

# Structures of the E46K Mutant-Type $\alpha$ -Synuclein Protein and Impact of E46K Mutation on the Structures of the Wild-Type $\alpha$ -Synuclein Protein

Olivia Wise-Scira,<sup>†,||</sup> Aquila Dunn,<sup>†,||</sup> Ahmet K. Aloglu,<sup>†</sup> Isin T. Sakallioğlu,<sup>‡</sup> and Orkid Coskuner<sup>\*,†,§,||</sup>

<sup>†</sup>Department of Chemistry, The University of Texas at San Antonio, San Antonio, Texas 78249, United States

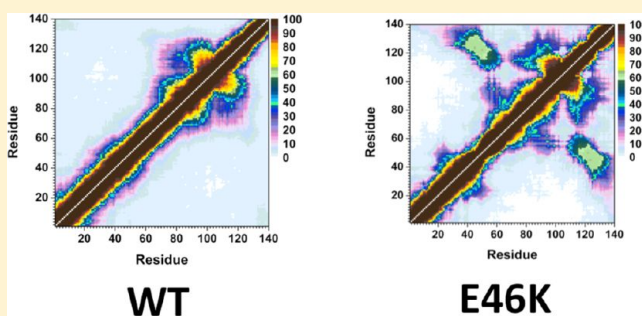
<sup>‡</sup>Department of Chemistry, Bilkent University, 06800 Bilkent, Ankara, Turkey

<sup>§</sup>Neurosciences Institute, The University of Texas at San Antonio, One UTSA Circle, San Antonio, Texas 78249, United States

## Supporting Information

**ABSTRACT:** The E46K genetic missense mutation of the wild-type  $\alpha$ -synuclein protein was recently identified in a family of Spanish origin with hereditary Parkinson's disease. Detailed understanding of the structures of the monomeric E46K mutant-type  $\alpha$ -synuclein protein as well as the impact of the E46K missense mutation on the conformations and free energy landscapes of the wild-type  $\alpha$ -synuclein are required for gaining insights into the pathogenic mechanism of Parkinson's disease. In this study, we use extensive parallel tempering molecular dynamics simulations along with thermodynamic calculations to assess the secondary and tertiary structural properties as well as the conformational preferences of the monomeric wild-type and E46K mutant-type  $\alpha$ -synuclein proteins in an aqueous solution environment. We also present the residual secondary structure component conversion stabilities with dynamics using a theoretical strategy, which we most recently developed. To the best of our knowledge, this study presents the first detailed comparison of the structural and thermodynamic properties of the wild-type and E46K mutant-type  $\alpha$ -synuclein proteins in an aqueous solution environment at the atomic level with dynamics. We find that the E46K mutation results not only in local but also in long-range changes in the structural properties of the wild-type  $\alpha$ -synuclein protein. The mutation site shows a significant decrease in helical content as well as a large increase in  $\beta$ -sheet structure formation upon E46K mutation. In addition, the  $\beta$ -sheet content of the C-terminal region increases significantly in the E46K mutant-type  $\alpha$ S in comparison to the wild-type  $\alpha$ S. Our theoretical strategy developed to assess the thermodynamic preference of secondary structure transitions indicates that this shift in secondary structure is the result of a decrease in the thermodynamic preference of turn to helix conversions while the coil to  $\beta$ -sheet preference increases for these residues. Long-range intramolecular protein interactions of the C-terminal with the N-terminal and NAC regions increase upon E46K mutation, resulting in more compact structures for the E46K mutant-type rather than wild-type  $\alpha$ S. However, the E46K mutant-type  $\alpha$ S structures are less stable than the wild-type  $\alpha$ S. Overall, our results show that the E46K mutant-type  $\alpha$ S has a higher propensity to aggregate than the wild-type  $\alpha$ S and that the N-terminal and C-terminal regions are reactive toward fibrillization and aggregation upon E46K mutation and we explain the associated reasons based on the structural properties herein. Small molecules or drugs that can block the specific residues forming abundant  $\beta$ -sheet structure, which we report here, might help to reduce the reactivity of these intrinsically disordered fibrillogenic proteins toward aggregation and their toxicity.

**KEYWORDS:**  $\alpha$ -synuclein, genetic missense mutation, free energy landscape, molecular dynamics



Parkinson's disease (PD) is a debilitating neurodegenerative disorder that affects approximately 1% of the population above age 65 and up to 5% of the population by age 85.<sup>1</sup> PD as well as some other neurodegenerative disorders are characterized by the formation of Lewy bodies, which are large proteinaceous inclusions composed mainly of the intrinsically disordered protein  $\alpha$ -synuclein ( $\alpha$ S).<sup>2–7</sup> In addition, some inherited forms of PD are associated with genetic missense or replication mutations of the  $\alpha$ S protein.<sup>8–11</sup> Recently, a genetic missense mutation of the 140 amino acid residue  $\alpha$ S protein resulting in the mutation of the 46th amino acid residue from

glutamic acid to lysine (E46K) was isolated in a family of Spanish origin.<sup>8</sup> Zarranz et al. linked this mutation to both familial PD and Lewy body dementia.<sup>8</sup> The aggregation of  $\alpha$ S is proposed to play a crucial role in the pathogenic mechanism of PD, especially the formation of the reportedly neurotoxic oligomeric species.<sup>12,13</sup> Furthermore, the monomeric and

Received: November 9, 2012

Accepted: January 2, 2013

Published: January 2, 2013



oligomeric structures reportedly determine the aggregation mechanism of  $\alpha$ S.<sup>14,15</sup>

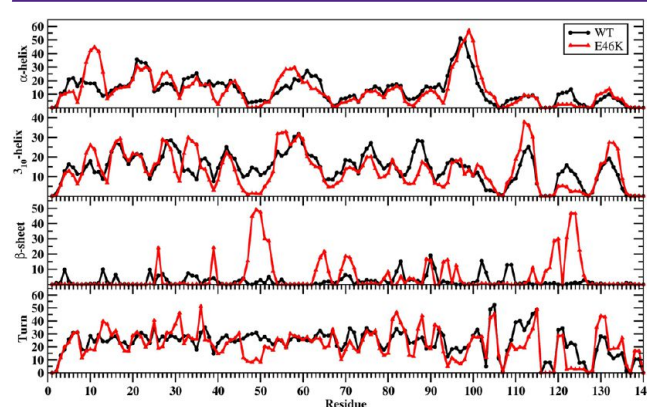
The E46K genetic missense mutation alters the aggregation mechanism and neurotoxicity of the wild-type  $\alpha$ S, but the structural properties and exact impact have been debated in the current literature. It was reported that the E46K mutant-type  $\alpha$ S is neurotoxic whereas the wild-type  $\alpha$ S is neuroprotective toward rat brains as well as for SH-SY5Y cells in the presence of environmental toxins.<sup>16</sup> Circular dichroism (CD), electron microscopy, and thioflavin T (ThT) fluorescence studies as well as sedimentation assays report that the E46K mutant-type  $\alpha$ S fibrillize more rapidly than the wild-type  $\alpha$ S.<sup>17–20</sup> However, Kamiyoshihara et al. reported that the E46K mutant-type  $\alpha$ S formed fibrils at a slower rate than the wild-type  $\alpha$ S even though the total number of fibrils formed is greater upon E46K mutation.<sup>21</sup> While some experimental studies report that the E46K mutant-type  $\alpha$ S fibril structure is morphologically similar to that of the wild-type  $\alpha$ S,<sup>17–19</sup> Choi et al. reported a more pronounced twisted appearance for the E46K mutant-type  $\alpha$ S in comparison to the wild-type  $\alpha$ S.<sup>20</sup> Understanding the monomeric and oligomeric wild- and mutant-type  $\alpha$ S structures as well as the effect of the E46K genetic missense mutation on these structures provides knowledge directly linked to the aggregation mechanism of  $\alpha$ S.<sup>18,22,23</sup> For example, specific secondary structure component formations, such as helix and  $\beta$ -sheet structures, in the structures of the monomeric intrinsically disordered fibrillogenic proteins are known to determine the fibrillization and aggregation mechanisms.<sup>18,24–26</sup> Gaining insights into these structural properties at the monomeric level can then be used for developing therapeutic agents that target the aggregation mechanism of  $\alpha$ S via blocking specific residues reactive toward aggregation. A detailed description of the dynamic structural properties of these species remains challenging with conventional experimental techniques due in part to complications provided by fast aggregation, rapid structural changes, and solvent effects. Despite, some experimental studies have reported the structural properties of the wild-type and E46K mutant-type  $\alpha$ S proteins. Regarding oligomeric structures, Fredenburg et al. reported that the formation of protofibrillar structures decreases upon E46K mutation of  $\alpha$ S but that the morphological characteristics of these structures are similar to those of its wild-type form.<sup>18</sup> Nuclear magnetic resonance (NMR) and CD studies reported that the secondary structure components of the free wild-type  $\alpha$ S protein are not greatly impacted upon E46K mutation.<sup>18,19,27,28</sup> In contrast, single molecule force (SMF) spectroscopy measurements by Brucale et al. indicated the  $\beta$ -sheet content of the E46K mutant-type  $\alpha$ S is significantly increased over the wild-type  $\alpha$ S.<sup>25</sup> NMR measurements also indicated that interactions between the N-terminal (Met1-Lys60) and C-terminal (Lys96-Ala140) regions are increased upon E46K mutation of the wild-type  $\alpha$ S.<sup>27</sup>

