

# Characterization of $\text{A}\beta$ Monomers through the Convergence of Ensemble Properties among Simulations with Multiple Force Fields

David J. Rosenman,<sup>†,‡</sup> Chunyu Wang,<sup>†,‡,§</sup> and Angel E. García\*,<sup>†,‡,¶,⊥</sup>

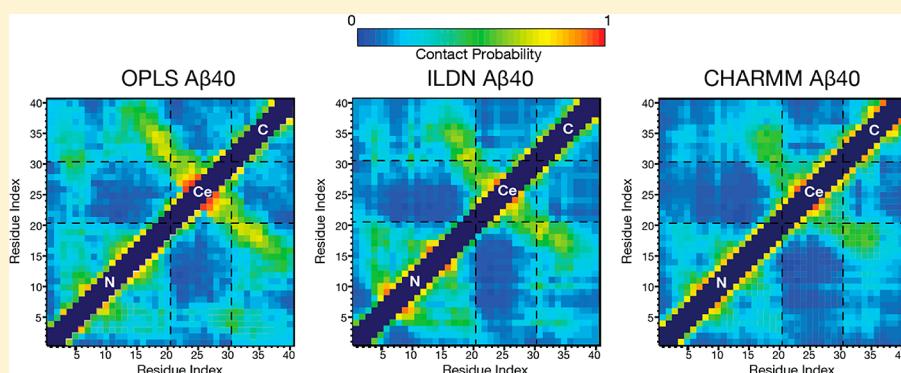
<sup>†</sup>Department of Biology, Rensselaer Polytechnic Institute, 110 Eighth Street, Troy, New York 12180, United States

<sup>‡</sup>Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, 110 Eighth Street, Troy, New York 12180, United States

<sup>§</sup>Graduate Program in Biochemistry and Biophysics, Rensselaer Polytechnic Institute, 110 Eighth Street, Troy, New York 12180, United States

<sup>¶</sup>Department of Physics, Applied Physics and Astronomy, Rensselaer Polytechnic Institute, 110 Eighth Street, Troy, New York 12180, United States

<sup>⊥</sup>Center for Non Linear Studies, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, United States



**ABSTRACT:** Amyloid  $\beta$  ( $\text{A}\beta$ ) monomers represent a base state in the pathways of aggregation that result in the fibrils and oligomers implicated in the pathogenesis of Alzheimer's disease (AD). The structural properties of these intrinsically disordered peptides remain unclear despite extensive experimental and computational investigations. Further, there are mutations within  $\text{A}\beta$  that change the way the peptide aggregates and are known to cause familial AD (FAD). Here, we analyze the ensembles of different isoforms ( $\text{A}\beta_{42}$  and  $\text{A}\beta_{40}$ ) and mutants (E22 $\Delta$ , D23N, E22K, E22G, and A2T in  $\text{A}\beta_{40}$ ) of  $\text{A}\beta$  generated with all-atom replica exchange molecular dynamics (REMD) simulations on the  $\mu\text{s}/\text{replica}$  time scale. These were run using three different force field/water model combinations: OPLS-AA/L and TIP3P ("OPLS"), AMBER99sb-ILDN and TIP4P-Ew ("ILDN"), as well as CHARMM22\* and TIP3P ("CHARMM"). Despite fundamental changes in simulation parameters, we find that the resulting ensembles demonstrate a strong convergence in structural properties. In particular, antiparallel contacts between L17–A21 and A30–L34 are prevalent in ensembles of  $\text{A}\beta_{40}$ , directly forming  $\beta$  sheets in the OPLS and ILDN combinations. A21–A30 commonly forms an interceding region that rarely interacts with the rest of the peptide. Further,  $\text{A}\beta_{42}$  contributes new  $\beta$  hairpin motifs involving V40–I41 in both OPLS and ILDN. However, the structural flexibility of the central region and the electrostatic interactions that characterize it are notably different between the different conditions. Further, for OPLS, each of the FAD mutations disrupts central bend character and increases the polymorphism of antiparallel contacts across the central region. However, the studied mutations in the ILDN set primarily encourage more global contacts involving the N-terminus and the central region, and promote the formation of new  $\beta$  topologies that may seed different aggregates involved in disease phenotypes. These differences aside, the large degree of agreement between simulation sets across multiple force fields provides a generalizable characterization of  $\text{A}\beta$  that is also consistent with experimental data and models.

## 1. INTRODUCTION

20–30% of the sequences in the human proteome are classified as intrinsically disordered proteins/peptides (IDPs); these amino acid sequences are highly polymorphic and fail to adopt a stable native structure under physiological conditions.<sup>1</sup> Several IDPs play detrimental roles in human pathologies like neurodegenerative diseases, HIV transmission, and diabetes due to their propensity to misfold or aggregate.<sup>2,3</sup> One of these is

amyloid  $\beta$  ( $\text{A}\beta$ ), a peptide that forms aggregates that are intimately implicated in the etiology of Alzheimer's disease (AD),<sup>4</sup> a terminal, neurodegenerative disease estimated to affect 36 million people worldwide.<sup>5</sup> Interestingly, small

**Received:** September 25, 2015

**Revised:** October 30, 2015

**Published:** November 12, 2015

changes to the physicochemistry of the  $\text{A}\beta$  are observed to drastically change the aggregation and toxicity properties of the peptide.<sup>6–8</sup> For example, the generation of the C-terminus of  $\text{A}\beta$  from the amyloid precursor protein (APP) is prone to variable cleavage by the enzyme  $\gamma$ -secretase, resulting in the production of  $\text{A}\beta$  alloforms that can range from 34 to 43 residues in length.<sup>9</sup> The two predominant variants in the brain,  $\text{A}\beta_{40}$  and  $\text{A}\beta_{42}$ , differ by only 2 amino acids at the C-terminus; however,  $\text{A}\beta_{42}$  forms unique aggregates<sup>6,7</sup> and is more amyloidogenic and toxic<sup>10–12</sup> than the  $\text{A}\beta_{40}$  variant.

