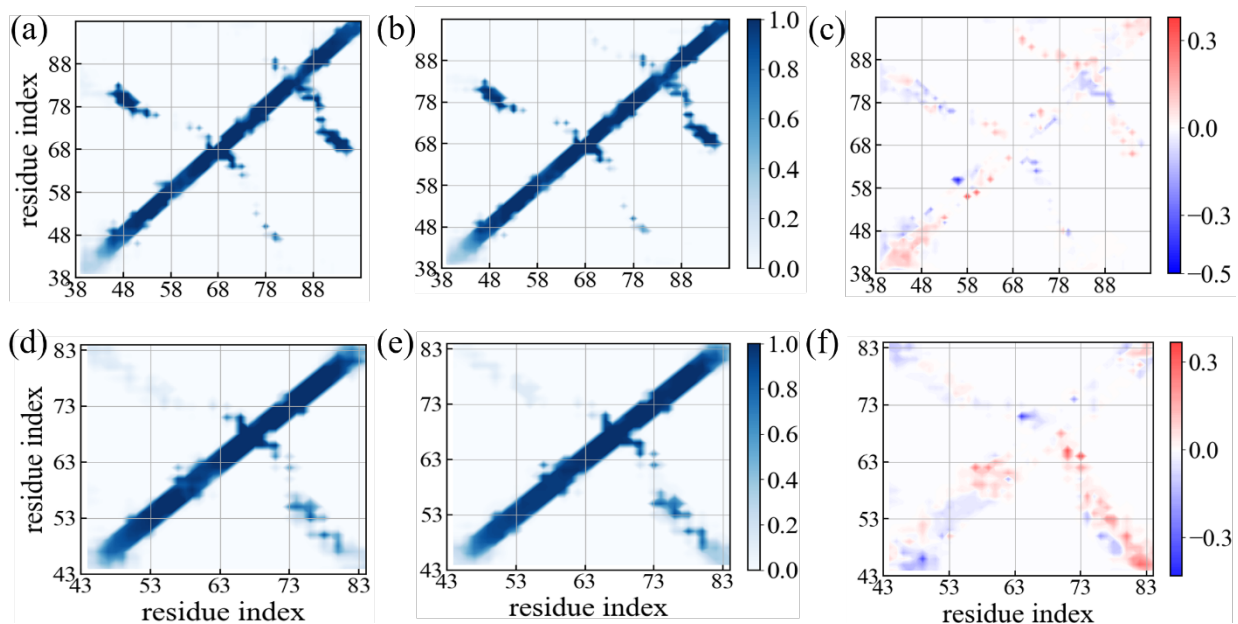
**Figure 7:** Representative final configurations extracted from simulations starting from (a) the experimentally determined twister-like  $\alpha$ -synuclein fibril model (PDB-ID: 6CU8) and (c) the extended model. Corresponding final snapshots extracted from simulations in the presence of SK9-segment are shown in (b), and (d). N- and C-terminus are represented by blue and red spheres, respectively. Only residues 43-83 are shown for the extended model configurations in (c) and (d). The time evolution of the RMSD in the simulation of these systems is shown in (e), and residue-wise RMSF in (f). We calculate RMSD and RMSF again only for the experimentally resolved region 43-83, i.e., ignoring the disordered and unresolved parts of the fibril models, considering all backbone atoms. Only a few typical errorbars are shown to make figures more readable.



**Figure 8:** The residue-wise normalized binding probability of the SK9 segment for  $\alpha$ -synuclein rod (a) and twister (b) polymorph. Data are averaged over the final 50 ns of each trajectory and shaded region represents the standard deviation.

