

# SCIENTIFIC REPORTS



OPEN

## Computational insights into the role of $\alpha$ -strand/sheet in aggregation of $\alpha$ -synuclein

Anand Balupuri, Kwang-Eun Choi &amp; Nam Sook Kang

The  $\alpha$ -synuclein is a major component of amyloid fibrils found in Lewy bodies, the characteristic intracellular proteinaceous deposits which are pathological hallmarks of neurodegenerative diseases such as Parkinson's disease (PD) and dementia. It is an intrinsically disordered protein that may undergo dramatic structural changes to form amyloid fibrils. Aggregation process from  $\alpha$ -synuclein monomers to amyloid fibrils through oligomeric intermediates is considered as the disease-causative toxic mechanism. However, mechanism underlying aggregation is not well-known despite several attempts. To characterize the mechanism, we have explored the effects of pH and temperature on the structural properties of wild-type and mutant  $\alpha$ -synuclein using molecular dynamics (MD) simulation technique. MD studies suggested that amyloid fibrils can grow by monomer. Conformational transformation of the natively unfolded protein into partially folded intermediate could be accountable for aggregation and fibrillation. An intermediate  $\alpha$ -strand was observed in the hydrophobic non-amyloid- $\beta$  component (NAC) region of  $\alpha$ -synuclein that could proceed to  $\alpha$ -sheet and initiate early assembly events. Water network around the intermediate was analyzed to determine its influence on the  $\alpha$ -strand structure. Findings of this study provide novel insights into possible mechanism of  $\alpha$ -synuclein aggregation and promising neuroprotective strategy that could aid alleviate PD and its symptoms.

The  $\alpha$ -sheet is an atypical secondary structure in proteins which is also recognized as pleated sheet or polar pleated sheet. This structure is formed by regular hydrogen bonding between adjacent  $\alpha$ -strands and was firstly proposed by Pauling and Corey in 1951<sup>1-3</sup>. In contrast with  $\alpha$ -helix and  $\beta$ -sheet structures, the  $\alpha$ -sheet residues are not constrained to a specific region of dihedral angles. The  $\alpha$ -strand consists of alternating dihedrals in the right-handed ( $\alpha_R$ ) and left-handed ( $\alpha_L$ ) helical regions of Ramachandran plot. Even though  $\alpha$ -sheet is not commonly witnessed in natural proteins, it is hypothesized to play a central role in various amyloid diseases<sup>4</sup>. The  $\alpha$ -sheet is proposed as a probable intermediate state in the formation of amyloid fibrils on the basis of molecular dynamics (MD) simulations studies. This structure is considered as toxic conformer in amyloid diseases. The  $\alpha$ -sheet structure has been observed for lysozyme, prion proteins,  $\beta$ 2-microglobulin and transthyretin. These peptides and proteins are involved in many protein misfolding and amyloid diseases<sup>4,5</sup>. They undergo a conformational change from predominantly random coil or  $\alpha$ -helix structures to the  $\beta$ -sheet structures which are largely present in amyloid fibrils. Peptide plane flipping mechanism has been reported for direct  $\alpha$ -sheet and  $\beta$ -sheet interconversion<sup>6,7</sup>. Apart from MD simulations, X-ray crystallography structures of designed peptides have confirmed  $\alpha$ -sheets. Furthermore, several experimentally determined crystal and NMR structures in Protein Data Bank (PDB) have been found to adopt  $\alpha$ -strand conformation<sup>4</sup>. Recent experimental studies have reported that designed anti- $\alpha$ -sheet peptides preferentially bind to toxic amyloid precursor (intermediate  $\alpha$ -sheet conformer) as compared to non-aggregated, nontoxic species or mature fibrils and inhibit aggregation of amyloid proteins<sup>8-11</sup>. These peptides provide evidence that  $\alpha$ -strands and  $\alpha$ -sheet are not just computational or theoretical artifacts and they do indeed exist.

The  $\alpha$ -synuclein is principal component of numerous protein aggregates known as Lewy bodies and Lewy neurites. Although Lewy bodies are major pathological hallmarks of Parkinson's disease (PD),  $\alpha$ -synuclein is associated with a number of neurodegenerative diseases known as synucleinopathies<sup>12</sup>. This includes Lewy body variant of Alzheimer's disease (LBVAD), PD dementia (PDD), neurodegeneration with brain iron accumulation type 1 (NBIA1), pure autonomic failure, Down's syndrome, complex of Guam and diffuse Lewy body disease (DLBD)<sup>13-21</sup>. The  $\alpha$ -synuclein is a protein of 140 residues devoid of cysteine and tryptophan. Primary structure

Graduate School of New Drug Discovery and Development, Chungnam National University, Daejeon, 305-764, Republic of Korea. Correspondence and requests for materials should be addressed to N.S.K. (email: [nskang@cnu.ac.kr](mailto:nskang@cnu.ac.kr))

of  $\alpha$ -synuclein is usually divided into three distinct regions: N-terminal (residues 1–60), central region (residues 61–95) and C-terminal (residues 96–140). N-terminal is amphipathic region while C-terminal mainly consists of acidic residues<sup>22</sup>. Most hydrophobic central region which is also termed as non-amyloid- $\beta$  component (NAC), folds into a  $\beta$ -sheet secondary structure and plays critical role in both the aggregation and cytotoxicity<sup>23</sup>. Prior studies suggested that  $\alpha$ -synuclein fibril core comprise of residues 30–110<sup>24,25</sup>. However, it was found that core domain of  $\alpha$ -synuclein could be even shorter and comprise of residues 71–82<sup>26–30</sup>.

$\alpha$ -synuclein is an intrinsically disordered protein that may undergo dramatic structural changes to form amyloid fibrils. Several conformational states for this protein have been observed. Aggregation process from monomers to amyloid fibrils through oligomeric intermediates is considered as the disease-causative toxic mechanism<sup>24</sup>. Various intrinsic and extrinsic factors which could either increase or decrease the rate of  $\alpha$ -synuclein aggregation *in vitro* are known. These factors include temperature, pH, point mutations (A30P, E46K and A53T), NAC region mutations, S-129 phosphorylation, N-terminal acetylation, organic solvents, metal ions and salts<sup>12,31–41</sup>. However, mechanism underlying aggregation is not well-known despite several attempts. Previous studies have suggested the role of a partially folded intermediate in aggregation and fibrillation of  $\alpha$ -synuclein<sup>31,42</sup>. Recently,  $\beta$ -hairpin formation in the region 38–53 (part of N-terminal) was proposed to be responsible for  $\alpha$ -synuclein aggregation<sup>22</sup>. However, the role of  $\alpha$ -strand/sheet in  $\alpha$ -synuclein aggregation has not been studied yet.

As  $\alpha$ -sheet structure was found to be intermediate state in the formation of fibrils for various amyloid proteins<sup>4,5</sup>. In the present article, we investigated the role of  $\alpha$ -strand/sheet in the aggregation of  $\alpha$ -synuclein. To characterize the mechanism, we have explored the effects of pH and temperature on the structural properties of wild-type and mutant  $\alpha$ -synuclein using molecular dynamics (MD) simulation technique. Furthermore, we studied the effect Ser-129 phosphorylation and N-terminal acetylation. Here, we described the structural and dynamic changes in  $\alpha$ -synuclein and proposed a model for aggregation and fibrillation of this protein. MD studies suggested that amyloid fibrils could grow by monomer. Conformational transformation of the natively unfolded protein into  $\alpha$ -strand/sheet intermediate could be accountable for fibrillation. Our results provide structural insights which could be utilized for the development of anti-PD drugs targeting the NAC region of  $\alpha$ -synuclein.

## Materials and Methods

**Structures of human  $\alpha$ -synuclein and mutants.** Several structures of  $\alpha$ -synuclein have been determined using X-ray, NMR and electron microscopy techniques, yet most of the structures are available for small fragments of  $\alpha$ -synuclein (<http://www.rcsb.org/>). Merely, three NMR structures exist for full-length  $\alpha$ -synuclein. Out of these, Protein Data Bank (PDB) codes 1XQ8<sup>43</sup> and 2KKW<sup>44</sup> are micelle-bound  $\alpha$ -synuclein structures. Recently, third NMR structure (PDB 2N0A)<sup>45</sup> was reported for a pathogenic fibril of full-length human  $\alpha$ -synuclein. We have used chain A of PDBs 2N0A and 2KKW for our study. Structures for the mutants were generated by computationally mutating the corresponding amino acid residue in the wild type NMR structure (PDB 2N0A)<sup>45</sup>. Mutagenesis Wizard of PyMOL (The PyMOL Molecular Graphics System, Schrödinger, LLC) was used for creating the point mutation. The resulting mutant structure was energy-minimized prior to MD simulation. Residue Ser-129 was phosphorylated using “Build and Edit Protein” module of the Discovery Studio 2017 R2 (BIOVIA, San Diego: Dassault Systèmes).

**Molecular dynamics simulations.** All the MD simulations were performed using GROMACS 5.1.3 package<sup>46</sup> with CHARMM27 force field<sup>47</sup>. Chain A of PDBs 2N0A<sup>45</sup> and 2KKW<sup>44</sup> were employed as the starting point for the simulations. MD simulation were performed at neutral and low pH (protonated His, Asp and Glu)<sup>5,48,49</sup>. High temperature is reported to accelerate unfolding and folding process without affecting the overall pathway<sup>5,50</sup>. Accordingly, simulations were carried out at high temperature (498 K) besides 300 K. Simulation were performed with periodic boundary conditions at constant temperature, pressure and fixed protonation states. A cubic box with TIP3P water molecules<sup>51</sup> was used to solvate the protein. Depending upon the protonation state of protein, appropriate number of counter-ions ( $\text{Na}^+$  and  $\text{Cl}^-$ ) were added to the system to maintain neutrality. System was then minimized using 50,000 steps of steepest descent algorithm to ensure that the system has no steric clashes, inappropriate geometry or and structure distortions. This was followed by 100 ps NVT (constant number of particles, volume, and temperature) and 1000 ps NPT (constant number of particles, pressure and temperature) equilibrations. Temperature was maintained using the v-rescale (modified Berendsen thermostat)<sup>52</sup> whereas pressure was maintained at 1 atm using the Parrinello-Rahman barostat<sup>53</sup>. The LINCS algorithm<sup>54</sup> was employed to constrain bond lengths. Long-range electrostatic interactions were treated with particle mesh Ewald summation method<sup>55</sup> and a cutoff of 12 Å was used for short-range interactions. Production runs were performed for 20 ns. The integration time step was set to 2 fs and the trajectory coordinates and energies were saved at 10 ps intervals. Built-in programs of GROMACS software package<sup>46</sup> were employed for the analysis.