Another example of how small chemical changes affect  $\text{A}\beta$  self-association is the collection of point mutations of  $\text{A}\beta$  that are involved in familial Alzheimer's disease (FAD). While there are genetic risk factors underlying AD in general, FAD represents a subset of cases in which mutations in known genes directly cause the disease. FAD is further distinguished by its early onset compared to sporadic AD: the former generally occurs in patients 30–65 years, while the latter usually affects people older than 65. Further, patients with FAD are particularly prone to other symptoms such as cerebral amyloid angiopathy (CAA), stroke, and hemorrhaging.<sup>13,14</sup> While FAD describes less than 5% of all cases,<sup>15</sup> FAD is known to be caused by mutations, usually autosomal dominant, of APP,  $\gamma$ -secretase, and other proteins that process APP.<sup>14</sup> These mutations have thus proved instrumental for understanding the mechanisms involved in AD.<sup>13,14</sup>

A subset of these FAD mutations exist in the  $\text{A}\beta$  sequence itself; while some of these mutants may affect the processing of  $\text{A}\beta$ , some are also known to change the aggregation properties of the  $\text{A}\beta$  in the brain.<sup>13,14</sup> These  $\text{A}\beta$  affiliated FAD mutations are often associated with changes in  $\text{A}\beta$  production and the types of aggregates formed, but the effects and clinical manifestations of these mutants are also substantially different from each other.<sup>13,14</sup> Mutations appear to cluster into a few regions in the  $\text{A}\beta$  sequence; the hotspot with the most catalogued mutations, A21–D23, occurs in the center of the peptide. This cluster contains two anionic side chains that are adjacent in peptide sequence (E22, D23). All of the mutations in this hotspot are affiliated with increased fibril formation except A21G and E22G.<sup>16–26</sup> It has also been observed that the Iowa substitution (D23N) in the context of  $\text{A}\beta_{40}$  forms fibrils with faster kinetics and substantially more polymorphism at the molecular level than wild type.<sup>22</sup> Further,  $\text{A}\beta_{40}$ -D23N's fibrils can possess an antiparallel cross- $\beta$  topology; this is currently the only  $\text{A}\beta$  variant observed to adopt this form.<sup>22,26</sup> Meanwhile, the more recently discovered Osaka deletion mutant (E22 $\Delta$ ) forms aggregates very rapidly in the context of  $\text{A}\beta_{40}$  and  $\text{A}\beta_{42}$ , requiring no lag phase and a substantially reduced critical concentration than wild type.<sup>23,25</sup> It also associates into aggregates with a unique morphology: E22 $\Delta$  is now believed to form a new type of short, single-stranded fibril,<sup>23–25</sup> differing from the two-stranded striated ribbon model for wild type fibrils. Besides the formation of differential aggregation types, it has been observed that the increased oligomerization order of FAD mutations in E22 and D23 has been linked to increased susceptibility to trypsin proteolysis and destabilization of a turn in the  $\text{A}\beta(21–30)$  fragment.<sup>21</sup> Since the monomer form can seed these aggregates, it is quite possible that these mutations alter the conformational landscape of  $\text{A}\beta$  at the monomer level.

One additional point mutant of  $\text{A}\beta_{40}$  studied here is the rare Icelandic mutation A2T, the first  $\text{A}\beta$  substitution found to protect against both Alzheimer's disease and age-related mental decline in the elderly.<sup>27</sup> Although this particular A2T mutant

(unlike the others simulated here) is proximate to the N-terminal cleavage site and has been observed to reduce both  $\text{A}\beta$  production and  $\beta$  cleavage activity by BACE1,<sup>27</sup> the effects of this mutant on  $\text{A}\beta$  structure, dynamics, and aggregation may also play a role in its phenotype of improving cognitive health in advanced age. A comprehensive table of the phenotypic behavior of point mutations of  $\text{A}\beta$  found in human subjects has been previously published.<sup>28</sup>

We hypothesize that the reason small modifications to  $\text{A}\beta$  drastically affect the way the peptide aggregates is linked to the intrinsic disorder of monomer state; small biases in the ensembles caused by these changes could allow the disordered peptide to sample unique structures that can seed new aggregation pathways. There have been many efforts to determine the structural properties of  $\text{A}\beta$  monomers experimentally and computationally; however, progress has been frustrated both by the high aggregation propensity and the aforementioned inability of  $\text{A}\beta$  monomers to establish a stable native state under physiological conditions. While a number of computational investigations have been published to investigate full-length  $\text{A}\beta$  monomeric ensembles (refer to citations 29–39 for examples), there is a disparity in the secondary and tertiary structural properties of the simulated ensembles reported in these studies. These discrepancies may arise from (among other factors) differences in simulation parameters, insufficient sampling, and differences in simulation analysis protocols. Among the first category, inconsistencies in force field have been documented to lead to widely varying results in unfolded and intrinsically disordered ensembles.<sup>40–43</sup>