In order to understand in more detail changes in fibril stability upon addition of SK9, we have looked into the change of contact pattern in both fibril geometries. In **Figure 9** we show contact probabilities for stacking contacts (i.e., such between residues located on neighboring layers of the same fold) for the rod-like fibril in the absence (**Figure 9a**) and in the presence of (**Figure 9b**) the SK9 segment. The corresponding probabilities for the twister fibril are shown in **Figure 9d** (absence of SK9) and in **Figure 9e** (presence of the SK9 segment). The differences between probabilities measured in the presence and absence of SK9 are shown in **Figure 9c** (rod fibril) and **Figure 9f** (twister fibril). The data shown in **Figure 9** are from the simulations using the same chain-length as in the experimentally determined fibril models. The corresponding probabilities as measured in simulations of the fibrils with extended chains are shown in **Supplemental Figure S4**. As shown already in **Table 2**, native interchain contacts are more preserved than intrachain contacts. Most of the native contact pairs in the NAC region are highly conserved over the simulation, but hydrophobic contacts involving F94, the T59-T72 contacts and such with the G67 and G68 residues in the turn region are seen with a frequency of less than 10%.

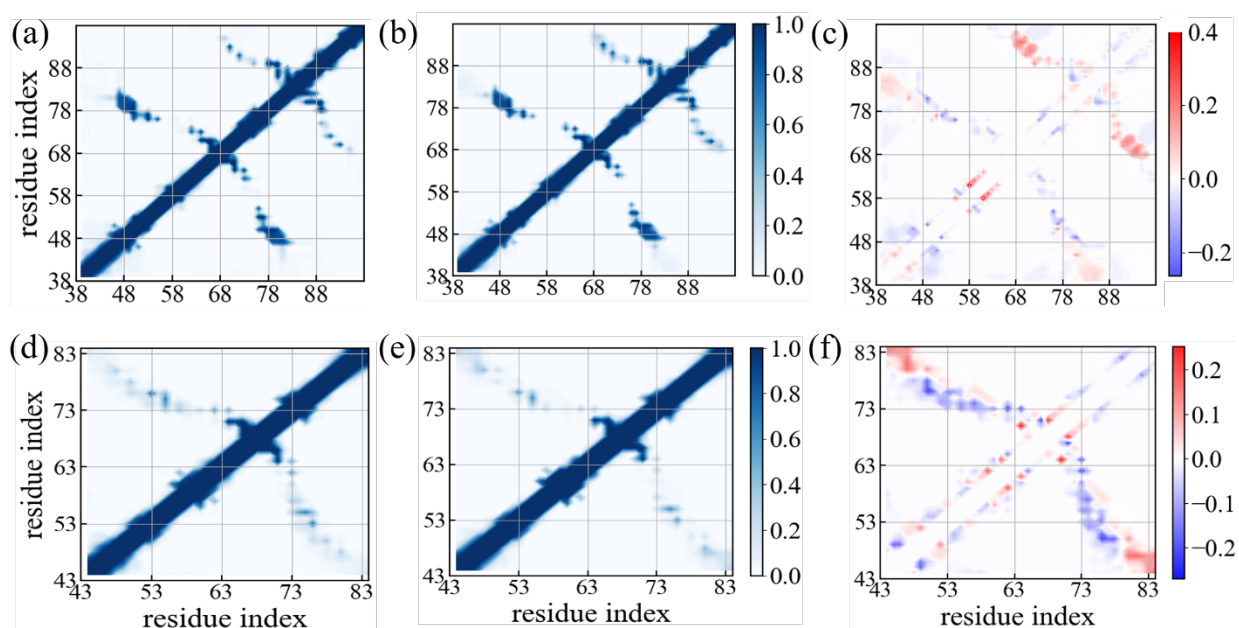
However, with the exception of a gain of contacts between residues in the segment of G68-A78 with the segment of residues I88-K97 no signal for changes in the stacking pattern upon binding of SK9 is found for the rod-like fibril. The exception can be explained by the preferential binding of SK9 to residues V70-A78, which stabilized this segment (see also the difference in intrachain contacts in **Figure 10c**) and leads to the observed stacking contacts.