Theoretical studies can provide insight into the atomic level dynamic structural changes of the wild-type  $\alpha$ S upon E46K mutation at the monomeric level in solution that are not easily observable with conventional experimental tools. However, theoretical studies investigating the E46K mutant-type  $\alpha$ S are mostly lacking or limited in their contents in the currently existing literature. For example, Sundar and co-workers reported the aggregation propensity of the E46K mutant-type  $\alpha$ S is intermediate of the other two genetic missense mutations of  $\alpha$ S via aggregation propensity calculations on structures obtained from classical molecular dynamics (MD) simulations

in explicit water.<sup>29</sup> Nevertheless, the impact of the E46K mutation on the structural and thermodynamic properties of the wild-type  $\alpha$ S as well as sampling of conformations are not presented in their study. Perlmutter et al. also presented classical MD simulations of the wild-type and three genetic missense mutant  $\alpha$ S proteins in the presence of either an sodium dodecyl sulfate (SDS) micelle or a phospholipid bilayer.<sup>30</sup> Based on these simulations, they reported that the E46K mutation added a hydrogen bond between  $\alpha$ S and the lipid or micelle that may be related to the increased bilayer affinity of E46K mutant-type  $\alpha$ S.<sup>20,30</sup> To the best of our knowledge, a detailed study presenting the impact of the E46K mutation on the structural and thermodynamic properties of the wild-type  $\alpha$ S in the free state at atomic level with dynamics has not yet been presented in the literature. Therefore, we performed all-atom replica exchange molecular dynamics (REMD) simulations on the full-length monomeric wild-type and E46K mutant-type  $\alpha$ S proteins in aqueous solution. Using these simulated structures, we calculated the secondary and tertiary structural components as well as the thermodynamic properties of the monomeric wild-type and E46K mutant-type  $\alpha$ S in an aqueous solution environment. Furthermore, we used our recently developed new theoretical strategy for gaining knowledge about the residual secondary structure component conversions at the atomic level with dynamics. Based on these results, we find that the E46K mutation greatly impacts the structural and thermodynamic properties of the monomeric wild-type  $\alpha$ S in aqueous solution.

## RESULTS AND DISCUSSION

The calculated secondary structure components per residue of the wild-type and E46K mutant-type  $\alpha$ S are depicted in Figure 1. Helical and  $\beta$ -sheet structures are proposed to play a crucial



**Figure 1.** WT and E46K  $\alpha$ S secondary structure components. Secondary structure abundances per residue for the wild-type (black) and E46K mutant-type (red)  $\alpha$ S. The abundances for the  $\pi$ -helix and coil structures are not displayed.

role in the aggregation process of intrinsically disordered fibrillogenic proteins in neurodegenerative diseases, and therefore, these secondary structure components are of special interest. The most abundant  $\alpha$ -helix formation in the N-terminal region of the wild-type  $\alpha$ S occurs at Glu20-Gln24 with 20%–35% abundance for the wild-type  $\alpha$ S (Figure 1). Even though the abundance of  $\alpha$ -helix structure of Glu20-Gln24 is similar in the E46K mutant-type  $\alpha$ S, the most abundant  $\alpha$ -helix structure in the N-terminal region of the E46K mutant-type  $\alpha$ S occurs at Ser9-Glu13, which is at least 15% more abundant than

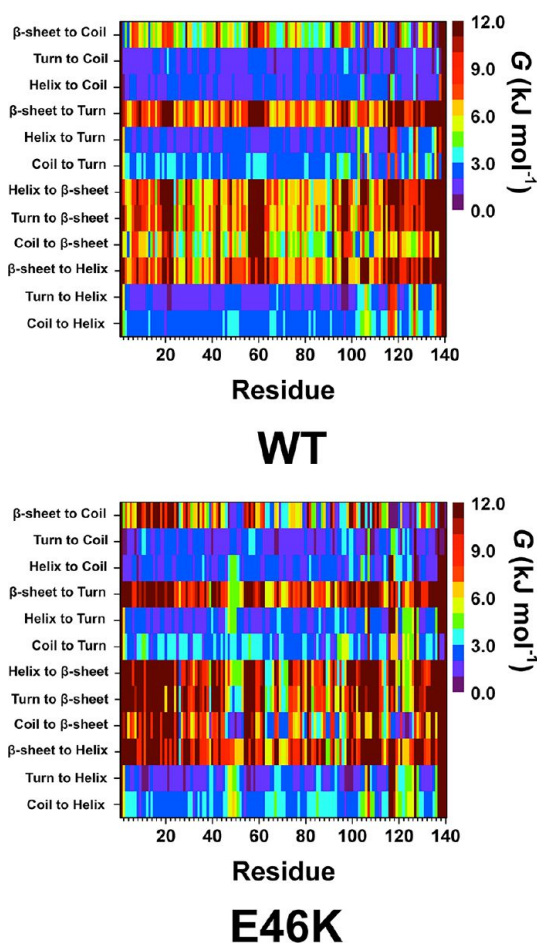
in the wild-type  $\alpha$ S. Furthermore, Ala27-Ala29 and Thr54-Thr59 also present an increase in  $\alpha$ -helical structure upon E46K mutation of the wild-type  $\alpha$ S. However, the vice versa trend is observed for residues Tyr39-Ser42. For the nonamyloid  $\beta$  component region (NAC; Glu61-Val95), residues Glu61-Thr64 of the wild-type  $\alpha$ S protein form abundant  $\alpha$ -helical structure. The formation of helical structure in the N-terminal and NAC regions is associated with lipid or vesicle binding of the  $\alpha$ S protein.<sup>31–34</sup> Therefore, differences in the  $\alpha$ -helix formation of the  $\alpha$ S protein as a result of the E46K mutation of the wild-type  $\alpha$ S may affect the binding of  $\alpha$ S with lipids or vesicles. Previous experimental studies reported that the E46K mutant-type  $\alpha$ S has a higher affinity for binding to negatively charged vesicles than the wild-type.<sup>20</sup> Our results support these experiments and further show that this reported higher affinity may be associated with the higher  $\alpha$ -helix content at Ser9-Glu13, Ala27-Ala29, and Thr54-Thr59 in the N-terminal region upon E46K mutation. Within the C-terminal region (Lys96-Ala140), the most prominent  $\alpha$ -helix formation appears at Lys96-Gly101 for both wild-type and E46K mutant-type  $\alpha$ S; however, the abundance at Gln99-Lys101 increases by up to 20% in E46K mutant-type  $\alpha$ S rather than in the wild-type  $\alpha$ S conformations. In addition, the  $\alpha$ -helical structure at Pro120-Glu123 almost disappears upon E46K mutation in the structures of the wild-type  $\alpha$ S protein.

Both the wild-type and E46K mutant-type  $\alpha$ S display abundant  $3_{10}$ -helix formation (>20% abundance) in the N-terminal region at Val15-Ala18, Glu20-Thr22, Ala27-Gly31, Gly41-Ser42, and Thr54-Lys60. Upon E46K mutation, we observe an increase in the  $3_{10}$ -helix formation at Ser9-Ala11, Lys32-Glu35, and Thr54-Ala56 by up to 15%. On the other side, the  $3_{10}$ -helix structure formation at residues Glu46-Gly51 almost disappears when Glu46 is replaced by Lys. In the NAC region, residues Val74-Val77 and Gly86-Ala89 of the wild-type  $\alpha$ S form abundant  $3_{10}$ -helix structure (>20%). Overall, the abundance of  $3_{10}$ -helix structure in the NAC region decreases upon E46K mutation except at residues Thr64, Asn65, Thr72, Thr81-Glu83, Ala90, and Ala91. Residues Ile112-Glu114 present the most abundant  $3_{10}$ -helix formation for both the wild-type and E46K mutant-type  $\alpha$ S; however, the abundance is slightly increased for the E46K mutant-type in comparison to the wild-type  $\alpha$ S. We also note an increase in  $3_{10}$ -helix formation at Gly132-Gln134 upon E46K mutation while the vice versa trend is observed for Pro120-Ala124 in the C-terminal region.

For the wild-type  $\alpha$ S, the most abundant  $\beta$ -sheet structures (up to 20%) are formed in parts of the NAC and C-terminal regions (Val70, Val71, Val82, Glu83, Ala89-Ala91, Lys102, Asn103, Pro108, and Gln109). We also observe  $\beta$ -sheet structure formation in the N-terminal region of the wild-type  $\alpha$ S at Phe4, Glu13, Val16, Gln24, Val26, Ala27, Thr33-Glu35, Val37, and Val52. Interestingly, the NAC and N-terminal residues (except Phe4) are located in regions considered part of the 11-mer repeats. There are seven regions of 11-mer repeats, sequences that assume right handed coiled coil conformations, in the N-terminal and NAC region of  $\alpha$ -synuclein.<sup>35</sup> It is suggested these repeats lower the propensity of  $\alpha$ -synuclein to form  $\beta$ -sheet due to their preference for  $\alpha$ -helix formation.<sup>35</sup> In comparison, the most prominent  $\beta$ -sheet structures (20%–50%) for the E46K mutant-type  $\alpha$ S are located in the N-terminal and C-terminal regions at residues Val26, Tyr39, Gly47-Ala53, Asp119, Pro120, and Asn122-Tyr125. In addition, Val63-Val66, Ala69-Thr72, Lys80, Gly93, Phe94, Lys96,

Glu114, Pro117, and Val118 in the NAC and C-terminal regions present  $\beta$ -sheet structure formation with 5%–20% abundances. Kessler et al. reported  $\alpha$ S was more fibrillogenic upon deletion of the N-terminal and C-terminal, thus implying the NAC region is more prone to aggregation.<sup>35</sup> In comparison with our results, the NAC region had minor  $\beta$ -sheet structure formation (5%–20% abundance), whereas the N-terminal and C-terminal exhibited a 20%–50%. Moreover, specific residues (Val37–Lys43, Val52–Thr59, Gln62–Val66, Gly68–Val77, and Ala90–Val95) were reported in NMR study conducted by Vilar et al. as the main regions to form  $\beta$ -sheet structures that compose the wild-type  $\alpha$ S fibrils.<sup>36</sup> The most abundant (up to 55%) turn structure formations occur in the C-terminal region of the wild-type but this is shifted to the N-terminal region in the E46K mutant-type  $\alpha$ S. Large discrepancies in the tendencies to form turn structure in the N-terminal region are noticed at Glu13, Gly14, Ala30, Gly31, Gly36, and Lys46-His50; in the NAC region at Val82, Ala85, and Gly86; and in the C-terminal region at Glu110-Ile112, Pro117, Val118, Asn122-Ala124, and Glu130, and Glu131 with a difference up to 20%.