To derive a model of  $\text{A}\beta$  monomers that is robust to the changes across multiple parameter sets, this work compares unbiased, all-atom, explicitly solvated REMD simulations performed over hundreds of ns/replica using different force field and water model combinations: OPLS-AA/L<sup>44</sup> and TIP3P,<sup>45</sup> AMBER99sb-ILDN<sup>46</sup> and TIP4P-Ew,<sup>47</sup> as well as CHARMM22\*<sup>40</sup> and modified CHARMM TIP3P (TIP3SP).<sup>45</sup> These are referred to in this paper as the "OPLS", "ILDN", and "CHARMM" conditions, respectively. Additionally, several alloforms ( $\text{A}\beta_{42}$  and  $\text{A}\beta_{40}$ ) and point mutants of  $\text{A}\beta_{40}$  (E22 $\Delta$ , D23N, and E22K) were investigated here with both the OPLS and ILDN sets. We have previously characterized wild type  $\text{A}\beta$  with the same OPLS combination described above.<sup>48</sup> However, unlike OPLS (to our knowledge), both AMBER99sb<sup>49</sup> (with \*<sup>50</sup> and/or ILDN<sup>46</sup> modifications) and CHARMM22\*<sup>40</sup> have been shown to be capable of reversibly folding multiple globular proteins, including those with different  $\alpha$  helix and  $\beta$  sheet compositions, as well as different folds.<sup>51–53</sup>

The parametrization philosophies of OPLS and AMBER type force fields are comparable; in particular, both force fields employ the following: a nearly identical functional form, very similar bonded parameters (peptide parameters for bond angles and dihedrals were originally adapted from the same force field published by Weiner et al.<sup>54</sup> for both OPLS and AMBER fields), similar fitting of Lennard-Jones parameters to match the experimental heat of vaporization for model liquid compounds, and a similar philosophy (though employing different mixing rules) for the rescaling of 1–4 nonbonded interactions.<sup>55–59</sup> Therefore, the most relevant difference between the two force fields is in the way partial charges are calculated. For new molecules, OPLS charges are usually derived using parameters transferred from previously fitted model compounds that are then either coparameterized with Lennard-Jones to reproduce

the experimental properties of neat liquid compounds and/or adjusted based on chemical intuition; either way, this method often requires no new explicit quantum mechanical (QM) fitting.<sup>55,57,59</sup> In contrast, AMBER charges are explicitly derived for each new compound by fitting single residue fragments *in vacuo*; this process entails optimizing the restricted electrostatic potential once the geometry of the molecule has been refined with *ab initio* QM methods.<sup>56,58,60</sup>

Additionally, we are also employing different water models for the parameter combinations investigated here. TIP3P is a three site model (one for each atom of water), while TIP4P-Ew has four sites.<sup>45,47</sup> The new fourth site is introduced to move the center of the charge density away from the center of the Lennard-Jones interaction for the oxygen atom: this entails introducing a negatively charged dummy atom located near the oxygen that bisects the normal H–O–H bond angle, as well as adjusting the partial charges of the oxygen and hydrogen atoms.<sup>45,47</sup> Although this extra site adds extra computational cost, its addition has been observed to better model the electrostatic potential of water, as observed through neutron diffraction and thermodynamics properties.<sup>45</sup> Besides changes in partial charge, the TIP3P and TIP4P-Ew possess identical bond length and bond angle properties for its hydrogens and oxygen, while both the Lennard-Jones repulsive and attractive terms are roughly ~10% weaker in TIP3P.<sup>45,47</sup>

The third force field considered, CHARMM, was chosen to address an emerging concern in the simulation of IDPs in the literature: it has been observed that modern force fields (including OPLS-AA and AMBER99sb-ILDN) underestimate the experimental radii of gyration as measured by small-angle X-ray scattering (SAXS) measurements for IDPs;<sup>41,42,61</sup> the CHARMM22\* force field has been shown to be the least prone to this bias (though even this and similar CHARMM force fields have been suggested to underestimate radii of gyration of unfolded proteins compared to experimental measures<sup>62,63</sup>). CHARMM force fields have a similar functional form but are parametrized using a fairly different protocol compared to those of AMBER and OPLS; in CHARMM, both charges and Lennard-Jones terms are iteratively optimized for each group to reproduce gas phase quantum mechanical energies.<sup>40,64</sup>

Our study reveals that all three force fields produce ensembles with a remarkable similarity in terms of tertiary structural properties, despite the use of fundamentally different parameters. Further, OPLS and ILDN demonstrate antiparallel  $\beta$  contacts among residues L17–A21, A30–L34, and, for  $\text{A}\beta$ 42, V40–I41, consistent with transient sampling of  $\beta$  sheets between these regions. Likewise, the central region (21–30) rarely interacts with the rest of the peptide except at the N-terminus and forms a loop-like structure in both simulations. Meanwhile, the CHARMM  $\text{A}\beta$ 40 simulations maintain similar antiparallel contacts and high  $\beta$  content at L17–A21, but are affiliated with a more extended ensemble characterized by turn regions at the G25–S26, G29–A30, G33–L34, and G37–G38 with intervening coil regions that dominate the C-terminal half of the peptide.