**Computation of dihedral angles.** The dihedral angles were obtained using gmx rama module of GROMACS software package<sup>46</sup>. This module selects the  $\varphi/\psi$  dihedral combinations from topology file and computes these as a function of time. Afterwards, in house R scripts were used to identify residues which satisfies  $\alpha_R$  and  $\alpha_L$  conformations criteria as well as number of such conformations. A residue was classified in  $\alpha_R$  conformation if its dihedral angles were within  $-180 < \varphi < 0$ ,  $-180 < \psi < 0$  i.e. (–, –) while, if the dihedral angles were within  $0 < \varphi < 180$ ,  $0 < \psi < 180$  i.e. (+, +), it was categorized in  $\alpha_L$  conformation. This is in accordance with a previous report on  $\alpha$ -strand<sup>4</sup>.

**Topological water network analysis.** Water molecules linked by hydrogen bonds form different types of hydrogen-bonded cyclic water-ring networks. Potential functions considered here involve a rigid TIP3P water model. The interactions between water molecules are commonly modeled with Lennard-Jones and Coulomb

Simulated system	Type	Counter ions	Temperature	pH	MD run
1	2N0A, WT	9 Na <sup>+</sup>	300 K	Neutral	20 ns
2	2N0A, WT	9 Na <sup>+</sup>	498 K	Neutral	20 ns
3	2N0A, WT	16 Cl <sup>-</sup>	300 K	Low	20 ns
4	2N0A, WT	16 Cl <sup>-</sup>	498 K	Low	20 ns
5	2N0A, Mutant (A30P)	9 Na <sup>+</sup>	300 K	Neutral	20 ns
6	2N0A, Mutant (A53T)	9 Na <sup>+</sup>	300 K	Neutral	20 ns
7	2N0A, Mutant (E46K)	7 Na <sup>+</sup>	300 K	Neutral	20 ns
8	2N0A, Mutant (A30P/A56P/A76P)	9 Na <sup>+</sup>	300 K	Neutral	20 ns
9	2N0A, Mutant (G73P)	9 Na <sup>+</sup>	300 K	Neutral	20 ns
10	2N0A, Mutant S129E	10 Na <sup>+</sup>	300 K	Neutral	20 ns
11	2N0A, Ser(P)-129	10 Na <sup>+</sup>	300 K	Neutral	20 ns
12	2N0A, ACE-N-terminal	10 Na <sup>+</sup>	498 K	Neutral	20 ns
13	2KKW, WT	9 Na <sup>+</sup>	498 K	Neutral	20 ns
14	2KKW, WT	16 Cl <sup>-</sup>	498 K	Low	20 ns

**Table 1.** MD simulations carried out in the current study. The aggregated simulation time is 280 ns.

potentials<sup>51</sup>. The interaction potential energy between water molecules *a* and *b* is expressed by the following equation:

$$v(a, b) = \sum_i^{on} \sum_j^{aon} \frac{b q_i q_j e^2}{r_{ij}} + \frac{A}{r_{oo}^{12}} - \frac{C}{r_{oo}^6}$$

Where  $r_{oo}$  denotes the distance between oxygen atoms. The electrostatic attraction is expressed as a coulombic force between two charges  $q_i$  and  $q_j$  with a distance  $r_{ij}$  between them. The van der Waals interaction is expressed as a function that simultaneously takes attraction and repulsion. Parameter *A* is the repulsive force of *i* and *j* while parameter *C* represents attraction and is an inherent value according to the type of the atom. The parameters *A* and *C* were selected to produce reasonable structural and energetic results for liquid water. The value of parameters are as follows:

$$\begin{aligned} A &= 582,000 \text{ kcal } \text{\AA}^{12} \text{ mol}^{-1} \\ C &= 595 \text{ kcal } \text{\AA}^6 \text{ mol}^{-1} \\ q_i &= -0.834e, q_j = 0.417e \end{aligned}$$

We selected the energy criterion of  $-2.25 \text{ kcal mol}^{-1}$  to determine the hydrogen bond between water molecules as this value correlates with the minimum value of the pair-energy distribution of potential<sup>51</sup>.

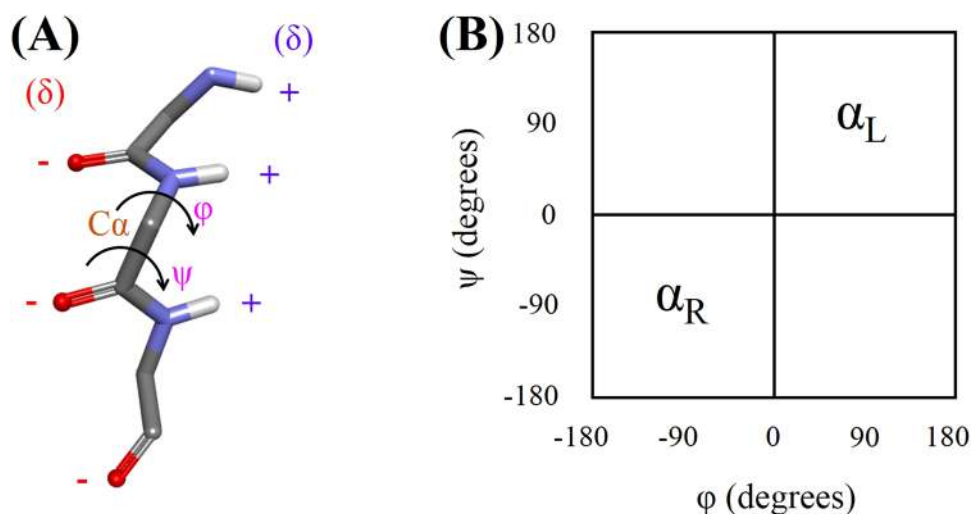
## Results and Discussion

$\alpha$ -synuclein is intrinsically disordered (natively unfolded) protein but how it is transformed into the highly organized fibrils is not well understood. Experimental studies demonstrated that  $\alpha$ -synuclein fibrillogenesis occurs in the similar way as  $\beta$ -amyloid (A $\beta$ ) fibrillogenesis by a multistep, nucleation-dependent polymerization process<sup>56</sup>. It requires seeding by an ordered nucleus, followed by the growth of oligomers through incorporation of further monomers. The first step involves primary nucleation that is the formation of a fibril nucleus from monomers. The formed nucleus grows during the fibril elongation step to produce mature fibrils, which subsequently catalyze the generation of new aggregates by secondary pathways. Generation of new aggregates by secondary pathways (fibril fragmentation or surface-catalyzed nucleation) are dependent on the concentration of existing fibrils. Secondary pathways lead to the formation of a new fibril nucleus which dissociates mature fibril in the final step and participate in polymerization process. Primary nucleation is a slow process while nucleation through secondary pathways is much more rapid<sup>57–59</sup>.

Here, we have explored the possibility of  $\alpha$ -strand formation in  $\alpha$ -synuclein and the apparent role of  $\alpha$ -strand/sheet in the aggregation and fibrillation of this protein. PDBs 2N0A (fibril state) and 2KKW (micelle-bound state) which represent two different conformations of  $\alpha$ -synuclein were selected for this study. To examine the  $\alpha$ -strand formation, we performed MD simulations of wild type (WT)  $\alpha$ -synuclein monomer at neutral and low pH at physiological (300 K) and high temperature (498 K). Furthermore, simulations of A30P, E46K, A53T, A30P/A56P/A76P and G73P variants were carried out to investigate the effect of these mutations on the  $\alpha$ -strand formation. Besides, phosphorylation effect was explored by phosphorylation-mimicking mutation S129E as well as by phosphorylating Ser-129<sup>39</sup>. Table 1 summarize the MD simulations performed in the present study. Residue sequences for the simulated systems are displayed in the Fig. 1. We also characterized the probable reasons for the  $\alpha$ -strand formation. At the end, aggregation mechanism of  $\alpha$ -synuclein on the basis of  $\alpha$ -strand/sheet structure is proposed.

<b>Residue sequence: WT</b>				<b>Residue sequence: G73P</b>							
1	MDVFMKGLSKA	KEGVVAAA	EKTQGV	AEEAGK	35	1	MDVFMKGLSKA	KEGVVAAA	EKTQGV	AEEAGK	35
36	GVLYVGSKT	KEGVVHG	VATVAE	KTKQVT	70	36	GVLYVGSKT	KEGVVHG	VATVAE	KTKQVT	70
71	VTGVTAVA	QKTVEG	AGSIAA	TGFVKD	105	71	VTGVTAVA	QKTVEG	AGSIAA	TGFVKD	105
106	GAPQEGILE	DMPVDP	DNEAYE	MPSEEG	140	106	GAPQEGILE	DMPVDP	DNEAYE	MPSEEG	140
<b>Residue sequence: A30P</b>				<b>Residue sequence: S129E</b>							
1	MDVFMKGLSKA	KEGVVAAA	EKTQGV	AEEA	35	1	MDVFMKGLSKA	KEGVVAAA	EKTQGV	AEEA	35
36	GVLYVGSKT	KEGVVHG	VATVAE	KTKQVT	70	36	GVLYVGSKT	KEGVVHG	VATVAE	KTKQVT	70
71	VTGVTAVA	QKTVEG	AGSIAA	TGFVKD	105	71	VTGVTAVA	QKTVEG	AGSIAA	TGFVKD	105
106	GAPQEGILE	DMPVDP	DNEAYE	MPSEEG	140	106	GAPQEGILE	DMPVDP	DNEAYE	MPSEEG	140
<b>Residue sequence: E46K</b>				<b>Residue sequence: A30P/A56P/A76P</b>							
1	MDVFMKGLSKA	KEGVVAAA	EKTQGV	AEEA	35	1	MDVFMKGLSKA	KEGVVAAA	EKTQGV	AEEA	35
36	GVLYVGSKT	KEGVVHG	VATVAE	KTKQVT	70	36	GVLYVGSKT	KEGVVHG	VATVAE	KTKQVT	70
71	VTGVTAVA	QKTVEG	AGSIAA	TGFVKD	105	71	VTGVTAVA	QKTVEG	AGSIAA	TGFVKD	105
106	GAPQEGILE	DMPVDP	DNEAYE	MPSEEG	140	106	GAPQEGILE	DMPVDP	DNEAYE	MPSEEG	140
<b>Residue sequence: A53T</b>				<b>Residue sequence: ACE-N-terminal</b>							
1	MDVFMKGLSKA	KEGVVAAA	EKTQGV	AEEA	35	1	ACE-MDVFMKGLSKA	KEGVVAAA	EKTQGV	AEEA	35
36	GVLYVGSKT	KEGVVHG	VATVAE	KTKQVT	70	36	GVLYVGSKT	KEGVVHG	VATVAE	KTKQVT	70
71	VTGVTAVA	QKTVEG	AGSIAA	TGFVKD	105	71	VTGVTAVA	QKTVEG	AGSIAA	TGFVKD	105
106	GAPQEGILE	DMPVDP	DNEAYE	MPSEEG	140	106	GAPQEGILE	DMPVDP	DNEAYE	MPSEEG	140