**Figure 9:** Residue-wise stacking contact probabilities measured in simulations of the rod-like fibril in the absence (a) and in the presence of the SK9 segment (b). Corresponding data measured in simulations of the twister-like fibril are shown in (d) and (e). The differences in contact probability resulting from the presence of SK9 are shown in (c) and (f), respectively. Data are averaged over the final 50 ns of each trajectory.

On the other hand, while native interchain contacts are only slightly less stable in the twister fibril than in the rod, we find a larger reduction in the frequency of native intrachains, see **Table 2**. Especially, native hydrophobic contacts involving the residue pairs V66-A69, V66-A70 and V74-V77, and native contacts involving glycine residues (G67, G73) in the inter-protofilament interface are disrupted. Upon SK9-binding, we observe a significant increase in stacking contacts

between residues of the segment K45-N65 with that of the segment G73-K80, which is compensated by a pronounced loss in the intrachain contacts in the same region, see **Figure 10** and **Supplemental Figure S5**. Hence, interaction with SK9 leads in the Twister fibril to a replacement of contacts between residues within chains by contacts between residues located on chains in neighboring layers. This may explain why the observed stabilization of the twister fibril geometry upon binding of SK9, seen in the RMSD plots, does not lead to more pronounced changes in the quantities shown in **Table 2**.



**Figure 10:** Residue-wise intrachain contact probabilities measured in simulations of the rod-like fibril in the absence (a) and in the presence of the SK9 segment (b). Corresponding data measured in simulations of the twister-like fibril are shown in (d) and (e). The differences in contact probability resulting from the presence of SK9 are shown in (c) and (f), respectively. Data are averaged over the final 50 ns of each trajectory.

## Conclusions

As virus-induced amyloid-formation has been observed *in vitro*<sup>5</sup>, and may suggest onset of Parkinson's Disease and related neurodegenerative diseases<sup>2-5</sup> as potential long-term risk for

COVID-19 patients, it is important to investigate how amyloidogenic regions of SARS-COV-2 initiate or accelerate formation of  $\alpha$ -synuclein aggregates, the cytotoxic agent in Parkinson's Disease.<sup>6</sup> For this purpose, we have studied, using all-atom molecular dynamics simulations, how interaction with a short segment (SK9) of the Envelope protein of SARS-COV-2 changes the conformational ensemble of  $\alpha$ -synuclein chains and the stability of two  $\alpha$ -synuclein fibril polymorphs, the rod and the twister structure.

We find that the presence of SK9 changes the ensemble of  $\alpha$ -synuclein not only toward more aggregation-prone conformations but that this effect likely leads to a preference for the rod fibril motif. Crucial for this effect is a change in the contact pattern that causes higher flexibility, reduced helix-propensity and larger exposure of residues, especially in the segment E46-A56 that form in the rod fibril polymorph (see **Figure 6**) the inter-protofilament interface. Especially important is the increase infrequency and lifetime of the two contacts E46-K80 and V52-A76. Both are found also in the experimentally resolved rod-fibril form, and the mutation E46K, which destroys the salt-bridge E46-K80, is one of the familial mutations that cause Parkinson's Disease and that shift the equilibrium from rod to twister fibrils.<sup>16</sup>

The outcome of our fibril simulations is less obvious. While we see some stabilization of the fibril geometries in the presence of SK9, the signal is weak, and it is only for the twister form correlated with a clear change in the contact pattern. Hence, it appears that the viral protein fragment SK9 may alter  $\alpha$ -synuclein amyloid formation but has little effect of the stability of pre-existing or newly-formed fibrils.