Interestingly, the force fields vary in how they model the structural and electrostatic nature of the central region, with OPLS strongly favoring high curvature bend secondary structure while ILDN and CHARMM sample a mixture of bend and turn structures in this region, with higher populations for helical hydrogen bonds. While the aforementioned central hairpin is prevalent in ILDN, the increased flexibility of this central region permits ILDN to sample more diverse

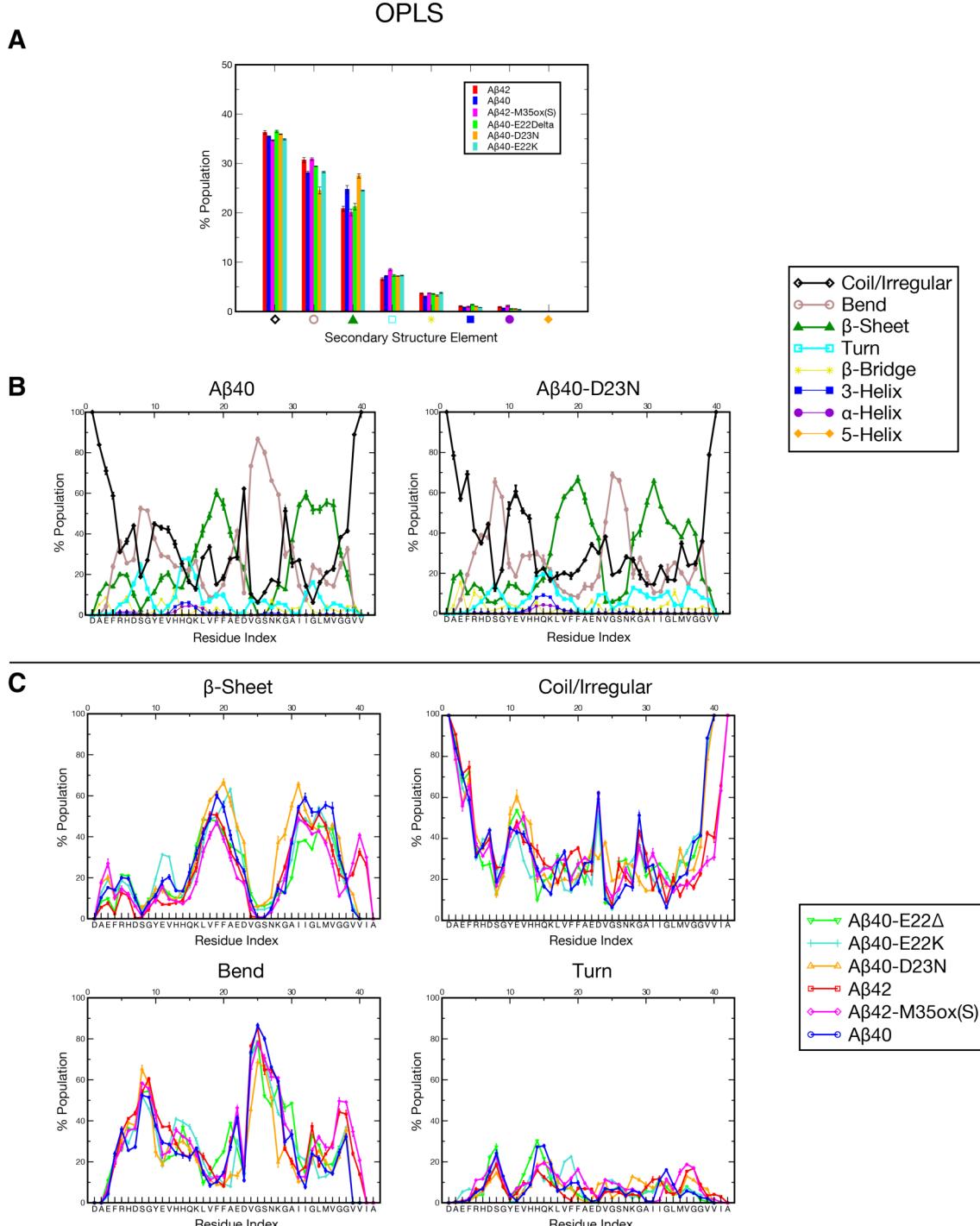
combinations of  $\beta$  sheets compared to those of OPLS. The nature of familial AD-linked residues E22 and D23 is also quite different between the ensembles, with the OPLS combination strongly populating the side chains of these residues in bend-promoting electrostatic interactions, while the ILDN and CHARMM combinations favor a more diverse role for these side chains and often have these interacting with solvent. As such, the effect of FAD-linked mutations of these residues is quite different among the force field sets. OPLS implicates these mutants in affecting the physicochemistry of the central bend and bonding across the central hairpin, but ILDN simulations indicate that the point substitutions investigated primarily promote the formation of both local and global  $\beta$  hairpins involving the N-terminus, at the cost of interactions of this region with the central hydrophobic cluster (CHC, L17–A21) more prevalent in wild type. The contribution of new  $\beta$  topologies with N-terminal residues seen in  $\text{A}\beta$  point mutants may be able to seed different pathogenic aggregates involved in familial AD. Despite these differences, the overall consensus between the extensively simulated ensembles reported in this study provides a force-field-independent characterization of the  $\text{A}\beta$  ensemble that also matches experimentally measured biases for  $\text{A}\beta$  monomers, and bears resemblance to experimental models of aggregates such as oligomers and fibrils (see Discussion). Further, the specific changes reported between the simulated ensembles here provide insight into the differences between force fields in the simulation of intrinsically disordered proteins that may facilitate force field refinement.