**Figure 1.** Residue sequences for the simulated systems.



**Figure 2.** (A) The  $\alpha$ -strand structure displaying the backbone dihedral angles and arrangement of the partial charges. Side chain and non-polar hydrogen atoms of the residues are not displayed for clarity. Phi ( $\phi$ ) is C-N-C $\alpha$ -C dihedral angle while Psi ( $\psi$ ) is N-C $\alpha$ -C-N dihedral angle. (B) Ramachandran or ( $\phi, \psi$ ) plot showing the right-handed ( $\alpha_R$ ) and left-handed ( $\alpha_L$ ) helical regions.

**Definition of the  $\alpha$ -strand structure.** The  $\alpha$ -strand is defined by an alternation of amino acid residues in helical,  $\alpha_R$  and  $\alpha_L$  conformations. Bifurcated hydrogen bonding between two nearby  $\alpha$ -strands forms an  $\alpha$ -sheet. Previous studies suggested various average ( $\phi, \psi$ ) for the  $\alpha$ -sheet structures<sup>4,6</sup>. In the present study, we analyzed the occurrence of  $\alpha$ -strand structure in  $\alpha$ -synuclein. Herein, successive residues were considered to adopt  $\alpha$ -strand conformation if they exist in  $\alpha_R\alpha_L\alpha_R$  or  $\alpha_L\alpha_R\alpha_L$  conformations. In an  $\alpha$ -strand, all the backbone carbonyl groups are oriented in the same direction on one side of the strand whereas all the backbone amino groups are oriented in the same direction on the reverse side of the strand. As a result, one edge of the  $\alpha$ -strand is negatively charged while the opposite edge is positively charged. Figure 2 shows the  $\alpha$ -strand structure, backbone dihedral angles, arrangement of the partial charges and Ramachandran or ( $\phi, \psi$ ) plot for right-handed ( $\alpha_R$ ) and left-handed ( $\alpha_L$ ) helical regions.

**Occurrence of  $\alpha$ -strand in the region 72–74.** Amyloidogenic NAC region is crucial for aggregation and cytotoxicity of  $\alpha$ -synuclein. This region is known to fold into a  $\beta$ -sheet secondary structure<sup>60</sup>. A hydrophobic stretch of 12 amino acids (residues 71–82) in the middle of NAC region is vital for filament assembly. Aggregation and neurotoxicity of  $\alpha$ -synuclein was abolished on deletion of residues 71–82. Moreover, when



isolated from rest of the structure, this 12 amino acid segment (residues 71–82) self-polymerize to form highly toxic amyloid fibrils and promote fibrillization of full-length human  $\alpha$ -synuclein *in vitro*. Interestingly, highly homologous human  $\beta$ -synuclein which lacks residues 74–84 does not aggregate *in vitro* and not found in Lewy bodies<sup>26–30</sup>. Besides, another closely related human  $\gamma$ -synuclein possess this critical region and forms fibrils like  $\alpha$ -synuclein<sup>61</sup>. These results indicate the significance of putative 12-residue core domain within the NAC region in fibril formation. A research group compared the *in vitro* aggregation properties of  $\alpha$ -,  $\beta$ -, and  $\gamma$ -synuclein<sup>61</sup>. They found that  $\gamma$ -synuclein fibril formation occurs by a nucleation-dependent mechanism as in  $\alpha$ -synuclein. However,  $\gamma$ -synuclein is intrinsically less fibrillogenic and fibril formation takes longer time than  $\alpha$ -synuclein. On the other hand,  $\beta$ -synuclein completely failed to fibrillize even under accelerated conditions for several weeks. They concluded that region between amino acids 72 and 84, which is lacking in  $\beta$ - but not  $\gamma$ -synuclein, may contain critical fibrillation determinators<sup>61</sup>.

In our WT simulations on 2N0A, an  $\alpha$ -strand ( $\alpha_R\alpha_L\alpha_R$ ) structure was observed in the region 72–74 for systems **2** (2N0A, WT, Neutral pH, 498 K) and **4** (2N0A, WT, Low pH, 498 K) which were simulated at high temperature. For determining  $\alpha$ -strand conformation, dihedral angles were monitored throughout the simulation and plotted. Variation in the dihedral angles of residues 72–74 for the systems **2** and **4** is plotted in Supplementary Figs S1 and S2, respectively. Occupancy of  $\alpha$ -strand during 20 ns simulation was found to be 14.84 and 10.85% for systems **2** and **4**, respectively. The position of  $\alpha$ -strand within sequence is correlated with the most amyloidogenic regions as determined experimentally<sup>26–30</sup>. Furthermore, these results are in agreement with previous experimental study which investigated the effect of pH and temperature on the human  $\alpha$ -synuclein structure and fibrillation<sup>31</sup>. A reversible transition of  $\alpha$ -synuclein into a partially folded intermediate was induced by either a reduction in pH or an elevation of temperature. The protein was partially folded under the low pH or high temperature conditions. Changes in the protein environment affect its charge and hydrophobic properties. Low pH minimizes the high net charge of this protein, thereby reduce intramolecular charge-charge repulsion and permit hydrophobic-driven collapse to a partially folded intermediate. High temperature provide a stronger hydrophobic driving force for folding.

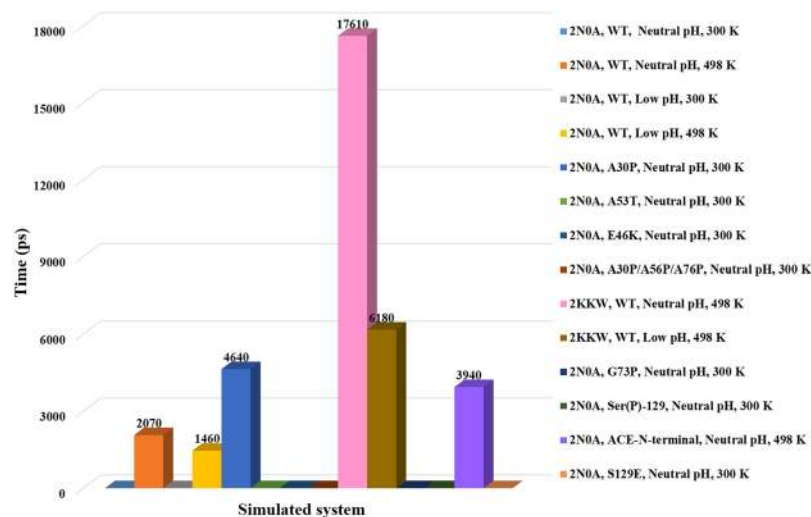
The  $\alpha$ -strand structure (region 72–74) was not observed for system **1** (2N0A, WT, Neutral pH, 300 K). This is supported by former studies which suggested that such structures are rare under normal physiological conditions<sup>4,5</sup>. Absence of this strand in system **3** (2N0A, WT, Low pH, 300 K) could be due to short simulation time (20 ns). Familial point mutation namely A30P, A53T and E46K which are located in the N-terminal region are reported to influence misfolding, aggregation and fibrillation of  $\alpha$ -synuclein<sup>62–65</sup>. However, effect of these PD-related point mutations on the folding behavior of this protein is unclear. Experimental studies revealed that A30P, A53T and E46K exerts strong, modest and subtle influences on local structural propensity, respectively<sup>66,67</sup>. We performed MD simulations on these mutants to inspect their influence on the  $\alpha$ -strand formation. Similar to WT systems **2** and **4**, mutant system **5** (2N0A, A30P, Neutral pH, 300 K) showed  $\alpha$ -strand ( $\alpha_R\alpha_L\alpha_R$ ) conformation for the residues 72–74. Furthermore, it was found that  $\alpha$ -strand conformation was extended to residue 75 in presence of A30P mutation. Consequently, a four-residue  $\alpha$ -strand ( $\alpha_R\alpha_L\alpha_R\alpha_L$ ) conformation was observed with high occupancy of 44.28% during 20 ns MD simulation. Variation in the dihedral angles of residues 72–75 for the system **5** is plotted in Supplementary Fig. S3. The  $\alpha$ -strand structure couldn't be observed for A53T and E46K in the short simulation time of 20 ns.

Mutations involving the NAC region such as A30P/A56P/A76P inhibit fibril formation<sup>38</sup>. We performed MD simulation on system **8** (2N0A, A30P/A56P/A76P, Neutral pH, 300 K) to explore the effect of this triple proline mutation on the  $\alpha$ -strand formation. During 20 ns simulation, no  $\alpha$ -strand structure was observed in the region 72–74 for system **8**. Additionally, we mutated central  $\alpha$ -strand residue from Gly73 to proline (G73P) in system **9** (2N0A, G73P, Neutral pH, 300 K). This system also failed to show  $\alpha$ -strand structure in the region 72–74. These results demonstrate that occurrence of  $\alpha$ -strand within the NAC region (residue 72–74) has substantial importance in the aggregation of  $\alpha$ -synuclein.