The preference for monomer conformations that likely seed the rod-like fibrils is interesting as the twister fibril is considered to be more cytotoxic.<sup>16</sup> Further simulations, using other amyloidogenic segments of SARS-COV-2 proteins, especially in the Spike-protein, are needed to see if this effect is specific for the SK9 segment. It will be especially interesting to compare our work with recent studies into the amyloidogenesis of Spike-protein fragments after endoproteolysis.<sup>39</sup>

## **AUTHOR INFORMATION**

### **Corresponding Author**

**Asis K. Jana** - *Department of Chemistry & Biochemistry, University of Oklahoma, Norman, Oklahoma 73019, United States*

**Ulrich H.E. Hansmann** - *Department of Chemistry & Biochemistry, University of Oklahoma, Norman, Oklahoma 73019, United States; Email: [uhansmann@ou.edu](mailto:uhansmann@ou.edu)*

### **Authors**

**Chance W. Lander** - *Department of Chemistry & Biochemistry, University of Oklahoma, Norman, Oklahoma 73019, United States*

**Andrew D. Chesney** - *Department of Chemistry & Biochemistry, University of Oklahoma, Norman, Oklahoma 73019, United States*

### **Notes**

The authors declare no competing financial interest.

### **Acknowledgment**

The simulations in this work were done using the SCHOONER cluster of the University of Oklahoma, XSEDE resources allocated under grant MCB160005 (National Science Foundation),

and TACC resources allocated under grant under grant MCB20016 (National Science Foundation).

We acknowledge financial support from the National Institutes of Health under grant GM120634.

## REFERENCES

- (1) Liu, J.; Tu, C.; Zhu, M.; Wang, J.; Yang, C.; Liu, W.; Xiong, B. The Clinical Course and Prognostic Factors of Severe COVID-19 in Wuhan, China: A Retrospective Case-Control Study. *Medicine* **2021**, *100* (8), e23996–e23996.
- (2) Cohen, M. E.; Eichel, R.; Steiner-Birmanns, B.; Janah, A.; Ioshpa, M.; Bar-Shalom, R.; Paul, J. J.; Gaber, H.; Skrahina, V.; Bornstein, N. M.; Yahalom, G. A Case of Probable Parkinson's Disease after SARS-CoV-2 Infection. *Lancet Neurol.*, **2020**, *19* (10), 804–805.
- (3) Brundin, P.; Nath, A.; Beckham, J. D. Is COVID-19 a Perfect Storm for Parkinson's Disease? *Trends Neurosci.*, **2020**, *43* (12), 931–933.
- (4) Ait Wahmane, S.; Achbani, A.; Ouhaz, Z.; Elatiqi, M.; Belmouden, A.; Nejmeddine, M. The Possible Protective Role of  $\alpha$ -Synuclein Against Severe Acute Respiratory Syndrome Coronavirus 2 Infections in Patients With Parkinson's Disease. *Mov. Disord.*, **2020**, *35* (8), 1293–1294.
- (5) Semerdzhiev, S. A.; Fakhree, M. A. A.; Segers-Nolten, I.; Blum, C.; Claessens, M. M. A. E. Interactions between SARS-CoV-2 N-Protein and  $\alpha$ -Synuclein Accelerate Amyloid Formation. *ACS Chem. Neurosci.*, **2022**, *13* (1), 143–150.