## 2. METHODS

### 2.1. Simulation Details.

All simulations in this study were generated with the replica exchange molecular dynamics (REMD) protocol, each using identical simulation parameters except water model, force field, and simulation duration. The REMD simulation protocol<sup>65,66</sup> begins with several initially identical copies (called replicas) that are each simulated over a set of increasing temperatures. Replicas that neighbor each other in temperature are exchanged with a probability governed by the Metropolis Monte Carlo criterion throughout the simulation. The “OPLS” simulations in this study use the OPLS-AA/L force field<sup>44</sup> and the TIP3P water model.<sup>45</sup> Three of the simulations discussed here,  $\text{A}\beta$ 42,  $\text{A}\beta$ 40, and  $\text{A}\beta$ 42-M35ox(s) with OPLS, were published previously,<sup>48</sup> but some data are reproduced here for the sake of comparison with new simulations. The  $\text{A}\beta$ 42-M35ox(S) system replaces methionine 35 with L-methioine-(S)-sulfoxide, and was simulated using the modifications to the OPLS force field discussed previously.<sup>48</sup> The “ILDN” simulations are coupled to the AMBER99sb-ILDN force field<sup>46</sup> and the TIP4P-Ew water model.<sup>47</sup> This combination is similar to one used in a previous study<sup>31</sup> (the only difference being the latter uses the older AMBER99sb<sup>67</sup> force field) that yielded an ensemble able to reproduce experimental NMR observables with reasonable accuracy. We chose nonetheless to use AMBER99sb-ILDN rather than the original AMBER99sb force field because the ILDN side chain torsion modifications have been shown to considerably improve agreement with both side chain and backbone NMR data for microsecond time scale simulations.<sup>46</sup> Finally, the “CHARMM” parameters include the CHARMM22\* force field<sup>40</sup> and the modified CHARMM TIP3P (TIP3SP) water model.<sup>45</sup>

All  $\text{A}\beta$  systems were initially constructed as fully extended structures with zwitterionic termini and protonation states that



**Figure 1.** Percent secondary structure for A $\beta$ , as annotated by DSSP,<sup>69,70</sup> averaged over the production ensemble, for simulations done with the OPLS condition (A) over all residues and (B and C) for each residue index. (B) Profile for all DSSP elements for two example A $\beta$  variants, A $\beta$ 40 and A $\beta$ 40-D23N. The dominant secondary structure elements, coil/irregular in black, bend in gray,  $\beta$  sheet in green, and turn in cyan, have been illustrated with thicker lines on the graph. (C) Plot comparing all of the A $\beta$  variants studied here for these four dominant DSSP elements. The sequence of wild type A $\beta$ 42 is on the lower x-axis, while residue numbers are on the upper x-axis.

correspond to pH 7. The sequences studied here include the following:

- A $\beta$ 42:  $^1\text{D}\text{AEFRHDSG}^{10}\text{YEVHHHQKL}V^{20}\text{FAEDVGSNKG}^{30}\text{AIIGLMVGGV}^{40}\text{VIA}$
- A $\beta$ 40:  $^1\text{D}\text{AEFRHDSG}^{10}\text{YEVHHHQKL}V^{20}\text{FAEDVGSNKG}^{30}\text{AIIGLMVGGV}^{40}\text{V}$

- A $\beta$ 42-M35ox(S):  $^1\text{D}\text{AEFRHDSG}^{10}\text{YEVHHHQKL}V^{20}\text{FAEDVGSNKG}^{30}\text{AIIGLXVG}^{40}\text{VIA}$  (where X = L-methionine-(S)-sulfoxide)
- A $\beta$ 40-E22 $\Delta$ :  $^1\text{D}\text{AEFRHDSG}^{10}\text{YEVHHHQKL}V^{20}\text{FA}-\text{DVGSNKG}^{30}\text{AIIGLMVGGV}^{40}\text{V}$  (the dash in this sequence represents the deleted 22nd residue; here, the numbering for this peptide keeps the residue numbers

consistent with the other peptides so index 22 is skipped; i.e., A21 is followed by D23)

- A $\beta$  4 0 - D 2 3 N :  $^1\text{D}\text{A}\text{E}\text{F}\text{R}\text{H}\text{D}\text{S}\text{G}^{1\text{0}}\text{Y}\text{E}\text{V}$ -HHQKLVF<sup>20</sup>FAENVGSNKG<sup>30</sup>AIIGLMVGGV<sup>40</sup>V
- A $\beta$  4 0 - E 2 2 K :  $^1\text{D}\text{A}\text{E}\text{F}\text{R}\text{H}\text{D}\text{S}\text{G}^{1\text{0}}\text{Y}\text{E}\text{V}$ -HHQKLVF<sup>20</sup>FAKDVGSNKG<sup>30</sup>AIIGLMVGGV<sup>40</sup>V
- A $\beta$  4 0 - E 2 2 G :  $^1\text{D}\text{A}\text{E}\text{F}\text{R}\text{H}\text{D}\text{S}\text{G}^{1\text{0}}\text{Y}\text{E}\text{V}$ -HHQKLVF<sup>20</sup>FAGDVGSNKG<sup>30</sup>AIIGLMVGGV<sup>40</sup>V
- A $\beta$  4 0 - A 2 T :  $^1\text{D}\text{T}\text{E}\text{F}\text{R}\text{H}\text{D}\text{S}\text{G}^{1\text{0}}\text{Y}\text{E}\text{V}$ -HHQKLVF<sup>20</sup>FAEDVGSNKG<sup>30</sup>AIIGLMVGGV<sup>40</sup>V

A $\beta$ 42, A $\beta$ 40, A $\beta$ 42-M35ox(S), A $\beta$ 40-E22 $\Delta$ , A $\beta$ 40-D23N, and A $\beta$ 40-E22K were simulated with OPLS, while A $\beta$ 42, A $\beta$ 40, A $\beta$ 40-E22 $\Delta$ , A $\beta$ 40-D23N, A $\beta$ 40-E22K, A $\beta$ 40-E22G, and A $\beta$ 40-A2T were investigated with ILDN. Only A $\beta$ 40 has been simulated using CHARMM parameters.