Phosphorylation of Ser-129 is reported to inhibit fibril formation<sup>39</sup>. We carried out simulations on phosphorylation-mimicking mutation S129E system **10** (2N0A, S129E, Neutral pH, 300 K) and phosphorylated Ser-129 system **11** (2N0A, Ser(P)-129, Neutral pH, 300 K). *In vitro* studies demonstrated that the phosphomimic S129E do not reproduce the effect of phosphorylation on the structural and aggregation properties of  $\alpha$ -synuclein and exhibits similar aggregation properties as the WT<sup>39</sup>. Our results are in agreement with the *in vitro* evaluations as both phosphomimic system **10** (2N0A, S129E, Neutral pH, 300 K) and WT system **1** (2N0A, WT, Neutral pH, 300 K) which were simulated under similar conditions didn't show  $\alpha$ -strand structure (region 72–74) during 20 ns MD simulations. Besides, no appearance of  $\alpha$ -strand structure in the phosphorylated system **11** (2N0A, Ser(P)-129, Neutral pH, 300 K) indicates that phosphorylation of Ser-129 suppress the assembly of  $\alpha$ -synuclein.

A recent paper described that amino-terminally acetylated  $\alpha$ -synuclein adopts compact conformations under physiological cell conditions. Such conformations shield aggregation-prone residues of NAC region from exposure to the cytoplasm and counteracts aggregation<sup>68</sup>. Our simulation on N-terminal acetylated system **12** (2N0A, ACE-N-terminal, Neutral pH, 498 K) also demonstrated that acetylation of N-terminal significantly reduced the number of  $\alpha$ -strand structures. Occupancy of  $\alpha$ -strand (region 72–74) during 20 ns simulation was found to be 2.50% for system **12**. Variation in the dihedral angles of residues 72–74 for the system **12** is plotted in Supplementary Fig. S4. Under the similar simulation conditions, non-acetylated system **2** (2N0A, WT, Neutral pH, 498 K) showed 14.84% occupancy. This demonstrate that N-terminal acetylation slows down the fibrillation.

Similar to 2N0A, simulations on 2KKW also showed  $\alpha$ -strand ( $\alpha_R\alpha_L\alpha_R$ ) structure in the region 72–74 for systems **13** (2KKW, WT, Neutral pH, 498 K) and **14** (2KKW, WT, Low pH, 498 K). However, 2KKW systems showed comparatively lower occupancy of  $\alpha$ -strand than 2N0A systems. Occupancy of  $\alpha$ -strand during 20 ns simulation was found to be 0.50 and 9.20% for systems **13** and **14**, respectively. Variation in the dihedral angles of residues 72–74 for the systems **13** and **14** is plotted in Supplementary Figs S5 and S6, respectively. Lower occupancy of



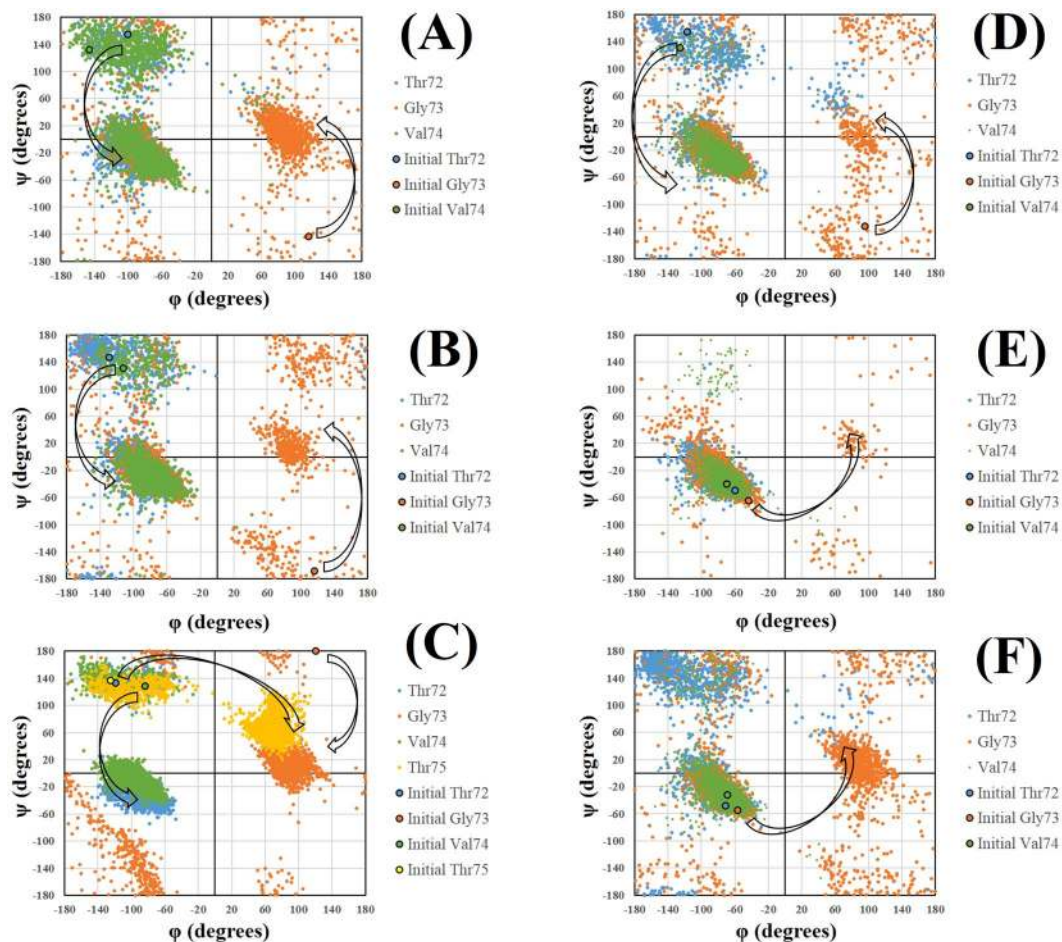
**Figure 3.** Comparison of  $\alpha$ -strand appearance time in the region 72–74 for the simulated systems.

2KKW (micelle-bound state) systems as compared to 2N0A (fibril state) systems indicates that fibril state produce relatively more  $\alpha$ -strand structures.

**$\alpha$ -strand appearance time.** We examined the  $\alpha$ -strand appearance time for the simulated systems (Fig. 3). As discussed before,  $\alpha$ -strand was not observed for systems 1 (2N0A, WT, Neutral pH, 300 K), 3 (2N0A, WT, Low pH, 300 K), 6 (2N0A, A53T, Neutral pH, 300 K), 7 (E46K, Neutral pH, 300 K), 8 (2N0A, A30P/A56P/A76P, Neutral pH, 300 K), 9 (2N0A, G73P, Neutral pH, 300 K), 10 (2N0A, S129E, Neutral pH, 300 K) and 11 (2N0A, Ser(P)-129, Neutral pH, 300 K) in the region 72–74. Simulation time corresponding to the first appearance of  $\alpha$ -strand structure in the region 72–74 was compared for systems 2 (2N0A, WT, Neutral pH, 498 K), 4 (2N0A, WT, Low pH, 498 K), 5 (2N0A, A30P, Neutral pH, 300 K), 12 (2N0A, ACE-N-terminal, Neutral pH, 498 K), 13 (2KKW, WT, Neutral pH, 498 K) and 14 (2KKW, WT, Low pH, 498 K). MD simulation results demonstrated high temperature was important factor for the  $\alpha$ -strand formation in WT  $\alpha$ -synuclein. In case of systems 2 (neutral pH) and 4 (low pH)  $\alpha$ -strand first appeared at 2070 and 1460 ps, respectively. The  $\alpha$ -strand formation occurred faster at low pH as compared to neutral pH. This is in agreement with pH-induced partial unfolding of amyloid proteins leading to  $\alpha$ -sheet structure<sup>4,48</sup>. First  $\alpha$ -strand appeared at relatively longer simulation time (4640 ps) for mutant system 5 (A30P), as this system was simulated at neutral pH and normal temperature. N-terminal acetylated system 12 showed first  $\alpha$ -strand at 3940 ps. Under the similar conditions, WT system 2 showed first  $\alpha$ -strand in slightly shorter time (2070 ps). This is in accordance with a previous *in vitro* study which reported that N-terminal acetylation delays the fibril assembly of  $\alpha$ -synuclein<sup>68</sup>. The first  $\alpha$ -strand appeared at 17610 and 6180 ps for systems 13 and 14, respectively. Delayed appearance of  $\alpha$ -strand in 2KKW systems as compared to 2N0A systems suggested that fibril state conformation (2N0A) folds more rapidly into intermediate  $\alpha$ -strand structure as compared to the micelle-bound single conformation (2KKW). This is supported by the nucleation-dependent polymerization process which stated that fibril-dependent secondary pathways are much more rapid than primary nucleation process where a fibril nucleus is formed from the monomers<sup>56–59</sup>.

**Transition from  $\beta$ -sheet to  $\alpha$ -helix region.** Ramachandran plots for 2N0A systems in Fig. 4A–D demonstrate that residues 72 and 74 underwent transition from  $\beta$ -sheet region to the right-handed  $\alpha$ -helix region for the systems 2, 4, 5 and 12. Besides, residue 73 displayed transition to left-handed  $\alpha$ -helix region from its diverse initial positions. Furthermore, residue 75 showed transition from  $\beta$ -sheet region to left-handed  $\alpha$ -helix region in case of mutant system 5 (Fig. 4C). Ramachandran plots for 2KKW systems in Fig. 4E,F show the transition of residue 73 from right-handed to left-handed  $\alpha$ -helix region. Sequential transition of backbone dihedral angles of the residues produced the  $\alpha$ -strand structure.