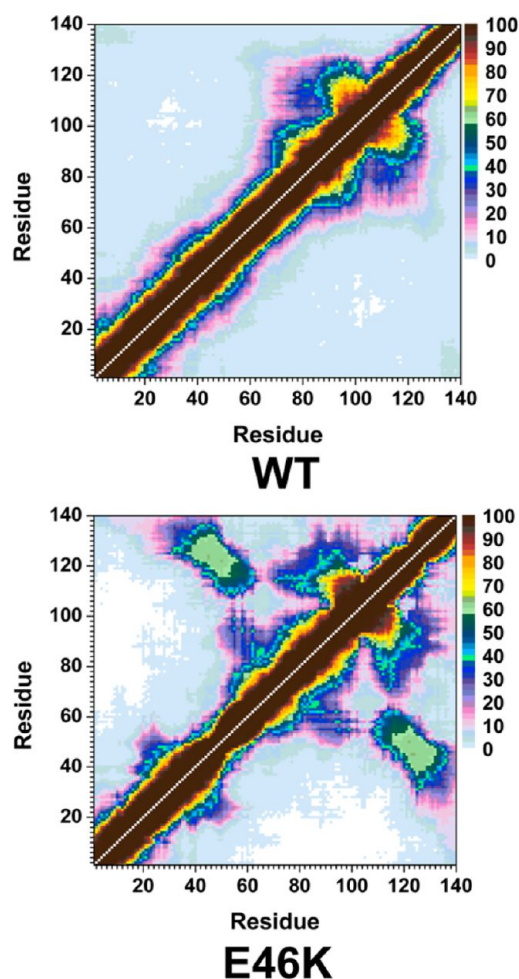
We also calculated the free energy change values associated with the transitions between two different secondary structure components for each residue of the wild-type and E46K mutant-type  $\alpha$ S in an aqueous solution medium (Figure 2) using a new theoretical strategy that we most recently developed.<sup>37</sup> These calculations overall indicate that turn or helix to coil transitions as well as turn to helix transitions are the most preferred transitions for the wild-type and E46K mutant-type  $\alpha$ -synuclein proteins. Regarding the formation of helical and  $\beta$ -sheet structures, transitions between turn to helix structures are thermodynamically most preferred for residues Val3-Val95 (N-terminal and NAC regions) of the wild-type  $\alpha$ S protein. Coil to  $\beta$ -sheet transitions at Phe4, Glu13, Ala90, and Ala91 are significantly less stable upon E46K mutation of the wild-type form. Helix to  $\beta$ -sheet transition per residues is slightly more stable in parts of the N-terminal and NAC regions in the structures of the E46K mutant-type  $\alpha$ S in comparison to the same regions in the wild-type  $\alpha$ S protein. Similar characteristics are also detected for the turn to  $\beta$ -sheet conversion per residues. Interestingly,  $\beta$ -sheet to helix conversion per residue with dynamics is more preferred in parts of the N-terminal and NAC regions of the E46K mutant-type rather than the same regions of the wild-type  $\alpha$ S protein. In the C-terminal region, turn to helix transitions for the wild-type  $\alpha$ S are more stable at Val95-Lys102, Gln109-Pro117, Pro120-Glu123, and Glu130-Tyr133 in the E46K mutant-type  $\alpha$ S structures. In addition, the preferred coil to  $\beta$ -sheet transitions at Lys102, Asn103, Pro108, and Gln109 in the C-terminal region of the wild-type  $\alpha$ S are less stable in wild-type  $\alpha$ S. Overall, CD and NMR measurements indicated that the secondary structural properties of the wild-type and E46K mutant-type  $\alpha$ S are similar to one another.<sup>18,19,27,28</sup> In contrast, SMF spectroscopy measurements indicated an number of  $\beta$ -sheet conformations upon E46K mutation of the wild-type  $\alpha$ S.<sup>25</sup> Furthermore, E46K mutant-type  $\alpha$ S is reported to convert to  $\beta$ -sheet conformations more readily than the wild-type  $\alpha$ S.<sup>17,19</sup> Our results agree with those experiments that reported an increased  $\beta$ -sheet conformation for the E46K mutant-type rather than wild-type  $\alpha$ S.<sup>17,19,25</sup> Moreover, we identify that this increase in  $\beta$ -sheet conformation occurs in the N-terminal and in the C-terminal regions of the E46K mutant-type  $\alpha$ S. The most preferred transition to a  $\beta$ -sheet



**Figure 2.** WT and E46K  $\alpha$ S secondary structure component transition stability. The stability of secondary structure transitions between two specific secondary structure components per residue for the wild-type (WT) and E46K mutant-type (E46K)  $\alpha$ S proteins. The color scale corresponds to the free energy value associated with specific secondary structure transition between two secondary structure components for a specific residue.

conformation for these residues occurs from a coil structure based on our thermodynamic calculations (Figure 2). The increased  $\beta$ -sheet content of the E46K mutant-type  $\alpha$ S also agrees with the reported increased aggregation propensity of  $\alpha$ S upon E46K mutation that was reported by some experiments given that the fibrillization and aggregation processes have been related to the formation of  $\beta$ -sheet.<sup>17–20</sup> Due to the proposed importance of  $\beta$ -sheet formation in the aggregation process of  $\alpha$ S, our results indicate that residues Val26, Tyr39, Gly47-Ala53, Asp119, Pro120, and Asn122-Tyr125 are most likely to be involved in the fibrillization and aggregation mechanisms of the E46K mutant-type  $\alpha$ S. This is in contrast to the findings of Kumar et al., who reported that residues Gly68-Ala78 and Gly36-Gly41 have the highest aggregation propensity in the E46K mutant-type  $\alpha$ S.<sup>29</sup> While some  $\beta$ -sheet formation occurs in these regions in our simulated structures of the E46K mutant-type  $\alpha$ S, we predict the highest aggregation propensity to be for residues Gly47-Ala53, Asp119, Pro120, and Asn122-Tyr125.

The probability of intramolecular protein interactions between residues in the structures of the wild-type and E46K mutant-type  $\alpha$ S proteins are presented in Figure 3. The NAC and C-terminal regions of the wild-type  $\alpha$ S protein interact

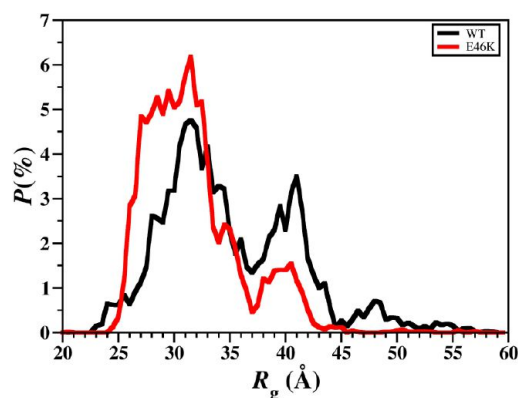


**Figure 3.** WT and E46K  $\alpha$ S tertiary structures. Calculated intramolecular interactions of the wild-type (WT) and the E46K mutant-type (E46K)  $\alpha$ S. The color scale corresponds to the probability ( $P$ ) of the distance between the heavy atoms (C, N, O, S) of a residue being  $\leq 20$  Å from each other.

strongly ( $>50\%$ ) via residues Gly86-Asn103 and Glu104-Asn122 (Figure 3, top). Strong intramolecular interactions between the NAC and C-terminal regions of the E46K mutant-type  $\alpha$ S are also observed between residues Val82-Asp98 and Glu104-Ala124 (Figure 3, bottom). In addition, an increase in moderately abundant interactions ( $\leq 50\%$ ) between Val70-Val82 and Glu110-Val118 as well as between Val82-Glu105 and Glu126-Ala140 occurs upon E46K mutation. Both the wild-type and E46K mutant-type  $\alpha$ S present strong intramolecular interactions with an abundance between residues Val70-Gly84 and Ala85-Leu100 within the NAC region. The E46K mutation also results in an increase in interactions between Val48-Gly67 and Val82-Lys102 by up to 40%. Furthermore, intramolecular interactions between the N-terminal and NAC regions (Gly7-Val66) with the C-terminal region (Gly106-Ala140) are up to 50% more abundant in the E46K mutant-type  $\alpha$ S than the wild-type  $\alpha$ S. In agreement with these findings, Rospigliosi et al. presented that interactions between the C-terminal region with the NAC and N-terminal regions are enhanced upon E46K mutation of the wild-type  $\alpha$ S.<sup>27</sup> Interactions within the N-terminal region are also increased in the E46K mutant-type  $\alpha$ S in comparison to the wild-type  $\alpha$ S between residues Val16-Ala30 and Gly36-Lys58. However, the weak intramolecular

interactions (<10%) between the N-terminal region (Met1-Glu20) with the NAC and C-terminal regions (Thr64-Pro120) completely disappear upon E46K mutation of the wild-type  $\alpha$ S protein. Interestingly, interactions between this region of the N-terminal (Met1-Glu20) and the NAC region have been used to identify the two different possible vesicle binding structures.<sup>38</sup> Experimental measurements have proposed that the  $\alpha$ S binds to vesicle and/or lipid structures in either a broken or extended helical conformation that can depend on the structure of the vesicle or lipid as well as on the protein conformation.<sup>38</sup> Specifically, the extended helical structure is reported to more likely bind to a flat micellar surface whereas the broken helix structure is more likely to bond to a spherical micellar structure.<sup>38</sup> The broken helical structure presents a distance of  $\sim 34$  Å between the N-terminal and NAC regions whereas this distance increases to  $\sim 67$  Å in the extended helix conformation.<sup>38</sup> Our reported results for the interactions between the NAC and N-terminal regions indicate that, upon E46K mutation, the distance is increased and therefore an extended conformation is more likely in the E46K mutant-type rather than wild-type  $\alpha$ S. As a result, we predict that the E46K mutant-type  $\alpha$ S may show a higher binding to the flat micellar surfaces than the wild-type whereas the vice versa trend is expected for spherical micellar surfaces.