Besides the choice of force field and water model, all simulations were constructed and run identically to Rosenman et al.,<sup>48</sup> with a few exceptions listed in the next two paragraphs. These simulations use the same number of water molecules (4947 waters), box type and size (5.4 nm cubic box), equilibration protocol, and other REMD simulation parameters. These simulations involve 52 replicas, spanning the temperatures 270.0–601.2 K. These simulations were also run to 1000 ns/replica for all systems (cumulative simulation time of 52  $\mu$ s per system), which we previously observed was necessary to reach an equilibrated ensemble. All simulations were run using GROMACS 4.5.3 or 4.5.5.<sup>68</sup>

One exception to this construction for ILDN simulations is A $\beta$ 42. For this run only, we used the final configurations for each of the 52 replicas of the simulation described by Sgourakis et al.<sup>31</sup> as the starting state. The AMBER99sb/TIP4P-Ew combination of the Sgourakis et al. simulation is very similar to the AMBER99sb-ILDN/TIP4P-Ew combination used for our ILDN run for this peptide, and no evidence of simulation instability was observed upon switching the force field. This A $\beta$ 42 ILDN system alone was simulated to 750 ns/replica under the AMBER99sb-ILDN force field. All other REMD simulation conditions were identical to those used for the other simulations discussed here (and to those in the Sgourakis et al. predecessor).

Another set of exceptions are the CHARMM simulations of A $\beta$ 40. For this parameter combination, two separate REMD runs were conducted up to 500 ns/replica. To accommodate the CHARMM force field, the Coulombic and Lennard-Jones cutoffs were both changed from 1.0 nm to 1.2 nm, a switching potential for the Lennard-Jones forces was introduced between 1.0–1.2 nm, the short range neighbor list cutoff was changed from 1.0 nm to 1.2 nm, and a long range neighbor list cutoff was introduced at 1.4 nm. Besides changes in run time, force field, and the aforementioned parameters, the first of the REMD simulations uses an identical methodology to the other simulations; this run is denoted as “same box” or “SB”. The second simulation (called “big box” or “BB”) uses a larger box size of 5.9 nm and 6846 water molecules as opposed to a box size of 5.4 nm and 4947 waters used in all other A $\beta$  REMD runs (after NPT equilibration). The CHARMM BB version was run with 66 replicas ranging 270.0–603.3 K, as opposed to the 52 replicas ranging 270.0–601.2 K used by SB and all other simulations. The BB run was done to accommodate the more extended structures expected to be produced by CHARMM22\* and reduce the likelihood of periodic boundary contacts, while the SB run was done to observe how that change in box size

(and number of replicas) might affect the ensemble. Indeed, it was observed that the BB run had 0.1% of frames in the production period as being less than or equal to 0.5 nm away from its periodic image, while SB had 1.5% of frames fall into this category.

The REMD simulations were each performed using either 208 or 416 CPUs of a Linux cluster based at Rensselaer Polytechnic Institute or the Stampede cluster at the Texas Advanced Computing Center at the University of Texas as part of the Extreme Science and Engineering Discovery Environment (XSEDE).

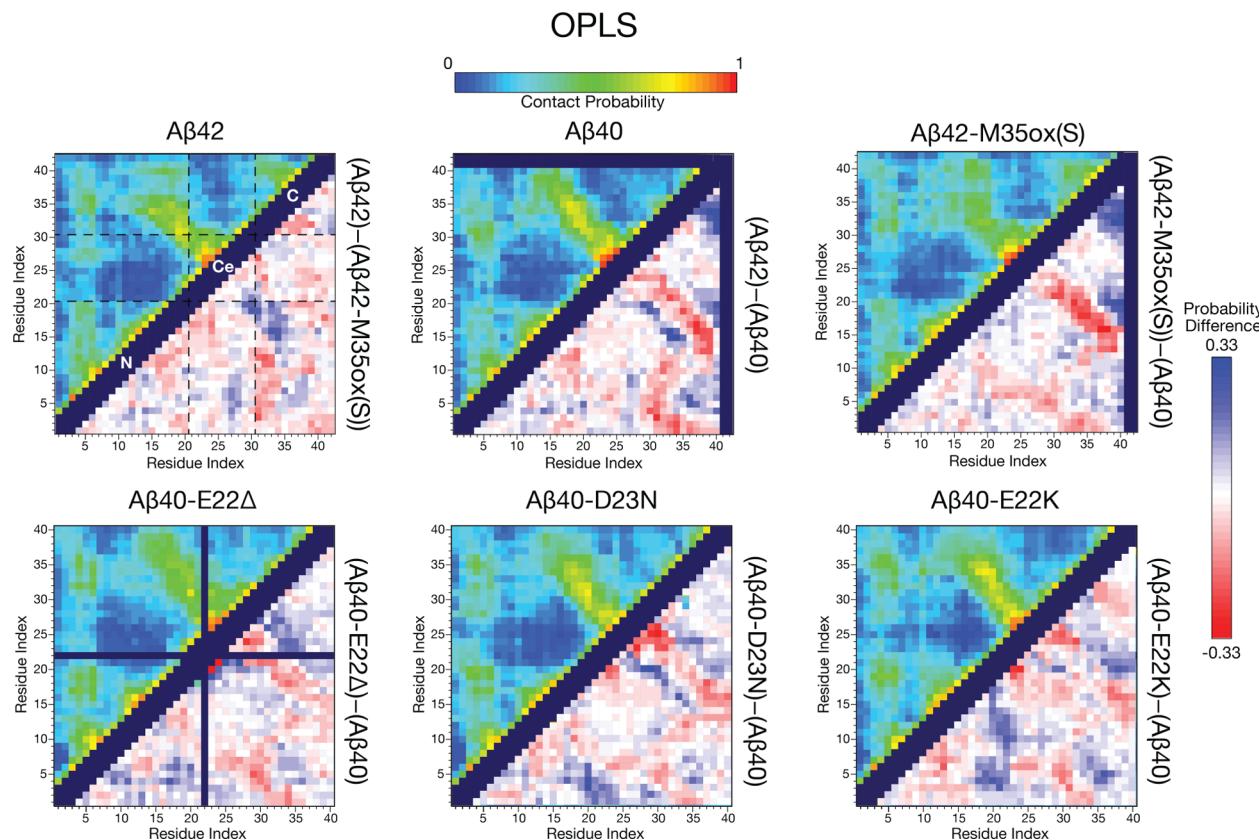
**2.2. Ensemble Analysis.** Analysis was done for all simulations over the following time intervals (referred to as the “production period” in the rest of this study): 400–750 ns/replica for ILDN A $\beta$ 42, 150–500 ns/replica for both CHARMM BB and SB simulations, and 400–1000 ns/replica for the other ILDN simulations. The enhanced sampling provided by the REMD protocol in tandem with the fact that analysis only begins after several hundred ns/replica into the run both mitigate the bias that comes from using the Sgourakis et al. starting conformation for the ILDN A $\beta$ 42 simulation. All of these ensembles were analyzed over 7 temperatures comprising the range 289.2–310.7 K; multiple temperatures were considered in order to enhance the sampling statistics. The only exception is CHARMM BB, which was analyzed over 7 temperatures ranging 290.1–309.1 K because of its different temperature set. The time intervals over the temperatures mentioned in this paragraph are referred to as the “production period” for each simulation.