**Peptide-plane flip.** As presented in Fig. 4, transition to  $\alpha$ -strand region happens sequentially *via* individual transition of backbone dihedral angles instead of all at once in a concerted manner. Such transitions involve rotation of peptide-plane with slight change in side chain orientation and can occur at both physiological and high temperature. However, transition is far slower at physiological temperature and need a longer simulation time<sup>4</sup>. Our results also support the statement as  $\alpha$ -strand was observed predominantly for high temperature simulations during 20 ns MD simulation. Peptide-plane phenomenon has been explained in detail elsewhere<sup>6,69</sup>. Peptide-plane flips back if the environment is nonconductive to the transition<sup>4</sup>. In compliance with previous studies<sup>4,6,7</sup>,  $\alpha$ -strand formed in our simulation was very dynamic as the peptide-plane flipped back and forth throughout the course of the simulation. As a result of flipping, carbonyl and amino groups are aligned in the  $\alpha$ -strand and two complementary partially charged interfaces are formed. In our study,  $\alpha$ -strand appeared in the region 72–74 for systems 2 (2N0A, WT, Neutral pH, 498 K), 4 (2N0A, WT, Low pH, 498 K), 5 (2N0A, A30P, Neutral pH, 300 K), 12 (2N0A, ACE-N-terminal, Neutral pH, 498 K), 13 (2KKW, WT, Neutral pH, 498 K) and 14 (2KKW, WT, Low



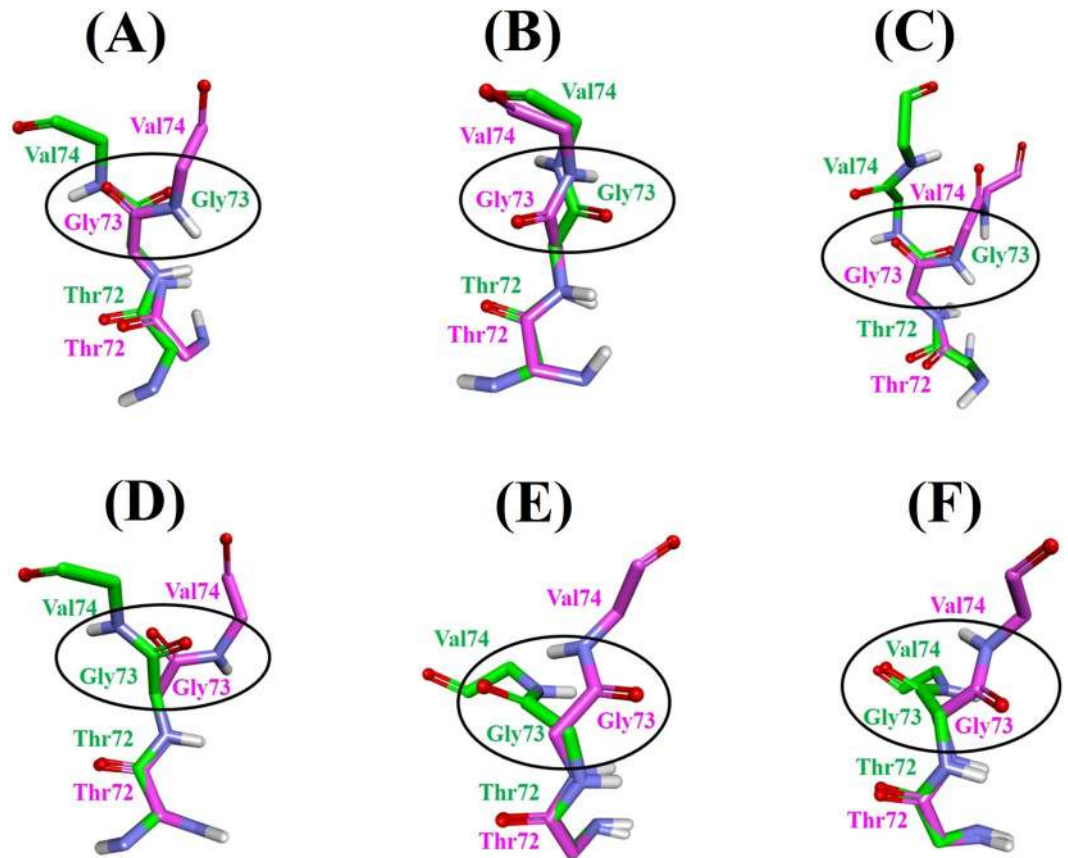
**Figure 4.** Distribution of dihedral angles for systems 2 (A), 4 (B), 5 (C), 12 (D), 13 (E) and 14 (F) in the region 72–74. Transition of  $\phi$  and  $\psi$  angles for  $\alpha$ -strand formation. Initial refers to the dihedrals of particular residue at the beginning of the production run.

pH, 498 K). For each system, we superimposed the first  $\alpha$ -strand and initial conformation at the beginning of the production run to visualize the peptide plane flip. As can be seen in Fig. 5, peptide-plane of Gly73–Val74 underwent flipping in all the  $\alpha$ -strand ( $\alpha_R\alpha_L\alpha_R$ ) forming systems. Additionally, rotation of Val74–Thr75 peptide-plane in system 5 (2N0A, A30P, Neutral pH, 300 K) lead to the extended  $\alpha$ -strand ( $\alpha_R\alpha_L\alpha_R\alpha_L$ ) formation.

**Evidence for  $\alpha$ -strand consisting of Thr, Gly and Val segment in the native proteins.** In MD results, we observed  $\alpha$ -strand structure for residues 72–74 which corresponds to Thr, Gly and Val amino acids, respectively. Since these are computational results, it is important to know whether such structure exist in experimentally determined protein structures. We searched PDB in order to find available structures where segment comprising of Thr, Gly and Val amino acids are present in  $\alpha_R\alpha_L\alpha_R$  conformation, respectively. Similar to our simulation results, there are 322 PDBs in this database that exhibit  $\alpha$ -strand structures. Moreover, in these 322 PDBs, there are 637 occurrences of  $\alpha$ -strands. These PDBs support the  $\alpha$ -strand ( $\alpha_R\alpha_L\alpha_R$ ) conformation displayed by Thr, Gly and Val segment (residues 72–74) during MD simulation of  $\alpha$ -synuclein. The  $\alpha$ -strand structure in the region 72–74 was not observed in any available PDBs of  $\alpha$ -synuclein. Hence, present work offers novel structural insights about this amyloid protein. Such information could be useful in understanding its aggregation mechanism.

**Proposed aggregation mechanism of  $\alpha$ -synuclein.** Traditional experimental methods often cannot unveil protein folding and unfolding routes or intermediate states. By contrast, computational methods enable the observation of such events and play crucial roles in probing the intermediates. MD simulation is a powerful computational tool and has been extensively used for several years to study the structural and dynamic properties of the proteins. MD simulations have shed an interesting light on the early events of protein folding and aggregation. Unfortunately, MD simulations are computationally intensive and thus often limited to pico-to-nano second timescales depending on the size of the system and computing resources. Protein folding or unfolding may require longer simulation time. Nonphysical temperatures have been suggested to accelerate such transitions<sup>5,50</sup>. Nevertheless, different conformations can be observed in the short besides the long time simulations.

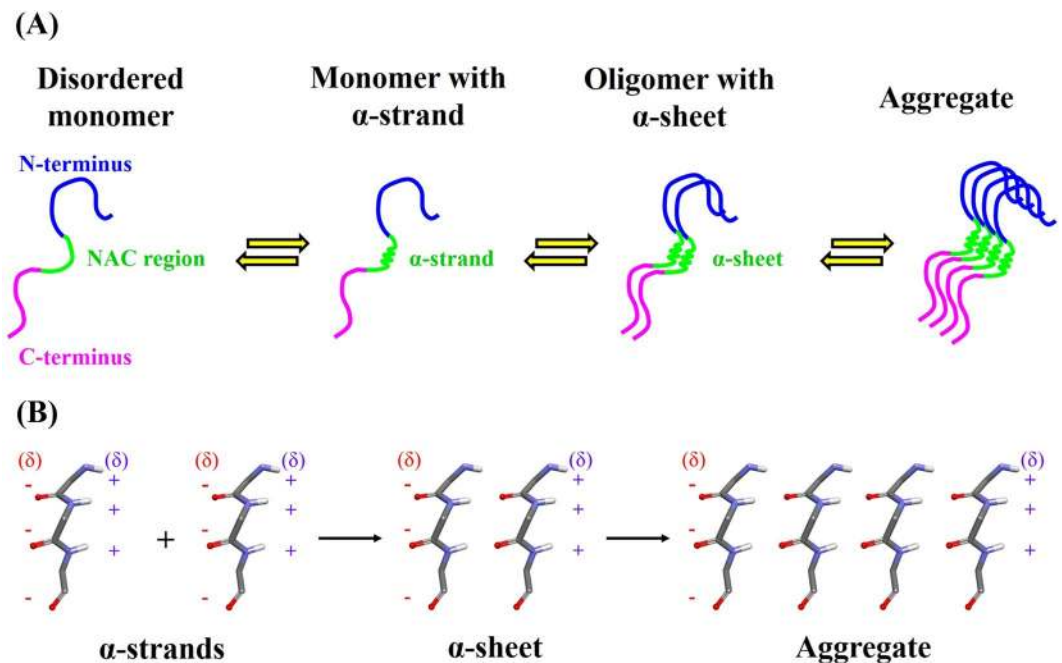




**Figure 5.** Peptide plane flipping. (A) System 2 (2N0A, WT, Neutral pH, 498 K), (B) System 4 (2N0A, WT, Low pH, 498 K), (C) System 5 (2N0A, A30P, Neutral pH, 300 K), (D) System 12 (2N0A, ACE-N-terminal, Neutral pH, 498 K), (E) System 13 (2KKW, WT, Neutral pH, 498 K) and (F) System 14 (2KKW, WT, Low pH, 498 K). The main chain atoms of the  $\alpha$ -strand residues are displayed whereas side chain and non-polar hydrogen atoms are not displayed for clarity. First  $\alpha$ -strand conformation is represented by pink stick model while initial conformation at the beginning of the production run is shown by green stick model. Flip of CO-NH plane is highlighted.