- (6) Spillantini, M. G.; Schmidt, M. L.; Lee, V. M.-Y.; Trojanowski, J. Q.; Jakes, R.; Goedert, M.  $\alpha$ -Synuclein in Lewy Bodies. *Nature* **1997**, *388* (6645), 839–840.
- (7) Frölich, L. Alzheimer's Disease – the 'Microbial Hypothesis' from a Clinical and Neuroimaging Perspective. *Psychiatry Res. Neuroimaging* **2020**, *306*, 111181.
- (8) Seaks, C. E.; Wilcock, D. M. Infectious Hypothesis of Alzheimer Disease. *PLoS Pathog.*, **2020**, *16* (11), e1008596–e1008596.
- (9) Eimer, W. A.; Vijaya Kumar, D. K.; Navalpur Shanmugam, N. K.; Rodriguez, A. S.; Mitchell, T.; Washicosky, K. J.; György, B.; Breakefield, X. O.; Tanzi, R. E.; Moir, R. D. Alzheimer's Disease-Associated  $\beta$ -Amyloid Is Rapidly Seeded by Herpesviridae to Protect against Brain Infection. *Neuron* **2018**, *99* (1), 56-63.e3.
- (10) Jana, A. K.; Greenwood, A. B.; Hansmann, U. H. E. Presence of a SARS-CoV-2 Protein Enhances Amyloid Formation of Serum Amyloid A. *J. Phys. Chem. B* **2021**, *125* (32), 9155–9167.
- (11) Ghosh, A.; Pithadia, A. S.; Bhat, J.; Bera, S.; Midya, A.; Fierke, C. A.; Ramamoorthy, A.; Bhunia, A. Self-Assembly of a Nine-Residue Amyloid-Forming Peptide Fragment of SARS Corona Virus E-Protein: Mechanism of Self Aggregation and Amyloid-Inhibition of HIAPP. *Biochemistry* **2015**, *54* (13), 2249–2261.
- (12) Ulmer, T. S.; Bax, A.; Cole, N. B.; Nussbaum, R. L. Structure and Dynamics of Micelle-Bound Human Alpha-Synuclein \*. *J. Biol. Chem.* **2005**, *280* (10), 9595–9603.
- (13) Eugene, C.; Laghaei, R.; Mousseau, N. Early Oligomerization Stages for the Non-Amyloid Component of  $\alpha$ -Synuclein Amyloid. *J. Chem. Phys.*, **2014**, *141* (13), 135103.

- (14) Trexler, A. J.; Rhoades, E. Function and Dysfunction of  $\alpha$ -Synuclein: Probing Conformational Changes and Aggregation by Single Molecule Fluorescence. *Mol. Neurobiol.*, **2013**, *47* (2), 622–631.
- (15) Bartels, T.; Choi, J. G.; Selkoe, D. J.  $\alpha$ -Synuclein Occurs Physiologically as a Helically Folded Tetramer That Resists Aggregation. *Nature* **2011**, *477* (7362), 107–110.
- (16) Li, B.; Ge, P.; Murray, K. A.; Sheth, P.; Zhang, M.; Nair, G.; Sawaya, M. R.; Shin, W. S.; Boyer, D. R.; Ye, S.; Eisenberg, D. S.; Zhou, Z. H.; Jiang, L. Cryo-EM of Full-Length  $\alpha$ -Synuclein Reveals Fibril Polymorphs with a Common Structural Kernel. *Nat. Commun.*, **2018**, *9* (1), 3609.
- (17) Guerrero-Ferreira, R.; Taylor, N. M. I.; Mona, D.; Ringler, P.; Lauer, M. E.; Riek, R.; Britschgi, M.; Stahlberg, H. Cryo-EM Structure of Alpha-Synuclein Fibrils. *elife* **2018**, *7*, e36402.
- (18) Chakraborty, R.; Chattopadhyay, K. Cryo-Electron Microscopy Uncovers Key Residues within the Core of Alpha-Synuclein Fibrils. *ACS Chem. Neurosci.*, **2019**, *10* (3), 1135–1136.
- (19) Kryshchak, A.; Moulton, J.; Billings, W. M.; Della Corte, D.; Fidelis, K.; Kwon, S.; Olechnovič, K.; Seok, C.; Venclovas, Č.; Won, J.; participants, C.-C. Modeling SARS-CoV-2 Proteins in the CASP-Commons Experiment. *Proteins Struct. Funct. Bioinform.*, **2021**, *89* (12), 1987–1996.
- (20) Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.*, **2010**, *31* (2), 455–461.