Exceptions stated above notwithstanding, all REMD simulations were analyzed over the production period in a similar manner as previously described by Rosenman et al.<sup>48</sup> Analysis was done using various programs in the GROMACS 4.5.3 or 4.5.5 package,<sup>68</sup> in addition to Python and shell scripts. The DSSP program<sup>69,70</sup> was used to annotate secondary structures over each REMD simulation. Probability contact maps illustrate, for each pair of residues, the likelihood of contact over the production period, where contact is defined to mean that their C $\alpha$ –C $\alpha$  distance is  $\leq$ 8.0 Å.<sup>71</sup> Hydrogen bonds were defined such that the atomic distance for the donor to acceptor is  $\leq$ 3.5 Å, while the angle between the hydrogen, donor, and acceptor must not exceed 60°. Salt bridges must occur between two oppositely charged atoms within 4.0 Å of each other.

### 3. RESULTS

**3.1. OPLS. 3.1.1. Secondary Structure.** Figure 1A maps the probability of each DSSP<sup>69,70</sup> secondary structure annotation for each of the A $\beta$  species over all residues and the production period. Overall, all of the A $\beta$  ensembles are observed to possess similar profiles. The propensity to form  $\beta$  sheets is the most variable element, having a ranking A $\beta$ 40-D23N > A $\beta$ 40 > A $\beta$ 40-E22K  $\gg$  A $\beta$ 40-E22 $\Delta$  > A $\beta$ 42 > A $\beta$ 42-M35ox(S). Bend and coil are the next most variable components; both largely compensate for the changes seen in  $\beta$  sheet content.

Illustrated in Figure 1B,C are the secondary structure contents per residue for each A $\beta$  variant averaged over the ensemble. We observed in simulations of A $\beta$ 40 that  $\beta$  character is robust in residues L17–A21 and A30–V36, while the intervening V24–K28 is dominated by bend character. This is in line with sampling of a  $\beta$ -hairpin-like motif, an observation supported by the biases in the contact maps reviewed in the next subsection. However, the different FAD mutants perturb



**Figure 2.** (Upper Corner) Contact probability maps for C $\alpha$  atoms for the production ensemble of simulations of different A $\beta$  isoforms and mutants done with REMD under the OPLS condition. Two residues are in contact on this map if their C $\alpha$ -C $\alpha$  distance is  $\leq 8.0 \text{ \AA}$ .<sup>71</sup> Regions of A $\beta$ , as defined in the text, are illustrated on the first contact map (A $\beta$ 42) (N = N-terminal, Ce = central, C = C-terminal). (Lower Corner) Difference between contact probability maps for selected pairs of A $\beta$  species with the OPLS condition, as listed in the vertical axis title for each map. Red colors indicate reduced contact probability for the minuend form compared to the subtrahend form, while blue colors indicate an increased likelihood of contact. For example, for the first difference contact map, larger probabilities for A $\beta$ 42-M35ox(S) are colored in red, and those for A $\beta$ 42 are in blue. Dark purple squares represent self or sequential contacts such as those between (i, i), (i, i + 1), and (i, i + 2) and invalid squares on the map that were not computed.

this profile; indeed, observation of the distribution of  $\beta$  sheet content in Figure 1C is considerably altered by the different A $\beta$  variants. A $\beta$ 40-D23N is the most disruptive mutant, significantly increasing  $\beta$  content in residues F19-D/N23 and K28-I31, while the content of bend is decreased throughout the V24-K28 region, particularly in residue K28. A $\beta$ 40-E22K increases  $\beta$  propensity in A21-D23, while A $\beta$ 40-E22 $\Delta$  reduces  $\beta$  content in the central hydrophobic cluster (CHC, L17-A21). Both E22 mutants, particularly E22 $\Delta$ , reduce  $\beta$  character in A30-V36, particularly among the earlier residues of that group. Consequentially, bend character is increased in G29-A30, opposite the effect of D23N. A $\beta$ 42 and A $\beta$ 42-M35ox(S) also somewhat weaken these features (particularly the latter). The main change seen in these two species compared to A $\beta$ 40, however, is that they demonstrate substantial probability to sample  $\beta$  sheets at residues V39-I41 and a bend at G37-G38. This suggests that the C-terminus of these two species contributes toward new  $\beta$  topologies in the ensemble.