Due to limited computing facilities, we performed MD simulations on  $\alpha$ -synuclein monomer for 20 ns and investigated the effects of the selected factors. This simulation time is rather short by current standards as such simulations can suffer from inadequate conformational sampling. Only  $\alpha$ -synuclein monomers were simulated and not the dimers as our lab lacks high performance computational resources. Beyond these limitations, our simulations revealed  $\alpha$ -strand formation in the critical NAC region (residues 72–74) of  $\alpha$ -synuclein. We propose that  $\alpha$ -strand structure is possibly involved in the aggregation of  $\alpha$ -synuclein. Figure 6A demonstrate how an  $\alpha$ -strand structure may act as an intermediate and as an initiation site for oligomerization. This hypothesis is supported by various experimental observations<sup>26–31,42</sup>. However, further experimental studies are required to confirm the role of  $\alpha$ -strand/sheet in aggregation and fibrillation of  $\alpha$ -synuclein. A number of experimental investigations have provided evidence that  $\alpha$ -synuclein exist as an intrinsically disordered monomer<sup>70</sup>. At physiological conditions,  $\alpha$ -synuclein monomers remain as disordered structures. However, under typical conditions such as elevated temperature or low pH or point mutations, region 72–74 of monomers occasionally adopts an  $\alpha$ -strand conformation. The  $\alpha$ -strands of monomers interact through hydrogen bonding to form oligomers with  $\alpha$ -sheet structures. Typical feature of the  $\alpha$ -sheet is the alignment of carbonyl and amino groups. Partial charges due to this alignment generate two complimentary charged interfaces, as shown in Fig. 6B. The  $\alpha$ -sheet structures facilitate self-assembly/aggregation of  $\alpha$ -synuclein into soluble protofibrils as a result of attractive forces between the interfaces with opposite charges. Subsequently, transition from  $\alpha$ -sheet to  $\beta$ -sheet *via* peptide plane flip causes transformation of soluble cytotoxic protofibrils to insoluble highly ordered amyloid fibrils. Besides,  $\alpha$ -sheet may proceed to the mature fibrils. This hypothesis is in agreement with the previous studies which reported a similar theory for various amyloid proteins such as transthyretin,  $\beta_2$ -microglobulin, lysozyme and prion protein<sup>4,5</sup>. Recently, cryo-electron microscopy (cryo-EM) structure of  $\alpha$ -synuclein fibrils was published<sup>71</sup>. In the three-dimensional structure, each  $\alpha$ -synuclein(1–121) molecule consisted of eight in-register parallel  $\beta$ -strands. This includes residues 42–46 ( $\beta$ 1), 48–49 ( $\beta$ 2), 52–57 ( $\beta$ 3), 59–66 ( $\beta$ 4), 69–72 ( $\beta$ 5), 77–82 ( $\beta$ 6), 89–92 ( $\beta$ 7), and 94–(102) ( $\beta$ 8). Hydrophobic residues including Val74 contributed to the stability of the protofilament<sup>71</sup>. Furthermore, it was suggested that the initial binding event of fibril elongation might be interactions involving residues 74–82. These results further support our findings that residue 72–74 are crucial for fibril formation.





**Figure 6.** Probable role of  $\alpha$ -strand in the aggregation of  $\alpha$ -synuclein. **(A)** Suggested stages of aggregation process. N-terminus, NAC region and C-terminus are highlighted in blue, green and pink, respectively. Biophysical studies have confirmed that isolated  $\alpha$ -synuclein monomer is intrinsically disordered *in vitro* and aggregation results in the formation of amyloid fibrils. Occasionally, region 72–74 of  $\alpha$ -synuclein monomer adopts an  $\alpha$ -strand conformation. The  $\alpha$ -strands of  $\alpha$ -synuclein oligomers interact through hydrogen bonding to form  $\alpha$ -sheet structure. Subsequently,  $\alpha$ -sheet acts as a possible nucleus that initiate aggregation and fibrillation of  $\alpha$ -synuclein. **(B)** A model for  $\alpha$ -sheet intermediate where partial charges on the interface are indicated by red (negative charges) and blue (positive charges) colors.

**Strategy for designing  $\alpha$ -synuclein aggregation inhibitors.** Comprehensive knowledge of the structure and function of a biological target is essential for its inhibitor design. The  $\alpha$ -synuclein lacks a well-defined native structure and its aggregation mechanism is ambiguous. X-ray fiber diffraction suggested that insoluble amyloid fibrils are comprised of cross  $\beta$ -sheet structure<sup>4,72</sup>. It is presumed that fibril formation involves a transition from amyloidogenic intermediate to  $\beta$ -sheet structure that is stabilized by intermolecular interactions<sup>4</sup>. Such protein-protein interactions generally involve shallow surfaces and lack well-defined ligand-binding pocket<sup>73</sup>. Accordingly, development of small drug-like inhibitors which could perhaps prevent or reverse the formation of amyloid fibrils is challenging. Hence, researchers are now focusing on peptide-based drug discovery. This includes rational design of short synthetic peptides capable of inhibiting  $\alpha$ -synuclein aggregation and design of  $\beta$ -sheet breakers which can bind and isolate  $\beta$ -sheet precursors<sup>74,75</sup>. Nowadays, short peptides can be rapidly and economically synthesized by chemical means and can be modified to improve binding properties. Degradation of peptides is anticipated to be less toxic as compared to synthetic small organic molecules. Various research groups are currently involved in designing peptides which possess the ability to bind to a specific region of  $\alpha$ -synuclein and disrupt its aggregation<sup>75–77</sup>.

In the present study, we have proposed an  $\alpha$ -synuclein aggregation mechanism on the basis of the  $\alpha$ -sheet/strand structure. Suggested hypothesis is supported by previous experimental and modeling studies<sup>4,5,26–31,42</sup>. According to the proposed mechanism, initially  $\alpha$ -strand formation occurs in the NAC region (residue 72–74) of  $\alpha$ -synuclein monomer. This strand is likely to form  $\alpha$ -sheet structure in the oligomers. Afterwards,  $\alpha$ -sheet acts as a promising nucleus that initiate and promote aggregation and fibrillation of  $\alpha$ -synuclein. We think that  $\alpha$ -strand/sheet formation within the NAC region has substantial importance and could be exploited for the rational design of PD drugs. Recently, various anti- $\alpha$ -sheet peptides were reported to inhibit amyloid fibril formation and its associated cytotoxicity<sup>8–11</sup>. Herein, we introduce  $\alpha$ -strand/sheet forming region (residue 72–74) of  $\alpha$ -synuclein that could be targeted by designed anti- $\alpha$ -sheet peptides. This approach may serve as a novel promising neuro-protective strategy that could aid alleviate PD and its symptoms.

**Topological water network analysis.** Proteins are strongly influenced by the surrounding water. Experiments as well as computer simulations revealed that water has a substantial impact on structure, stability, dynamics and function of a protein. Several theoretical studies have explored the effect of water molecules on the structural properties of proteins. However, role of water in the protein folding or unfolding transition mechanism is not well-known<sup>78</sup>. Our research group has previously predicted kinase selectivity on the basis of water-ring network analysis<sup>79,80</sup>. Here, we examined the role of the water network in the  $\alpha$ -strand formation within the NAC region of  $\alpha$ -synuclein. We focused on the small water-ring network namely three-membered water-ring for the

System	Non- $\alpha$ -strand structure						$\alpha$ -strand structure				
	TWN	Thr72	Gly73	Val74	Thr75	Total (N)	Thr72	Gly73	Val74	Thr75	Total (N)
System 2 (2N0A, WT, Neutral pH, 498 K)	15 Å	15.20	14.75	14.52		44.47	13.26	13.16	13.78		40.20
	20 Å	44.23	43.67	43.24		131.14	40.49	40.65	42.16		123.30
System 4 (2N0A, WT, Low pH, 498 K)	15 Å	12.93	12.53	13.02		38.48	13.18	13.75	13.96		40.89
	20 Å	40.66	40.47	41.38		122.51	41.00	42.41	43.48		126.89
System 5 (2N0A, A30P, Neutral pH, 300 K)	15 Å	8.34	8.62	7.75	7.05	31.76	7.18	7.53	6.69	6.16	27.56
	20 Å	25.28	25.49	23.89	22.96	97.62	23.51	24.06	22.23	21.82	91.62
System 12 (2N0A, ACE-N-terminal, Neutral pH, 498 K)	15 Å	4.72	5.02	5.12		14.86	3.66	3.41	3.31		10.38
	20 Å	13.66	13.82	13.80		41.28	11.67	12.23	12.17		36.07
System 13 (2KKW, WT, Neutral pH, 300 K)	15 Å	4.75	3.75	4.25		12.75	3.31	2.92	3.30		9.53
	20 Å	10.67	11.25	11.25		33.17	10.50	10.33	10.75		31.58
System 14 (2KKW, WT, Low pH, 300 K)	15 Å	4.51	4.24	3.97		12.72	4.92	4.38	4.01		13.31
	20 Å	12.89	12.55	12.41		37.85	13.78	13.67	13.54		40.99

**Table 2.** Comparison of the average number of three-membered water-rings in the absence and presence of the  $\alpha$ -strand structure.

topological water network (TWN) analysis. The water molecules within 15 and 20 Å radius sphere of  $\alpha$ -strand residues were obtained from MD trajectory. We calculated the number of three-membered water rings around the  $\alpha$ -strand residues and subsequently compared the simulated systems. TWN analysis results are summarized in Table 2. It is evident from the table that total average number of three-membered water rings ( $N$ ) formed around  $\alpha$ -strand structures were comparatively lower than the non- $\alpha$ -strand structures for the neutral pH simulations. The  $N$  values of non- $\alpha$ -strand structures for the system 2 (2N0A, WT, Neutral pH, 498 K) within 15 and 20 Å of residues 72–74 were 44.47 and 131.14, respectively. On the other hand, for the same system,  $\alpha$ -strand structures showed  $N$  values of 40.20 and 123.30 within 15 and 20 Å of region 72–74, respectively. Similarly,  $N$  values of non- $\alpha$ -strand and  $\alpha$ -strand structures for the system 5 (2N0A, A30P, Neutral pH, 300 K) were 31.76 and 27.56, respectively, within 15 Å of residues 72–75 while  $N$  values were 97.62 and 91.62, respectively, within 20 Å radius. The  $N$  values of non- $\alpha$ -strand structures for the system 12 (2N0A, ACE-N-terminal, Neutral pH, 498 K) within 15 and 20 Å of residues 72–74 were 14.86 and 41.28, respectively. For the same system,  $\alpha$ -strand structures showed  $N$  values of 10.38 and 36.07 within 15 and 20 Å of region 72–74, respectively. The  $N$  values of non- $\alpha$ -strand and  $\alpha$ -strand structures for the system 13 (2KKW, WT, Neutral pH, 300 K) were 12.75 and 9.53, respectively, within 15 Å of residues 72–74 whereas  $N$  values were 33.17 and 31.58, respectively, within 20 Å radius.