The radius of gyration ( $R_g$ ) of a protein yields information about the relative compactness of the overall protein structure, which has been linked to the stability as well as the amount of intramolecular interactions within a protein structure. Figure 4



**Figure 4.** WT and E46K  $\alpha$ S radius of gyration values. Distribution of the radius of gyration ( $R_g$ ) values of the wild-type (black) and E46K mutant-type (red)  $\alpha$ S for the structures obtained after convergence.

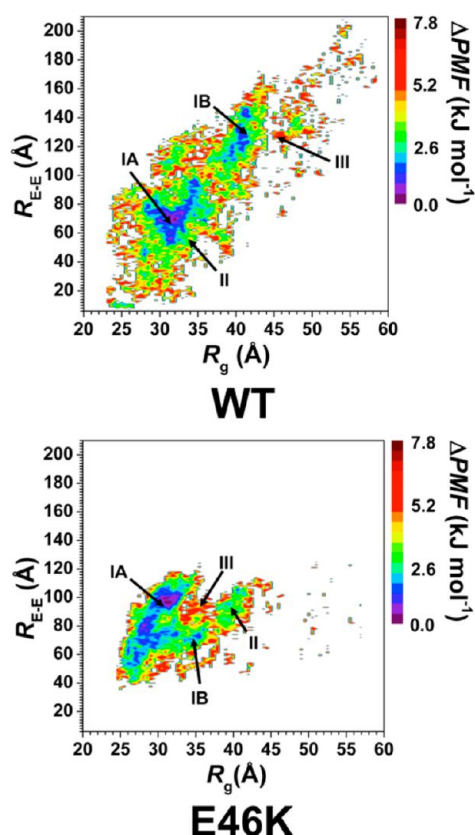
depicts the probability distribution of the  $R_g$  values for the structures of the wild-type and E46K mutant-type  $\alpha$ S. Based on this data, the probability of finding structures with  $R_g$  values between 25.5 and 32.5 Å is higher for the E46K mutant-type rather than wild-type  $\alpha$ S. The vice versa trend is observed for

structures with  $R_g$  values varying between 35.5 and 55.0 Å. Overall, these trends agree with the E46K mutant-type  $\alpha$ S protein having a more compact structure than the wild-type  $\alpha$ S, as stated above. In agreement with these findings as well as those from the tertiary structures of these two proteins, the increased retention time of the E46K mutant-type  $\alpha$ S in comparison to the wild-type  $\alpha$ S in SEC suggests that the conformations of the E46K mutant-type  $\alpha$ S are more compact than those of the wild-type  $\alpha$ S.<sup>18</sup> Overall, these findings suggest that the E46K mutant-type  $\alpha$ S structures possess a greater number of intramolecular interactions and are more stable than the wild-type  $\alpha$ S structures in aqueous solution. In order to quantitatively assess the relative stability of the wild-type and E46K mutant-type  $\alpha$ S structures, conformational Gibbs free energy calculations were performed on the wild-type and E46K mutant-type  $\alpha$ S structures. The difference between these Gibbs free energy values yields an estimate of the degree of difference in the stability of the two protein structures. The average thermodynamic values, presented in Table 1, indicate that the conformations of the wild-type  $\alpha$ S are thermodynamically more stable over the structures of the E46K mutant-type  $\alpha$ S protein by 114.5 kJ mol<sup>-1</sup>. Additionally, the entropic ( $TS$ ) and enthalpic ( $H$ ) contributions to the conformational Gibbs free energies indicate that the wild-type  $\alpha$ S is entropically and enthalpically more preferred than the E46K mutant-type  $\alpha$ S protein by 52.5 and 61.8 kJ mol<sup>-1</sup>, respectively. The decreased thermodynamic preference of the E46K mutant-type  $\alpha$ S structures in comparison to the wild-type  $\alpha$ S conformations indicates that the aggregation rate of the wild-type  $\alpha$ S increases upon E46K mutation. This result is in agreement with those experimental studies that also reported an increased aggregation rate of the E46K mutant-type  $\alpha$ S in comparison to the wild-type  $\alpha$ S.<sup>17–20</sup>

We also assessed the conformational preferences of the wild-type and E46K mutant-type  $\alpha$ S proteins using potential of mean force (PMF) surfaces along the coordinates of  $R_g$  and end-to-end distance ( $R_{E-E}$ ), which is the distance between the N-terminus backbone N atom and the C-terminal backbone carboxyl C atom (Figure 5), following our previous studies on intrinsically disordered fibrillogenic proteins.<sup>37,39,40</sup> The PMF surfaces determine the most favorable protein conformations based on specific structural parameters, such as  $R_g$  and  $R_{E-E}$  via capturing the solvent effects. Different intrinsically disordered proteins structures may have similar thermodynamic properties, which results in the presence of multiple favorable basins along the PMF surface. Analyzing the structures in these different basins provides insights into the conformational properties based on specific structural parameters that influence the favorability of the wild-type and E46K mutant-type  $\alpha$ S conformations in aqueous solution. The PMF surface of the wild-type  $\alpha$ S shows two favorable basins located at  $R_g$  values varying between 28.0 and 35.5 Å (basin IA) and between 37.4

**Table 1.** Calculated Total Energy ( $E_{total}$ ), Solvation Free Energy ( $G_{sol}$ ), Average Enthalpy ( $H$ ), Entropy ( $TS$ ), and Gibbs Free Energy ( $G$ ) Values for the Wild-Type and E46K Mutant-Type  $\alpha$ -Synuclein Proteins

peptide	$\langle E_{total} \rangle$ (kJ mol <sup>-1</sup> )	$\langle G_{sol} \rangle$ (kJ mol <sup>-1</sup> )	$\langle H \rangle$ (kJ mol <sup>-1</sup> )	$-T\langle S \rangle$ (kJ mol <sup>-1</sup> )	$\langle G \rangle$ (kJ mol <sup>-1</sup> )
WT	7146.5 ( $\pm 174.3$ )	-16704.1 ( $\pm 169.7$ )	-9557.4 ( $\pm 18.4$ )	-7043.8 ( $\pm 8.4$ )	-16601.4 ( $\pm 13.0$ )
E46K	5260.6 ( $\pm 450.8$ )	-14756.2 ( $\pm 499.2$ )	-9495.6 ( $\pm 60.9$ )	-6991.3 ( $\pm 13.7$ )	-16486.9 ( $\pm 54.2$ )



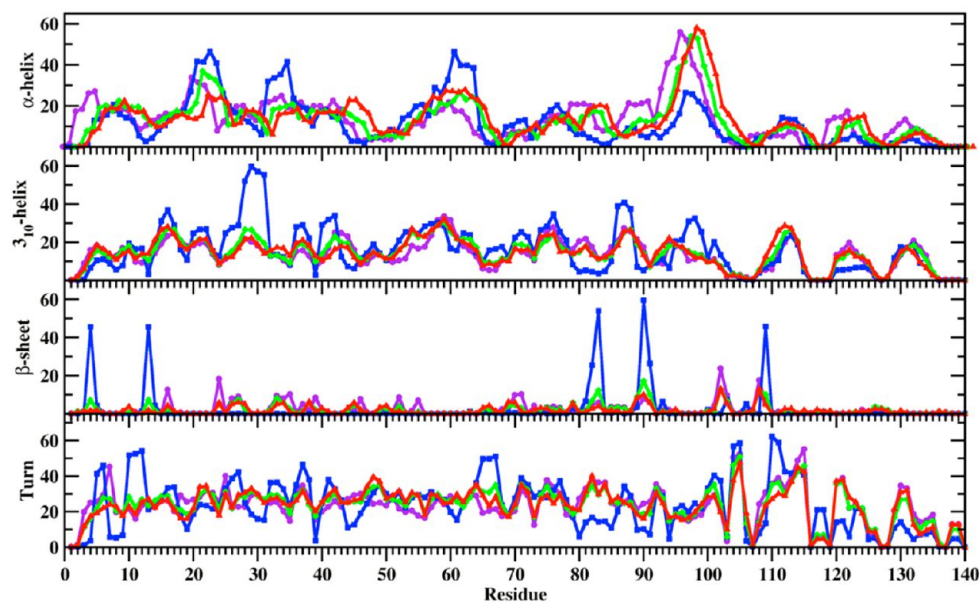
**Figure 5.** WT and E46K  $\alpha$ S PMF surfaces. Change in the potential of mean force ( $\Delta$ PMF) of the wild-type (WT) and E46K mutant-type (E46K)  $\alpha$ S along the coordinates of radius of gyration ( $R_g$ ) and end-to-end distance ( $R_{E-E}$ ) in units of  $\text{kJ mol}^{-1}$ . The reference values for the PMF surfaces are 6.73 and 6.48  $\text{kJ mol}^{-1}$  for the wild-type and E46K mutant-type  $\alpha$ S, respectively.

and 42.0 Å (basin IB) and at  $R_{E-E}$  values varying between 51 and 98 Å (basin IA) and between 108 and 148 Å (basin IB). The difference in PMF values of the region between these two basins ( $>1 k_B T$ ) indicates that transitions between these two favorable PMF basins might require overriding large energy barriers (Figure 5, top). Two favorable PMF basins are also observed for the E46K mutant-type  $\alpha$ S on the PMF surface (Figure 5, bottom). The first basin (basin IA) is located at  $R_g$  values between 25.8 and 33.2 Å and  $R_{E-E}$  values between 50 and 108 Å. The second favorable PMF basin (basin IB) is located at  $R_g$  values between 33.3 and 36.4 Å and  $R_{E-E}$  values between 68 and 79 Å. Interestingly, the  $R_g$  and  $R_{E-E}$  values of basin IA are similar for the wild-type and E46K mutant-type  $\alpha$ S proteins; however, the location of basin IB is shifted upon E46K mutation of the wild-type  $\alpha$ S. Furthermore, our results suggest that conformational transitions between these two favorable PMF basins do not require overriding large energy barriers upon E46K mutation. We also observe significant differences in the secondary structural properties of the structures located in the favorable basins for both the wild-type and E46K mutant-type  $\alpha$ S (Figure 6). For example, the C-terminal region of the wild-type  $\alpha$ S structures located in basin IA presents the most abundant  $\alpha$ -helix formation at residues Phe94-Gly101 (up to 56%), which is decreased by 30% in the basin IB structures (Figure 6, top). In contrast, the most abundant  $\alpha$ -helix structure is observed in the N-terminal and NAC regions of the basin IB structures of the wild-type  $\alpha$ S at residues Glu20-