- (21) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.
- (22) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D. CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat. Methods* **2017**, *14* (1), 71–73.
- (23) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.*, **1983**, *79* (2), 926–935.
- (24) Wang, W.; Hansmann, U. H. E. Stability of Human Serum Amyloid A Fibrils. *J. Phys. Chem. B* **2020**, *124* (47), 10708–10717.
- (25) Pandey, P.; Nguyen, N.; Hansmann, U. H. E. D-Retro Inverso Amylin and the Stability of Amylin Fibrils. *J. Chem. Theory Comput.*, **2020**, *16* (8), 5358–5368.
- (26) Piana, S.; Donchev, A. G.; Robustelli, P.; Shaw, D. E. Water Dispersion Interactions Strongly Influence Simulated Structural Properties of Disordered Protein States. *J. Phys. Chem. B* **2015**, *119* (16), 5113–5123.
- (27) Robustelli, P.; Piana, S.; Shaw, D. E. Developing a Molecular Dynamics Force Field for Both Folded and Disordered Protein States. *Proc. Natl. Acad. Sci.*, **2018**, *115* (21), E4758 - E4766.
- (28) Miyamoto, S.; Kollman, P. A. Settle: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem.*, **1992**, *13* (8), 952–962.

- (29) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.*, **1997**, *18* (12), 1463–1472.
- (30) Rao, J. N.; Kim, Y. E.; Park, L. S.; Ulmer, T. S. Effect of Pseudorepeat Rearrangement on Alpha-Synuclein Misfolding, Vesicle Binding, and Micelle Binding. *J. Mol. Biol.*, **2009**, *390* (3), 516–529.
- (31) Han, B.; Liu, Y.; Ginzinger, S. W.; Wishart, D. S. SHIFTX2: Significantly Improved Protein Chemical Shift Prediction. *J. Biomol. NMR* **2011**, *50* (1), 43–57.
- (32) Yu, H.; Han, W.; Ma, W.; Schulten, K. Transient  $\beta$ -Hairpin Formation in  $\alpha$ -Synuclein Monomer Revealed by Coarse-Grained Molecular Dynamics Simulation. *J. Chem. Phys.* **2015**, *143* (24), 243142.
- (33) Ramis, R.; Ortega-Castro, J.; Casanovas, R.; Mariño, L.; Vilanova, B.; Adrover, M.; Frau, J. A Coarse-Grained Molecular Dynamics Approach to the Study of the Intrinsically Disordered Protein  $\alpha$ -Synuclein. *J. Chem. Inf. Model.*, **2019**, *59* (4), 1458–1471.
- (34) Sokolenko, S.; Aucoin, M. G. A Correction Method for Systematic Error in  $(1)H$ -NMR Time-Course Data Validated through Stochastic Cell Culture Simulation. *BMC Syst. Biol.*, **2015**, *9*, 51.
- (35) Somavarapu, A. K.; Kepp, K. P. The Dependence of Amyloid- $\beta$  Dynamics on Protein Force Fields and Water Models. *ChemPhysChem* **2015**, *16* (15), 3278–3289.
- (36) Esteban-Martín, S.; Silvestre-Ryan, J.; Bertoncini, C. W.; Salvatella, X. Identification of Fibril-like Tertiary Contacts in Soluble Monomeric  $\alpha$ -Synuclein. *Biophys. J.*, **2013**, *105* (5), 1192–1198.

- (37) Pawar, A. P.; Dubay, K. F.; Zurdo, J.; Chiti, F.; Vendruscolo, M.; Dobson, C. M. Prediction of “Aggregation-Prone” and “Aggregation-Susceptible” Regions in Proteins Associated with Neurodegenerative Diseases. *J. Mol. Biol.*, **2005**, *350* (2), 379–392.
- (38) Mirecka, E. A.; Shaykhalishahi, H.; Gauhar, A.; Akgül, Ş.; Lecher, J.; Willbold, D.; Stoldt, M.; Hoyer, W. Sequestration of a B-Hairpin for Control of A-Synuclein Aggregation. *Angew. Chem. Int. Ed.*, **2014**, *53* (16), 4227–4230.
- (39) Nyström, S.; Hammarström, P. Amyloidogenesis of SARS-CoV-2 Spike Protein. *bioRxiv* **2021**, 2021.12.16.472920.