On the contrary, for the low pH simulation,  $\alpha$ -strand structures showed relatively higher  $N$  values as compared to the non- $\alpha$ -strand structures. The  $N$  values of  $\alpha$ -strand structures for the system 4 (2N0A, WT, Low pH, 498 K) within 15 and 20 Å of residues 72–74 were 40.89 and 126.89, respectively. However, non- $\alpha$ -strand structures for the same system showed  $N$  values of 38.48 and 122.51 within 15 and 20 Å of region 72–74, respectively. The  $N$  values of non- $\alpha$ -strand and  $\alpha$ -strand structures for the system 14 (2KKW, WT, Low pH, 300 K) were 12.72 and 13.31, respectively, within 15 Å of residues 72–74 whereas  $N$  values were 37.85 and 40.99, respectively, within 20 Å radius. Protonation state of the amino acid residues is effected at the low pH that could influence the hydrogen-bonded cyclic water-ring networks. As discussed in the methodology section, His, Asp and Glu residues were protonated for low pH simulation. Accordingly, overall charge on the  $\alpha$ -synuclein was affected (Table 1). This could be the probable reason for the dissimilarity in the TWN analysis for neutral and low pH simulations. To get more insight into exactly how surrounding water molecules contribute to the  $\alpha$ -strand formation, further studies are advised. Overall, our TWN analysis demonstrated the influence of water network on the occurrence of  $\alpha$ -strand in NAC region of the  $\alpha$ -synuclein and suggested that water molecules play an active role in the protein folding or unfolding process.

## Conclusion

Various experiments provide evidence that a partially folded intermediate is responsible for  $\alpha$ -synuclein fibril formation under low pH and high temperature conditions. Besides, previous studies on amyloid proteins such as lysozyme, prion proteins,  $\beta$ 2-microglobulin and transthyretin stated that  $\alpha$ -strand/sheet may represent the intermediate as such structure nicely explains many experimental observations. The  $\alpha$ -synuclein is also an amyloid protein that forms fibrils. However,  $\alpha$ -strand/sheet structure is not yet reported for  $\alpha$ -synuclein. We investigated the occurrence of  $\alpha$ -strand/sheet structure in this amyloid protein through all-atom MD simulations under different conditions. Interestingly,  $\alpha$ -strand structure was observed for residue 72–74 of  $\alpha$ -synuclein monomer. These residues are the part of 12 amino acid segment (residues 71–82) within the NAC region of  $\alpha$ -synuclein. This region plays a critical role in both the aggregation and cytotoxicity of  $\alpha$ -synuclein. Moreover, when isolated from rest of the structure, this segment (residues 71–82) self-polymerize to form highly toxic amyloid fibrils and promote fibrillization of full-length human  $\alpha$ -synuclein *in vitro*. Aggregation and neurotoxicity of  $\alpha$ -synuclein is abolished on deletion of residues 71–82. Accordingly, occurrence of  $\alpha$ -strand within the NAC region (residue 72–74) of  $\alpha$ -synuclein has substantial importance. Furthermore, TWN analysis around  $\alpha$ -strand residues indicated that existence of such structure is influenced by surrounding water network. Based on the  $\alpha$ -strand structure in the region 72–74, we proposed a plausible aggregation mechanism for  $\alpha$ -synuclein that is supported

by former experimental findings. According to the suggested mechanism, occurrence of  $\alpha$ -strand in the monomer lead to  $\alpha$ -sheet formation in the oligomers which initiates and promote aggregation and fibrillation of  $\alpha$ -synuclein. Hence, crucial region 72–74 could be targeted by designed anti- $\alpha$ -sheet peptides. Such peptides have been found to inhibit aggregation of various amyloid proteins. This promising strategy could inhibit amyloid fibril formation in  $\alpha$ -synuclein and could be beneficial for the therapeutic treatment of PD.

## Data Availability

No datasets were generated or analyzed during the current study.

## References

- Pauling, L. & Corey, R. B. The pleated sheet, a new layer configuration of polypeptide chains. *Proc. Natl. Acad. Sci. USA* **37**, 251–256 (1951).
- Pauling, L. & Corey, R. B. The structure of feather rachis keratin. *Proc. Natl. Acad. Sci. USA* **37**, 256–261 (1951).
- Pauling, L. & Corey, R. B. Configurations of polypeptide chains with favored orientations around single bonds two new pleated sheets. *Proc. Natl. Acad. Sci. USA* **37**, 729–740 (1951).
- Daggett, V.  $\alpha$ -Sheet: the toxic conformer in amyloid diseases? *Acc. Chem. Res.* **39**, 594–602 (2006).
- Armen, R. S., DeMarco, M. L., Alonso, D. O. & Daggett, V. Pauling and Corey's  $\alpha$ -pleated sheet structure may define the prefibrillar amyloidogenic intermediate in amyloid disease. *Proc. Natl. Acad. Sci. USA* **101**, 11622–11627 (2004).
- Milner-White, J. E., Watson, J. D., Qi, G. & Hayward, S. Amyloid formation may involve  $\alpha$ -to  $\beta$  sheet interconversion via peptide plane flipping. *Structure* **14**, 1369–1376 (2006).
- Yang, M., Lei, M., Yordanov, B. & Huo, S. Peptide plane can flip in two opposite directions: Implication in amyloid formation of transthyretin. *J. Phys. Chem. B* **110**, 5829–5833 (2006).
- Grillo-Bosch, D. *et al.* Retro-Enantio N-Methylated Peptides as  $\beta$ -Amyloid Aggregation Inhibitors. *ChemMedChem* **4**, 1488–1494 (2009).
- Hopping, G. *et al.* Designed  $\alpha$ -sheet peptides inhibit amyloid formation by targeting toxic oligomers. *Elife* **3**, <https://doi.org/10.7554/eLife.01681> (2014).
- Maris, N. L., Shea, D., Bleem, A., Bryers, J. D. & Daggett, V. Chemical and physical variability in structural isomers of an L/D  $\alpha$ -sheet peptide designed to inhibit amyloidogenesis. *Biochemistry* **57**, 507–510 (2017).
- Bleem, A., Francisco, R., Bryers, J. D. & Daggett, V. Designed  $\alpha$ -sheet peptides suppress amyloid formation in *Staphylococcus aureus* biofilms. *NPJ Biofilms Microbiomes* **3**, 16, <https://doi.org/10.1038/s41522-017-0025-2> (2017).
- Uversky, V. N. & Eliezer, D. Biophysics of Parkinson's disease: structure and aggregation of  $\alpha$ -synuclein. *Curr. Protein Pept. Sci.* **10**, 483–499 (2009).
- Trojanowski, J. Q., Goedert, M., Iwatsubo, T. & Lee, V. M. Fatal attractions: abnormal protein aggregation and neuron death in Parkinson's disease and Lewy body dementia. *Cell Death Differ.* **5**, 832–837 (1998).
- Takeda, A. *et al.* Abnormal accumulation of NACP/alpha-synuclein in neurodegenerative disorders. *Am. J. Pathol.* **152**, 367–372 (1998).
- Lücking, C. & Brice, A. Alpha-synuclein and Parkinson's disease. *Cell. Mol. Life Sci.* **57**, 1894–1908 (2000).
- Spillantini, M. G. *et al.*  $\alpha$ -Synuclein in Lewy bodies. *Nature* **388**, 839–840 (1997).
- Spillantini, M. G. *et al.* Filamentous  $\alpha$ -synuclein inclusions link multiple system atrophy with Parkinson's disease and dementia with Lewy bodies. *Neurosci. Lett.* **251**, 205–208 (1998).
- Wakabayashi, K., Yoshimoto, M., Tsuji, S. & Takahashi, H.  $\alpha$ -Synuclein immunoreactivity in glial cytoplasmic inclusions in multiple system atrophy. *Neurosci. Lett.* **249**, 180–182 (1998).
- Gai, W., Power, J., Blumbergs, P. & Blessing, W. Multiple-system atrophy: a new  $\alpha$ -synuclein disease? *Lancet* **352**, 547–548 (1998).
- Arawaka, S., Saito, Y., Murayama, S. & Mori, H. Lewy body in neurodegeneration with brain iron accumulation type 1 is immunoreactive for  $\alpha$ -synuclein. *Neurology* **51**, 887–889 (1998).
- Wakabayashi, K., Matsumoto, K., Takayama, K., Yoshimoto, M. & Takahashi, H. NACP, a presynaptic protein, immunoreactivity in Lewy bodies in Parkinson's disease. *Neurosci. Lett.* **239**, 45–48 (1997).
- 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.* **143**, 243142, <https://doi.org/10.1063/1.4936910> (2015).
- Rodriguez, J. A. *et al.* Structure of the toxic core of  $\alpha$ -synuclein from invisible crystals. *Nature* **525**, 486–490 (2015).
- Vilar, M. *et al.* The fold of  $\alpha$ -synuclein fibrils. *Proc. Natl. Acad. Sci. USA* **105**, 8637–8642 (2008).
- Miake, H., Mizusawa, H., Iwatsubo, T. & Hasegawa, M. Biochemical characterization of the core structure of  $\alpha$ -synuclein filaments. *J. Biol. Chem.* **277**, 19213–19219 (2002).
- Giasson, B. I., Murray, I. V., Trojanowski, J. Q. & Lee, V. M.-Y. A hydrophobic stretch of 12 amino acid residues in the middle of  $\alpha$ -synuclein is essential for filament assembly. *J. Biol. Chem.* **276**, 2380–2386 (2001).
- Periquet, M., Fulga, T., Myllykangas, L., Schlossmacher, M. G. & Feany, M. B. Aggregated  $\alpha$ -synuclein mediates dopaminergic neurotoxicity *in vivo*. *J. Neurosci.* **27**, 3338–3346 (2007).
- El-Agnaf, O. *et al.* Aggregates from mutant and wild-type  $\alpha$ -synuclein proteins and NAC peptide induce apoptotic cell death in human neuroblastoma cells by formation of  $\beta$ -sheet and amyloid-like filaments. *FEBS Lett.* **440**, 71–75 (1998).
- Han, H., Weinreb, P. H. & Lansbury, P. T. Jr. The core Alzheimer's peptide NAC forms amyloid fibrils which seed and are seeded by  $\beta$ -amyloid: is NAC a common trigger or target in neurodegenerative disease? *Chem. Biol.* **2**, 163–169 (1995).
- Biere, A. L. *et al.* Parkinson's disease-associated  $\alpha$ -synuclein is more fibrillogenic than  $\beta$ - and  $\gamma$ -synuclein and cannot cross-seed its homologs. *J. Biol. Chem.* **275**, 34574–34579 (2000).
- Uversky, V. N., Li, J. & Fink, A. L. Evidence for a partially folded intermediate in  $\alpha$ -synuclein fibril formation. *J. Biol. Chem.* **276**, 10737–10744 (2001).
- Li, J., Uversky, V. N. & Fink, A. L. Effect of familial Parkinson's disease point mutations A30P and A53T on the structural properties, aggregation, and fibrillation of human  $\alpha$ -synuclein. *Biochemistry* **40**, 11604–11613 (2001).
- Greenbaum, E. A. *et al.* The E46K mutation in  $\alpha$ -synuclein increases amyloid fibril formation. *J. Biol. Chem.* **280**, 7800–7807 (2005).
- Narhi, L. *et al.* Both familial Parkinson's disease mutations accelerate  $\alpha$ -synuclein aggregation. *J. Biol. Chem.* **274**, 9843–9846 (1999).
- Conway, K. *et al.* Accelerated oligomerization by Parkinson's disease linked  $\alpha$ -synuclein mutants. *Ann. N. Y. Acad. Sci.* **920**, 42–45 (2000).
- Uversky, V. N., Li, J. & Fink, A. L. Metal-triggered structural transformations, aggregation, and fibrillation of human  $\alpha$ -synuclein a possible molecular link between parkinson's disease and heavy metal exposure. *J. Biol. Chem.* **276**, 44284–44296 (2001).
- Munishkina, L. A., Henriques, J., Uversky, V. N. & Fink, A. L. Role of protein–water interactions and electrostatics in  $\alpha$ -synuclein fibril formation. *Biochemistry* **43**, 3289–3300 (2004).
- Karpinar, D. P. *et al.* Pre-fibrillar  $\alpha$ -synuclein variants with impaired  $\beta$ -structure increase neurotoxicity in Parkinson's disease models. *EMBO J.* **28**, 3256–3268 (2009).
- Paleologou, K. E. *et al.* Phosphorylation at Ser-129 but not the phosphomimics S129E/D inhibits the fibrillation of  $\alpha$ -synuclein. *J. Biol. Chem.* **283**, 16895–16905 (2008).