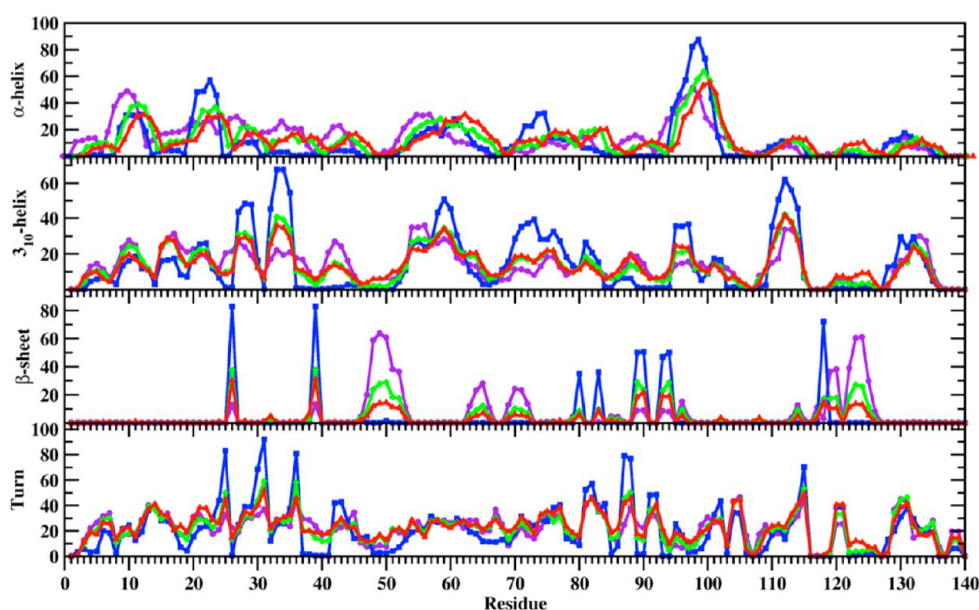
Val26, Lys32-Gly36 and Lys28-Thr64 (up to 46%). Although  $\alpha$ -helix structure is also present at these residues in the basin IA structures, the abundance is 10% less than the abundance in the basin IB wild-type  $\alpha$ S structures. Furthermore, the  $\alpha$ -helix content in the N-terminal and NAC regions of the basin IA wild-type  $\alpha$ S structures at Met5, Lys6, Glu28, Ala29, Leu38, Val40-Lys43, Glu61, Lys80-Glu83, and Ala89-Thr92 decreases in the basin IB structures. Furthermore, we detect that the  $\alpha$ -helix abundance at Glu20-Gln24 and Lys32-Glu35 decreases with decreasing preference along the PMF surface for the basin IB structures (basin IB > basin II > basin III) of the wild-type  $\alpha$ S. The formation of  $3_{10}$ -helix structure also varies between structures of the two favorable PMF basins. The  $3_{10}$ -helix content at residues Val15-Ala17, Gly25-Gly31, Gly36-Leu38, Val40-Ser42, Gly86-Ile88, Val95-Gln99, and Gly101 increases by up to 40% in the basin IB structures rather than the basin IA structures of the wild-type  $\alpha$ S. Despite, residues Lys43-Lys45, Gln79-Ala85, and Asp121-Tyr125 present the vice versa trend for  $3_{10}$ -helix abundance between the basin IA and basin IB structures of the wild-type  $\alpha$ S. However, we do not note any significant trends in the  $3_{10}$ -helix content along the PMF surface for either favorable PMF basin. We also note large variations in the  $\beta$ -sheet structure formed in the basin IA and basin IB structures of the wild-type  $\alpha$ S. For example, the basin IA structures present abundant  $\beta$ -sheet formation (up to 22%) at residues Val16, Gln24, Glu35, Val70, Val71, Lys102, and Pro108. However, the basin IB structures present more abundant  $\beta$ -sheet structure (up to 60%) at residues Phe4, Glu13, Val82, Glu83, Ala90, Ala91, and Gln109. Furthermore, the  $\beta$ -sheet structure at Phe4, Glu13, Val82, Glu83, Ala90, and Ala91 of the basin IB structures decreases along the PMF surface (basin IB > basin II > basin III).

Regarding the secondary structure components of the E46K mutant-type  $\alpha$ S in the two most favorable PMF basins (Figure 6B), we note that the abundant (up to 50%)  $\alpha$ -helix formation at Gly7-Lys12, Gly25-Glu28, Gly31-Val37, and Ala53-Glu57 in the basin IA structures decreases by up to 20% in the basin IB structures. On the other hand, the  $\alpha$ -helix structure at Glu20-Gln24, Thr72-Thr75, and Lys96-Gly101 is up to 40% more abundant in basin IB compared to the basin IA structures. Moreover, residues Leu8-Ala11 and Ala53-Glu57 of the basin IA structures and Glu20-Gln24 and Lys96-Gln99 of the basin IB structures present decreasing  $\alpha$ -helix content along the PMF surface (basin IA > basin II > basin III or basin IB > basin II or basin III, respectively). The formation of  $3_{10}$ -helix structure also displays significant variances between the basin IA and basin IB structures. For example, the  $3_{10}$ -helix content at residues Ser9-Ala11, Val15-Ala19, Gly25, Val26, Gly36-Lys45, Thr54-Ala56, Tyr133, and Gln134 is at least 5% more abundant in the basin IA rather than basin IB structures of the E46K mutant-type  $\alpha$ S. However, the basin IB structures present up to 50% more abundant  $3_{10}$ -helix structure at residues Ala27-Ala29, Lys32-Glu35, Lys58-Glu61, Ala69-Glu83, Val95-Lys97, Glu110-Glu114, and Pro128-Gly132 than in the basin IA structures. We also note that residues Ala27-Ala29 and Lys32-Glu35 display decreasing  $3_{10}$ -helix content along the PMF surface while the same trend is only observed for residues Thr54-Ala56 in the basin IA structures. The formation of  $\beta$ -sheet structure also presents large variations between the basin IA and basin IB structures of the E46K mutant-type  $\alpha$ S. The basin IB structures present abundant (up to 85%)  $\beta$ -sheet structure at residues Val26, Tyr39, Lys80, Glu83, Ala89, Ala90, Gly93, Phe94, and Val118 that is decreased by 60% in the basin IA structures. On

## WT



## E46K

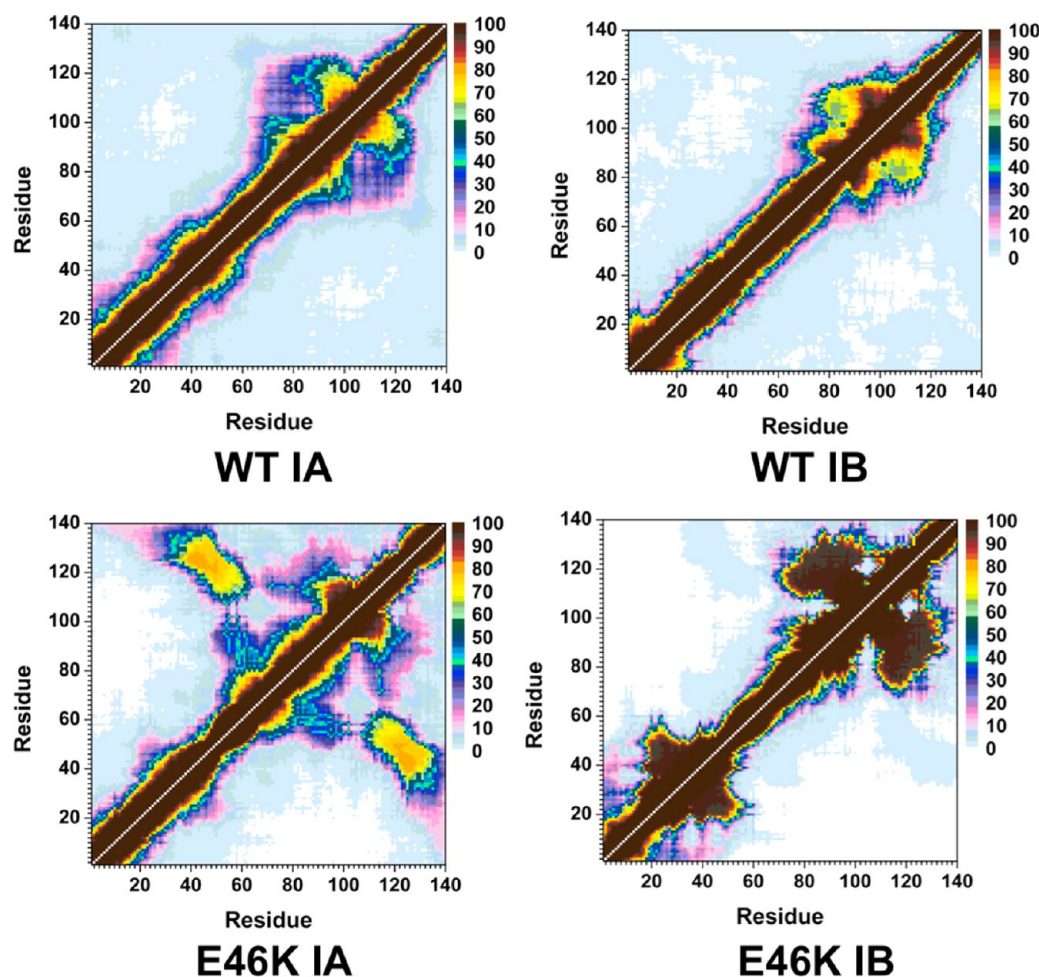


**Figure 6.** WT and E46K  $\alpha$ S secondary structure components of the PMF basin structures. Calculated secondary structure abundances per residue for the structures of the wild-type (WT) and E46K mutant-type (E46K)  $\alpha$ S located in the basin IA (purple), basin IB (blue), basin II (green), and basin III (red) along the PMF surface. The abundances for the  $\pi$ -helix and coil structures are not displayed.

the other side,  $\beta$ -sheet formation occurs at Gly47-Val52, Val63-Val66, Ala69-Thr72, and Asp119-Glu126 with up to 62% more abundance in the basin IA rather than basin IB structures of the E46K mutant-type  $\alpha$ S. Moreover, these residues also display decreasing  $\beta$ -sheet content along the PMF surface for each of their respective basins.

Notable differences in the tertiary structural properties of the basin IA and basin IB structures of the wild-type and E46K mutant-type  $\alpha$ S are also evident (Figure 7). The basin IA structures of the wild-type  $\alpha$ S (Figure 7, top left) show abundant intramolecular interactions between Gly86-Glu104 in the NAC and C-terminal regions with Glu105-Ala124 in the C-terminal region and Val70-Glu83 in the NAC region (between

50% and 95%). The N-terminal and C-terminal regions also interact with the NAC region with up to 45% abundance via residues Val55-Gly86 and Glu104-Tyr136. The N-terminal (Leu8-Val16) and C-terminal (Gly132-Ala140) interactions are also observed with low abundance (5%) in the basin IA structures of the wild-type  $\alpha$ S. In parallel with the basin IA structures, the NAC and C-terminal regions of the basin IB structures of the wild-type  $\alpha$ S interact abundantly via residues Thr72-Ile88 with Ile88-Val118 and residues Ala90-Asn103 with Glu104-Pro120. However, the low abundance interactions between residues Met1-Ala18 in the N-terminal region with the NAC and C-terminal regions in the basin IA structures completely disappear in the basin IB structures. Instead,



**Figure 7.** WT and E46K  $\alpha$ S tertiary structures for the favorable PMF basin structures. Calculated intramolecular interactions of the wild-type  $\alpha$ S structures located in basin IA (WT IA) and basin IB (WT IB) of the PMF surface and of the E46K mutant-type  $\alpha$ S structures located in basin IA (E46K IA) and basin IB (E46K IB) of the PMF surface. The color scale corresponds to the probability ( $P$ ) of the distance between the heavy atoms (C, N, O, S) of a residue being  $\leq 20$  Å from each other.

residues Met1-Lys10 interact with residues Gly14-Gly25 with up to 40% abundance in the basin IB structures of the wild-type  $\alpha$ S. The tertiary structure properties of the basin IA (Figure 7, bottom left) and basin IB (Figure 7, bottom right) structures of the E46K mutant-type  $\alpha$ S also display significant differences. The highly abundant intramolecular interactions between residues Val66-Gly106 and Ala107-Ala140 in the basin IB structures decrease by up to 70% in the basin IA structures of the E46K mutant-type  $\alpha$ S. We also note the same trend for interactions between residues Lys10-Val37 and Leu38-Lys60 as well as between Met1-Lys10 and Val36-Val49 in the N-terminal region of the E46K mutant-type basin IB structures. On the other side, we observe that the interactions of residues Met1-Asn65 with Gly106-Ala140 and residues Val48-Gly68 with Val82-Lys102 are up to 65% more abundant in the structures of basin IA rather than basin IB.

Herein we present the structures of the E46K mutant-type  $\alpha$ S and the impact of the E46K mutation on the structural and thermodynamic properties of the wild-type  $\alpha$ S protein in an aqueous solution environment. The helical content of the wild-type and E46K mutant-type  $\alpha$ S shows an increase for a few residues in the N-terminal and C-terminal regions but a slight decrease in the NAC region upon E46K mutation. The formation of  $\beta$ -sheet structures is significantly increased for the

E46K mutant-type  $\alpha$ S in comparison to the wild-type  $\alpha$ S, especially in parts of the N-terminal and in the C-terminal regions. This result indicates that the N-terminal and C-terminal regions are more likely to participate in the aggregation mechanism of the E46K mutant-type  $\alpha$ S rather than the wild-type  $\alpha$ S. The use of our newly developed theoretical strategy for identifying the thermodynamic preference of secondary structure component transitions reveal that helix to  $\beta$ -sheet as well as  $\beta$ -sheet to helix conversions are more stable in parts of the N-terminal and NAC regions of the E46K mutant-type  $\alpha$ S in comparison to the wild-type  $\alpha$ S protein. Moreover, coil to  $\beta$ -sheet conversion is more stable in parts of the C-terminal region of the wild-type  $\alpha$ S rather than the E46K mutant-type  $\alpha$ S protein. Analyses of the tertiary structures present that the E46K mutation increases the likelihood of long-range intramolecular interactions of the C-terminal region with the N-terminal and NAC regions of  $\alpha$ S. Due to the decrease in N-terminal and NAC region interactions as well as the increase in helical content of the N-terminal region upon E46K mutation, our results also predict that the E46K mutant-type  $\alpha$ S is more likely to bind to vesicles and/or lipid membranes in an extended helical conformation than the wild-type  $\alpha$ S protein. Furthermore, the tertiary structures and  $R_g$  calculations yield that the conformations of the E46K mutant-



type  $\alpha$ S are overall more compact than those of the wild-type form. The abundant intramolecular interactions and compactness of the E46K mutant-type  $\alpha$ S, evaluation of the conformational thermodynamic preferences of the E46K mutant-type and wild-type  $\alpha$ S proteins present that the E46K mutant-type  $\alpha$ S is structurally and thermodynamically more stable than those of the wild-type form. Therefore, the E46K mutant-type  $\alpha$ S structures are more likely to aggregate than the wild-type  $\alpha$ S conformations, which is in agreement with some experiments reporting similar characteristics, but here we present the information at the atomic level with dynamics in an aqueous solution environment. Furthermore, the PMF surfaces of both proteins present two favorable conformational basins with different secondary and tertiary structural properties. However, transitions between these two basins are predicted to occur more readily for the E46K rather than the wild-type  $\alpha$ S due to the decreased energy barrier upon E46K mutation of the wild-type  $\alpha$ S. Overall, our results show that the E46K mutation significantly impacts the structural properties of the wild-type  $\alpha$ S protein not only at the mutation site but also throughout the protein structure. Furthermore, the resulting differences in the structural and thermodynamic properties of the wild-type  $\alpha$ S mutation can be associated with the observed differences in the conformational properties and aggregation kinetics of the two  $\alpha$ S proteins. To the best of our knowledge, this study presents the first detailed insights, at the atomic level with dynamics in an aqueous solution environment, about the structures of the E46K mutant-type  $\alpha$ S and the impact of the E46K mutation on the structures of the wild-type  $\alpha$ S in terms of secondary structure transition stabilities per residue and tertiary structure properties along with conformational free energy landscapes. The information presented herein fills the currently existing gap in our fundamental knowledge about the E46K mutant-type  $\alpha$ S protein and how this mutation impacts the wild-type form. Given that the  $\beta$ -sheet formation has been linked directly to the aggregation process and toxicity of  $\alpha$ S, we propose that the residues we report herein with abundant  $\beta$ -sheet formation should be blocked via small molecules or antibodies to reduce the toxicity.

## METHODS

REMD simulations enable increased conformational sampling of protein structures by using replicas at different temperatures to overcome energy barriers between different minimum energy conformations of the same protein.<sup>41,42</sup> This special sampling is especially necessary in the case of intrinsically disordered proteins, such as  $\alpha$ S, which can possess numerous minimum energy conformations due to their disordered nature. Therefore, we performed REMD simulations on the wild-type and E46K mutant-type full-length  $\alpha$ S proteins using the AMBER ff99SB force field parameters with the AMBER 11 software package.<sup>43,44</sup> The Onufriev-Bashford-Case Generalized Born implicit solvent model was utilized to avoid possible errors due to confined aqueous volume and heat capacity variations observed in explicit solvent REMD simulations.<sup>39,40,45</sup> The replica temperatures were exponentially distributed between 283 and 490 K to yield an approximate exchange probability of 0.70 for both simulations, resulting in a total number of 56 replicas.<sup>46</sup> Starting from fully extended structures, the wild-type and E46K mutant-type  $\alpha$ S structures were equilibrated for 500 ns for each replica. REMD simulations were then performed for a total of 40 ns with exchanges between replicas attempted every 5 ps. The time step was set to 2 fs, and trajectories were saved every 500 steps. The temperature was controlled using Langevin dynamics with a collision frequency of 2 ps<sup>-1</sup>.<sup>47,48</sup> Long-range interactions were treated with the particle mesh Ewald (PME) method with a cutoff value of 450 Å, and

bonds to hydrogen atoms were constrained using the SHAKE algorithm. The convergence of both simulations was confirmed to occur after 20 ns of simulation per replica by the cumulative secondary structure abundance (see the Supporting Information), as shown in our previous studies.<sup>37,39,40,49</sup> All structural and thermodynamic properties of the wild-type and E46K mutant-type  $\alpha$ S were determined from the structures obtained from the 20 ns of simulation after convergence was reached for the replica closest to physiological temperature (310 K). The secondary structure components of both  $\alpha$ S proteins were determined using the DSSP program, which includes hydrogen bond and dihedral angle criteria.<sup>50</sup> In addition, we analyzed the thermodynamic preference of secondary structure transitions between two different secondary structure components with recently developed theoretical strategy.<sup>37</sup> This new strategy utilizes conditional probability to justify the potential of mean force. Specifically, the probability of a secondary structure transition is defined as  $P(t_{i \rightarrow j}|S_i)$ , where  $P(t_{i \rightarrow j})$  is the probability of a transition from one specific ( $i$ ) to another specific secondary structure component ( $j$ ) per residue and  $P(S_i)$  is the probability of a transition between two secondary structure components leading to the specific formation of  $j$  for a specific residue. The secondary structure components are defined based on hydrogen bond and backbone dihedral angle criteria. The hydrogen bond criteria which we implements uses the  $E_{\text{HB}} = E_r E_t E_b$  relationship in which the hydrogen bond energy ( $E_{\text{HB}}$ ) depends on the spatial arrangements; the distance-dependent energy term ( $E_r$ ) is calculated via  $E_r = (A/r^8) + (B/r^6)$  with  $A = 3Er^8$  kcal Å<sup>8</sup> mol<sup>-1</sup> and  $B = 4Er^6$  kcal Å<sup>8</sup> mol<sup>-1</sup>, where  $r$  and  $E$  represent the hydrogen bond distance and its energy, respectively. The angular spatial energy terms are calculated using eq 1:

$$E_p = \begin{cases} (0.9 + 0.1 \sin 2t_i) \cos t_o, & 0^\circ < t_i < 90^\circ \\ K_1(K_2 - \cos^2 t_i)^3 \cos t_o, & 90^\circ < t_i < 110^\circ \\ 0, & t_i > 110^\circ \end{cases} \quad (1)$$