40. Kang, L., Janowska, M. K., Moriarty, G. M. & Baum, J. Mechanistic insight into the relationship between N-terminal acetylation of  $\alpha$ -synuclein and fibril formation rates by NMR and fluorescence. *PLoS One* **8**, e75018, <https://doi.org/10.1371/journal.pone.0075018> (2013).
41. Narkiewicz, J., Giachin, G. & Legname, G. *In vitro* aggregation assays for the characterization of  $\alpha$ -synuclein prion-like properties. *Prion* **8**, 19–32 (2014).
42. Fink, A. L. The aggregation and fibrillation of  $\alpha$ -synuclein. *Acc. Chem. Res.* **39**, 628–634 (2006).
43. Ulmer, T. S., Bax, A., Cole, N. B. & Nussbaum, R. L. Structure and dynamics of micelle-bound human  $\alpha$ -synuclein. *J. Biol. Chem.* **280**, 9595–9603 (2005).
44. Rao, J. N., Jao, C. C., Hegde, B. G., Langen, R. & Ulmer, T. S. A combinatorial NMR and EPR approach for evaluating the structural ensemble of partially folded proteins. *J. Am. Chem. Soc.* **132**, 8657–8668 (2010).
45. Tuttle, M. D. *et al.* Solid-state NMR structure of a pathogenic fibril of full-length human  $\alpha$ -synuclein. *Nat. Struct. Mol. Biol.* **23**, 409–415 (2016).
46. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1**, 19–25 (2015).
47. Bjelkmar, P., Larsson, P., Cuendet, M. A., Hess, B. & Lindahl, E. Implementation of the CHARMM force field in GROMACS: analysis of protein stability effects from correction maps, virtual interaction sites, and water models. *J. Chem. Theory Comput.* **6**, 459–466 (2010).
48. Armen, R. S., Alonso, D. O. & Daggett, V. Anatomy of an amyloidogenic intermediate: conversion of  $\beta$ -sheet to  $\alpha$ -sheet structure in transthyretin at acidic pH. *Structure* **12**, 1847–1863 (2004).
49. Steward, R. E., Armen, R. S. & Daggett, V. Different disease-causing mutations in transthyretin trigger the same conformational conversion. *Protein Eng. Des. Sel.* **21**, 187–195 (2008).
50. Day, R., Bennion, B. J., Ham, S. & Daggett, V. Increasing temperature accelerates protein unfolding without changing the pathway of unfolding. *J. Mol. Biol.* **322**, 189–203 (2002).
51. 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.* **79**, 926–935 (1983).
52. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101, <https://doi.org/10.1063/1.2408420> (2007).
53. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
54. Hess, B., Bekker, H., Berendsen, H. J. & Fraaije, J. G. LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).
55. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
56. Wood, S. J. *et al.*  $\alpha$ -Synuclein fibrillogenesis is nucleation-dependent implications for the pathogenesis of Parkinson's disease. *J. Biol. Chem.* **274**, 19509–19512 (1999).
57. Buell, A. K. *et al.* Solution conditions determine the relative importance of nucleation and growth processes in  $\alpha$ -synuclein aggregation. *Proc. Natl. Acad. Sci. USA* **111**, 7671–7676 (2014).
58. Aleksis, R., Oleskovs, F., Jaudzems, K., Pahnke, J. & Biverstål, H. Structural studies of amyloid- $\beta$  peptides: Unlocking the mechanism of aggregation and the associated toxicity. *Biochimie* **140**, 176–192 (2017).
59. Cohen, S. I. *et al.* Proliferation of amyloid- $\beta$ 42 aggregates occurs through a secondary nucleation mechanism. *Proc. Natl. Acad. Sci. USA* **110**, 9758–9763 (2013).
60. Gallegos, S., Pacheco, C., Peters, C., Opazo, C. M. & Aguayo, L. G. Features of alpha-synuclein that could explain the progression and irreversibility of Parkinson's disease. *Front. Neurosci.* **9**, 59, <https://doi.org/10.3389/fnins.2015.00059> (2015).
61. Biere, A. L. *et al.* Parkinson's disease associated  $\alpha$ -synuclein is more fibrillogenic than  $\beta$ - and  $\gamma$ -Synuclein and cannot cross-seed its homologs. *J. Biol. Chem.* **275**, 34574–34579 (2000).
62. Krüger, R. *et al.* AlaSOPro mutation in the gene encoding  $\alpha$ -synuclein in Parkinson's disease. *Nature Genet.* **18**, 106–108 (1998).
63. Zarranz, J. J. *et al.* The new mutation, E46K, of  $\alpha$ -synuclein causes parkinson and Lewy body dementia. *Ann. Neurol.* **55**, 164–173 (2004).
64. Lemkau, L. R. *et al.* Site-specific perturbations of alpha-synuclein fibril structure by the Parkinson's disease associated mutations A53T and E46K. *PLoS One* **8**, e49750, <https://doi.org/10.1371/journal.pone.0049750> (2013).
65. Berrocal, R., Vasquez, V., Krs, S. R., Gadad, B. S. & Rao, K.  $\alpha$ -Synuclein misfolding versus aggregation relevance to Parkinson's disease: Critical assessment and modeling. *Mol. Neurobiol.* **51**, 1417–1431 (2015).
66. Bussell, R. & Eliez, D. Residual structure and dynamics in Parkinson's disease-associated mutants of  $\alpha$ -synuclein. *J. Biol. Chem.* **276**, 45996–46003 (2001).
67. Fredenburg, R. A. *et al.* The impact of the E46K mutation on the properties of  $\alpha$ -synuclein in its monomeric and oligomeric states. *Biochemistry* **46**, 7107–7118 (2007).
68. Theillet, F.-X. *et al.* Structural disorder of monomeric  $\alpha$ -synuclein persists in mammalian cells. *Nature* **530**, 45–50 (2016).
69. Hayward, S. Peptide-plane flipping in proteins. *Protein Sci.* **10**, 2219–2227 (2001).
70. Alderson, T. R. & Markley, J. L. Biophysical characterization of  $\alpha$ -synuclein and its controversial structure. *Intrinsically Disord. Proteins* **1**, 18–19 (2013).
71. Guerrero-Ferreira, R. *et al.* Cryo-EM structure of alpha-synuclein fibrils. *Elife* **7**, e36402, <https://doi.org/10.7554/eLife.36402> (2018).
72. Blake, C. & Serpell, L. Synchrotron X-ray studies suggest that the core of the transthyretin amyloid fibril is a continuous  $\beta$ -sheet helix. *Structure* **4**, 989–998 (1996).
73. Mason, J. M. Design and development of peptides and peptide mimetics as antagonists for therapeutic intervention. *Future Med. Chem.* **2**, 1813–1822 (2010).
74. Mason, J. M. & Fairlie, D. P. Toward peptide-based inhibitors as therapies for Parkinson's disease. *Future Med. Chem.* **7**, 2103–2105 (2015).
75. El-Agnaf, O. M. *et al.* A strategy for designing inhibitors of  $\alpha$ -synuclein aggregation and toxicity as a novel treatment for Parkinson's disease and related disorders. *FASEB J.* **18**, 1315–1317 (2004).
76. Madine, J., Doig, A. J. & Middleton, D. A. Design of an N-methylated peptide inhibitor of  $\alpha$ -synuclein aggregation guided by solid-state NMR. *J. Am. Chem. Soc.* **130**, 7873–7881 (2008).
77. Bodles, A. M., El-Agnaf, O. M., Greer, B., Guthrie, D. J. & Irvine, G. B. Inhibition of fibril formation and toxicity of a fragment of  $\alpha$ -synuclein by an N-methylated peptide analogue. *Neurosci. Lett.* **359**, 89–93 (2004).
78. Karvounis, G., Nerukh, D. & Glen, R. C. Water network dynamics at the critical moment of a peptide's  $\beta$ -turn formation: A molecular dynamics study. *J. Chem. Phys.* **121**, 4925–4935 (2004).
79. Jang, W. D., Lee, M. H. & Kang, N. S. Quantitative assessment of kinase selectivity based the water-ring network in protein binding sites using molecular dynamics simulations. *J. Mol. Liq.* **221**, 316–322 (2016).
80. Jang, W. D., Kim, J.-T. & Kang, N. S. The analysis of water network for kinase selectivity based on the MD simulations. *J. Mol. Liq.* **191**, 37–41 (2014).



## Acknowledgements

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Science, ICT & Future Planning (2017R1A2B4002827).

## Author Contributions

A.B. designed the work, performed the simulations, analyzed the results and wrote the manuscript. K.E.C. assisted in the analysis. N.S.K. guided and approved the research work.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-37276-1>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019