The spatial angles  $t_i$  and  $t_o$  are computed based on the distance deviations between the hydrogen atom and the electron lone pair electron bisector that is located within the lone pair orbital plane and the distance from this plane, respectively. The values  $K_1$  and  $K_2$  are calculated via  $K_1 = (0.9/\cos^6 110^\circ)$  and  $K_2 = \cos^2 110^\circ$ , respectively. The method that we adopt for identifying the residual secondary structures has been used extensively. However, to the best of our knowledge, the current literature does not include a new theoretical strategy now a software package for determining the stability of secondary structure component transitions. We calculate these thermodynamic stabilities using our ProtMet software package eq 2:

$$\text{PMF} = -k_B T \ln Z(\lambda) \quad (2)$$

where  $k_B$  is the Boltzmann constant,  $T$  is the temperature, and  $Z(\lambda)$  is the ratio of the probability of transitions between two distinct secondary structure components per residue to the probability of overall transitions between two, three, or four components leading to the formation of a specific residual secondary structure.

Intramolecular interactions between residues were considered to occur if at least one heavy atom (C, O, N, S) of a specific residue was at least  $\leq 20$  Å from any heavy atom of another residue. A cutoff distance of 20 Å was utilized in order to account for any long-range intramolecular interactions observed experimental from residual dipolar coupling measurements. The thermodynamic properties of the wild-type and E46K mutant-type  $\alpha$ S were assessed using two different methods: the MMPBSA method and PMF calculations.<sup>39,40,51</sup> The MM/PBSA method calculates the conformational Gibbs free energy ( $G$ ) of each structure based on the potential energy ( $E_{\text{total}}$ ), solvation free energy ( $G_{\text{sol}}$ ), and estimated entropy ( $S$ ) of each structure at a specific temperature ( $T$ ) (eq 3).<sup>52</sup>

$$G = E_{\text{total}} + G_{\text{sol}} - TS \quad (3)$$

$G_{\text{sol}}$  is calculated from the sum of the electrostatic and nonpolar contributions. In order to calculate the electrostatic contribution,

dielectric constant values of 1 and 80 are used for the protein and solvent environment, respectively. The entropic contribution to  $G$  was estimated using normal-mode analysis (NMA).<sup>53</sup> The PMF surfaces are determined along the coordinates of radius of gyration ( $R_g$ ) and the end-to-end distance ( $R_{E-E}$ ), as shown in our previous studies on intrinsically disordered proteins.<sup>37,39,40</sup>

## ■ ASSOCIATED CONTENT

### ● Supporting Information

Convergence figures for the wild-type and E46K mutant-type  $\alpha$ S. Convergence figures for the secondary structure transition stabilities and the PMF surfaces of both peptides. Description of the tertiary structural properties of both peptides for the least favorable basins on the PMF surfaces. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [orkid.coskuner@utsa.edu](mailto:orkid.coskuner@utsa.edu).

### Author Contributions

<sup>||</sup>O. C., O.W.-S., and A.D. contributed equally.

### Author Contributions

O.C., O.W.-S., A.D., A.A., and I.S. wrote the paper. O.C. designed the simulations. A.D. and O.W.-S. performed the simulations. O.C. and O.W.-S. designed and coded the developed secondary structure transition software.

### Funding

This research was supported by an allocation and computing resources provided by the National Science Foundation (Grant No. TG-CHE110044). The authors are thankful for the financial support from the University of Texas at San Antonio.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

The calculations and simulations were performed on Kraken at the National Institute for Computational Sciences, Texas Advanced Computing Center (TACC) and the University of Texas at San Antonio Computational Biology Initiative.

## ■ ABBREVIATIONS

Parkinson's disease, PD;  $\alpha$ -synuclein,  $\alpha$ S; circular dichroism, CD; thioflavin T, ThT; nuclear magnetic resonance, NMR; single molecule force, SMF; molecular dynamics, MD; sodium dodecyl sulfate, SDS; replica exchange molecular dynamics, REMD; nonamyloid- $\beta$  component, NAC; radius of gyration,  $R_g$ ; size exclusion chromatography, SEC; end-to-end distance,  $R_{E-E}$ ; potential of mean force, PMF; Molecular Mechanics/Poisson-Boltzmann surface area, MM/PBSA; normal-mode analysis, NMA

## ■ REFERENCES

- (1) Eriksen, J. L., Dawson, T. M., Dickson, D. W., and Petrucelli, L. (2003) Caught in the act: alpha-synuclein is the culprit in Parkinson's disease. *Neuron* 40, 453–456.
- (2) Spillantini, M. G., Schmidt, M. L., Lee, V. M. Y., Trojanowski, J. Q., Jakes, R., and Goedert, M. (1997)  $\alpha$ -Synuclein in Lewy bodies. *Nature* 388, 839–840.
- (3) Goedert, M. (2001) Alpha-synuclein and neurodegenerative diseases. *Nat. Rev. Neurosci.* 2, 492–501.
- (4) Polymeropoulos, M. H., Higgins, J. J., Golbe, L. I., Johnson, W. G., Ide, S. E., DiIorio, G., Sanges, G., Stenroos, E. S., Pho, L. T., Schaffer, A. A., Lazzarini, A. M., Nussbaum, R. L., and Duvoisin, R. C.

(1996) Mapping of a gene for Parkinson's disease to chromosome 4q21-q23. *Science* 274, 1197–1199.

(5) Clayton, D. F., and George, J. M. (1998) The synucleins: a family of proteins involved in synaptic function, plasticity, neurodegeneration and disease. *Trends Neurosci.* 21, 249–254.

(6) Kruger, R., Muller, T., and Riess, O. (2000) Involvement of alpha-synuclein in Parkinson's disease and other neurodegenerative disorders. *J. Neural Transm.* 107, 31–40.

(7) Maroteaux, L., Campanelli, J. T., and Scheller, R. H. (1988) Synuclein - a Neuron-Specific Protein Localized to the Nucleus and Presynaptic Nerve-Terminal. *J. Neurosci.* 8, 2804–2815.

(8) Zarranz, J. J., Alegre, J., Gomez-Esteban, J. C., Lezcano, E., Ros, R., Ampuero, I., Vidal, L., Hoenicka, J., Rodriguez, O., Atares, B., Llorens, V., Tortosa, E. G., del Ser, T., Munoz, D. G., and de Yebenes, J. G. (2004) The new mutation, E46K, of alpha-synuclein causes Parkinson and Lewy body dementia. *Ann. Neurol.* 55, 164–173.

(9) Singleton, A. B., Farrer, M., Johnson, J., Singleton, A., Hague, S., Kachergus, J., Hulihan, M., Peuralinna, T., Dutra, A., Nussbaum, R., Lincoln, S., Crawley, A., Hanson, M., Maraganore, D., Adler, C., Cookson, M. R., Muentner, M., Baptista, M., Miller, D., Blacato, J., Hardy, J., and Gwinn-Hardy, K. (2003) alpha-synuclein locus triplication causes Parkinson's disease. *Science* 302, 841–841.

(10) Kruger, R., Kuhn, W., Muller, T., Woitalla, D., Graeber, M., Kosel, S., Przuntek, H., Epplen, J. T., Schols, L., and Riess, O. (1998) Ala30Pro mutation in the gene encoding alpha-synuclein in Parkinson's disease. *Nat. Genet.* 18, 106–108.

(11) Polymeropoulos, M. H., Lavedan, C., Leroy, E., Ide, S. E., Dehejia, A., Dutra, A., Pike, B., Root, H., Rubenstein, J., Boyer, R., Stenroos, E. S., Chandrasekharappa, S., Athanassiadou, A., Papapetropoulos, T., Johnson, W. G., Lazzarini, A. M., Duvoisin, R. C., DiIorio, G., Golbe, L. I., and Nussbaum, R. L. (1997) Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science* 276, 2045–2047.

(12) Winner, B., Jappelli, R., Maji, S. K., Desplats, P. A., Boyer, L., Aigner, S., Hetzer, C., Loher, T., Vilar, M. a., Campioni, S., Tzitzilonis, C., Soragni, A., Jessberger, S., Mira, H., Consiglio, A., Pham, E., Masliah, E., Gage, F. H., and Riek, R. (2011) In vivo demonstration that  $\alpha$ -synuclein oligomers are toxic. *Proc. Natl. Acad. Sci. U.S.A.* 108, 4194–4199.

(13) Outeiro, T. F., Putcha, P., Tetzlaff, J. E., Spoelgen, R., Koker, M., Carvalho, F., Hyman, B. T., and McLean, P. J. (2008) Formation of Toxic Oligomeric  $\alpha$ -Synuclein Species in Living Cells. *PLoS One* 3, e1867.

(14) Li, J., Uversky, V. N., and Fink, A. L. (2002) Conformational behavior of human alpha-synuclein is modulated by familial Parkinson's disease point mutations A30P and A53T. *Neurotoxicology* 23, 553–567.

(15) Li, J., Uversky, V. N., and Fink, A. L. (2001) 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.

(16) Choong, C. J., and Say, Y. H. (2011) Neuroprotection of alpha-synuclein under acute and chronic rotenone and maneb treatment is abolished by its familial Parkinson's disease mutations A30P, A53T and E46K. *Neurotoxicology* 32, 857–863.

(17) Ono, K., Ikeda, T., Takasaki, J., and Yamada, M. (2011) Familial Parkinson disease mutations influence alpha-synuclein assembly. *Neurobiol. Dis.* 43, 715–724.

(18) Fredenburg, R. A., Rospigliosi, C., Meray, R. K., Kessler, J. C., Lashuel, H. A., Eliezer, D., and Lansbury, P. T. (2007) The impact of the E46K mutation on the properties of alpha-synuclein in its monomeric and oligomeric states. *Biochemistry* 46, 7107–7118.

(19) Greenbaum, E. A., Graves, C. L., Mishizen-Eberz, A. J., Lupoli, M. A., Lynch, D. R., Englander, S. W., Axelsen, P. H., and Giasson, B. I. (2005) The E46K mutation in alpha-synuclein increases amyloid fibril formation. *J. Biol. Chem.* 280, 7800–7807.

(20) Choi, W., Zibae, S., Jakes, R., Serpell, L. C., Davletov, B., Crowther, R. A., and Goedert, M. (2004) Mutation E46K increases

phospholipid binding and assembly into filaments of human alpha-synuclein. *FEBS Lett.* 576, 363–368.

(21) Kamiyoshihara, T., Kojima, M., Ueda, K., Tashiro, M., and Shimotakahara, S. (2007) Observation of multiple intermediates in alpha-synuclein fibril formation by singular value decomposition analysis. *Biochem. Biophys. Res. Commun.* 355, 398–403.

(22) Bertocini, C. W., Jung, Y. S., Fernandez, C. O., Hoyer, W., Griesinger, C., Jovin, T. M., and Zweckstetter, M. (2005) Release of long-range tertiary interactions potentiates aggregation of natively unstructured alpha-synuclein. *Proc. Natl. Acad. Sci. U.S.A.* 102, 1430–1435.

(23) Bertocini, C. W., Fernandez, C. O., Griesinger, C., Jovin, T. M., and Zweckstetter, M. (2005) Familial mutants of  $\alpha$ -synuclein with increased neurotoxicity have a destabilized conformation. *J. Biol. Chem.* 280, 30649–30652.

(24) Zhu, M., and Fink, A. L. (2003) Lipid binding inhibits  $\alpha$ -synuclein fibril formation. *J. Biol. Chem.* 278, 16873–16877.

(25) Brucale, M., Sandal, M., Di Maio, S., Rampioni, A., Tessari, I., Tosatto, L., Bisaglia, M., Bubacco, L., and Samori, B. (2009) Pathogenic Mutations Shift the Equilibria of  $\alpha$ -Synuclein Single Molecules towards Structured Conformers. *ChemBioChem* 10, 176–183.

(26) Bussell, R., Jr, and Eliezer, D. (2001) Residual structure and dynamics in Parkinson's disease-associated mutants of  $\alpha$ -synuclein. *J. Biol. Chem.* 276, 45996–46003.

(27) Rospigliosi, C. C., McClendon, S., Schmid, A. W., Ramlall, T. F., Barre, P., Lashuel, H. A., and Eliezer, D. (2009) E46K Parkinson's-Linked Mutation Enhances C-Terminal-to-N-Terminal Contacts in alpha-Synuclein. *J. Mol. Biol.* 388, 1022–1032.

(28) Ferreon, A. C. M., Moran, C. R., Ferreon, J. C., and Deniz, A. A. (2010) Alteration of the  $\alpha$ -Synuclein Folding Landscape by a Mutation Related to Parkinson's Disease. *Angew. Chem., Int. Ed.* 49, 3469–3472.

(29) Kumar, S., Sarkar, A., and Sundar, D. (2009) Controlling aggregation propensity in A53T mutant of alpha-synuclein causing Parkinson's disease. *Biochem. Biophys. Res. Commun.* 387, 305–309.

(30) Perlmutter, J. D., Braun, A. R., and Sachs, J. N. (2009) Curvature Dynamics of alpha-Synuclein Familial Parkinson Disease Mutants. Molecular Simulations of the Micelle- And Bilayer-Bound Forms. *J. Biol. Chem.* 284, 7177–7189.

(31) Eliezer, D., Kutluay, E., Bussell, R., and Browne, G. (2001) Conformational properties of alpha-synuclein in its free and lipid-associated states. *J. Mol. Biol.* 307, 1061–1073.

(32) Davidson, W. S., Jonas, A., Clayton, D. F., and George, J. M. (1998) Stabilization of alpha-synuclein secondary structure upon binding to synthetic membranes. *J. Biol. Chem.* 273, 9443–9449.

(33) Jo, E. J., McLaurin, J., Yip, C. M., St, George-Hyslop, P., and Fraser, P. E. (2000) alpha-synuclein membrane interactions and lipid specificity. *J. Biol. Chem.* 275, 34328–34334.

(34) Georgieva, E. R., Ramlall, T. F., Borbat, P. P., Freed, J. H., and Eliezer, D. (2010) The Lipid-binding Domain of Wild Type and Mutant alpha-Synuclein. Compactness and Interconversion between the Broken and Extended Helix Forms. *J. Biol. Chem.* 285, 28261–28274.

(35) Kessler, J. C., Rochet, J. C., and Lansbury, P. T. (2003) The N-terminal repeat domain of alpha-synuclein inhibits beta-sheet and amyloid fibril formation. *Biochemistry* 42, 672–678.

(36) Vilar, M., Chou, H. T., Luhrs, T., Maji, S. K., Riek-Loher, D., Verel, R., Manning, G., Stahlberg, H., and Riek, R. (2008) The fold of alpha-synuclein fibrils. *Proc. Natl. Acad. Sci. U.S.A.* 105, 8637–8642.

(37) Coskuner, O., Wise-Scira, O., Perry, G., and Kitahara, T. (2012) The Structure of the E22 $\Delta$  Mutant-Type Amyloid- $\beta$  Alloforms and the Impact of E22 $\Delta$  Mutation on the Structures of the Wild-Type Amyloid- $\beta$  Alloforms. *ACS Chem. Neurosci.*, published online October 25, 1012. DOI: 10.1021/cn300149j.

(38) Ferreon, A. C. M., Gambin, Y., Lemke, E. A., and Deniz, A. A. (2009) Interplay of alpha-synuclein binding and conformational switching probed by single-molecule fluorescence. *Proc. Natl. Acad. Sci. U.S.A.* 106, 5645–5650.

(39) Wise-Scira, O., Xu, L., Kitahara, T., Perry, G., and Coskuner, O. (2011) Amyloid-beta peptide structure in aqueous solution varies with fragment size. *J. Chem. Phys.* 135, 205101.

(40) Wise-Scira, O., Xu, L., Perry, G., and Coskuner, O. (2012) Structures and free energy landscapes of aqueous zinc(II)-bound amyloid-beta(1–40) and zinc(II)-bound amyloid-beta(1–42) with dynamics. *J. Biol. Inorg. Chem.* 17, 927–938.

(41) Sugita, Y., and Okamoto, Y. (1999) Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* 314, 141–151.

(42) Zhang, W., Wu, C., and Duan, Y. (2005) Convergence of replica exchange molecular dynamics. *J. Chem. Phys.* 123, 154105.

(43) Simmerling, C., Hornak, V., Abel, R., Okur, A., Strockbine, B., and Roitberg, A. (2006) Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins* 65, 712–725.

(44) Case, D., Darden, T., Cheatham, T., Simmerling, C., Wang, J., Duke, R., Luo, R., Walker, R., Zhang, W., and Merz, K. (2010) AMBER 11, University of California, San Francisco.

(45) Case, D. A., Onufriev, A., and Bashford, D. (2004) Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins* 55, 383–394.

(46) Patriksson, A., and van der Spoel, D. (2008) A temperature predictor for parallel tempering simulations. *Phys. Chem. Chem. Phys.* 10, 2073–2077.

(47) Allen, M. P., and Tildesley, D. J. (1999) *Computer simulation of liquids*, Clarendon Press: Oxford.

(48) Frenkel, D., and Smit, B. (2002) *Understanding molecular simulation: from algorithms to applications*, Vol. 1, Academic Press: San Diego, CA.

(49) Fawver, J. N., Duong, K. T., Wise-Scira, O., Petrofes Chapa, R., Schall, H. E., Coskuner, O., Zhu, X., Colom, L. V., and Murray, I. V. J. (2012) Probing and Trapping a Sensitive Conformation: Amyloid- $\beta$  Fibrils, Oligomers, and Dimers. *J. Alzheimer's Dis.* 32, 197–215.

(50) Kabsch, W., and Sander, C. (1983) Dictionary of Protein Secondary Structure - Pattern-Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers* 22, 2577–2637.

(51) Lee, M. R., Duan, Y., and Kollman, P. A. (2000) Use of MM-PB/SA in estimating the free energies of proteins: Application to native, intermediates, and unfolded villin headpiece. *Proteins: Struct., Funct., Genet.* 39, 309–316.

(52) Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S. H., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W., Donini, O., Cieplak, P., Srinivasan, J., Case, D. A., and Cheatham, T. E. (2000) Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Acc. Chem. Res.* 33, 889–897.

(53) Case, D. A. (1994) Normal-Mode Analysis of Protein Dynamics. *Curr. Opin. Struct. Biol.* 4, 285–290.