In general, the N-terminus (D1-Q15) features substantially more coil/irregular content in OPLS simulations, though E3-H6 and Y10-H13 can populate  $\beta$  sheets (up to  $\sim 30\%$  in A $\beta$ 40-E22K and A $\beta$ 42-M35ox(S)) while H6-G9 and H14-K16 contribute considerable bend/turn character. Lastly,  $\alpha$  helices are conspicuously absent over the A $\beta$  sequence; only minor

populations (<20% when combining both 3 helix and  $\alpha$  helix) are observed at H13-L17 for all A $\beta$  ensembles.

**3.1.2. Tertiary Structure.** The maps in the upper corner of Figure 2 display the probabilities of C $\alpha$ -C $\alpha$  contact between all pairs of residues for different A $\beta$  alloforms and mutants. These maps reveal a characteristic tertiary structure of A $\beta$  defined by three different regions: N-terminal (residues D1-F20), central (residues A21-A30), and C-terminal (residues I31-V40/A42). Defining the regions in this way shows that there are frequent contacts between residues of each region, except the C-terminus in A $\beta$ 40 and its point mutants. However, while the N-terminal and C-terminal regions interact frequently, the central region sparsely interacts with the other two regions, except the distal N-terminus (residues F4-H6). As we have covered in this previous paragraph, this central region is rich in bend character (especially V24-K28) and bridges two  $\beta$  sheet forming regions (the CHC and A30-V36); its role as a loop in  $\beta$  hairpin structures in the ensemble would explain its restriction from the rest of the peptide.

For A $\beta$ 40, we further observe high probability antidiagonal contacts between Q15-G25 and N27-V39 (most populated between K16-F20 and A30-M35), further suggesting the sampling of antiparallel  $\beta$  hairpins between these sections. For ease of reference, we term this the “main antidiagonal”. Difference contact maps (Figure 2, lower corner) reveal

**Table 1. Central Region Electrostatic Interactions of  $\text{A}\beta$  Involving the Side Chains of E22 and D23 for the OPLS Simulation Set<sup>a</sup>**

| donor–acceptor atoms                   | % for OPLS $\text{A}\beta42$             | % for OPLS $\text{A}\beta40$             | % for OPLS $\text{A}\beta42\text{-M35ox(S)}$ |
|--|--|--|--|
| Hydrogen Bonds                         |  |  |  |
| V24NH–D/N23O <sub>δ</sub>              | 27                                       | 32                                       | 25   |
| G25NH–D/N23O <sub>δ</sub>              | 57                                       | 61                                       | 51   |
| S26NH–D/N23O <sub>δ</sub>              | 59                                       | 65                                       | 54   |
| S26H <sub>γ</sub> –D/N23O <sub>δ</sub> | 32                                       | 37                                       | 32   |
| N27NH–D/N23O <sub>δ</sub>              | 31                                       | 36                                       | 29   |
| K28NH–D/N23O <sub>δ</sub>              | 17                                       | 25                                       | 18   |
| Salt Bridges                           |  |  |  |
| K28N <sub>ζ</sub> –E/K22O <sub>ε</sub> | 10                                       | 23                                       | 20   |
| K28N <sub>ζ</sub> –D/N23O <sub>δ</sub> | 4  | 2  | 5  |
| donor–acceptor atoms                   | % for OPLS $\text{A}\beta40\text{-E22Δ}$ | % for OPLS $\text{A}\beta40\text{-D23N}$ | % for OPLS $\text{A}\beta40\text{-E22K}$     |
| Hydrogen Bonds                         |  |  |  |
| V24NH–D/N23O <sub>δ</sub>              | 29                                       | 7  | 25   |
| G25NH–D/N23O <sub>δ</sub>              | 57                                       | 26                                       | 52   |
| S26NH–D/N23O <sub>δ</sub>              | 61                                       | 4  | 50   |
| S26H <sub>γ</sub> –D/N23O <sub>δ</sub> | 39                                       | 0  | 32   |
| N27NH–D/N23O <sub>δ</sub>              | 33                                       | 4  | 25   |
| K28NH–D/N23O <sub>δ</sub>              | 14                                       | 0  | 13   |
| Salt Bridges                           |  |  |  |
| K28N <sub>ζ</sub> –E/K22O <sub>ε</sub> | N/A                                      | 8  | N/A  |
| K28N <sub>ζ</sub> –D/N23O <sub>δ</sub> | 2  | 2  | 5  |

<sup>a</sup>Listed populations represent average bond frequencies for equivalent atoms in the side chain (for example, “D23O<sub>δ</sub>” is the average of interactions involving the D23O<sub>δ1</sub> and D23O<sub>δ2</sub> atoms). Because the D23N mutation eliminates the salt bridge, for  $\text{A}\beta40\text{-D23N}$  only, populations for the interaction between K28N<sub>ζ</sub> and N23O<sub>δ</sub> in the final row follow the hydrogen bond criteria, not salt bridge criteria.

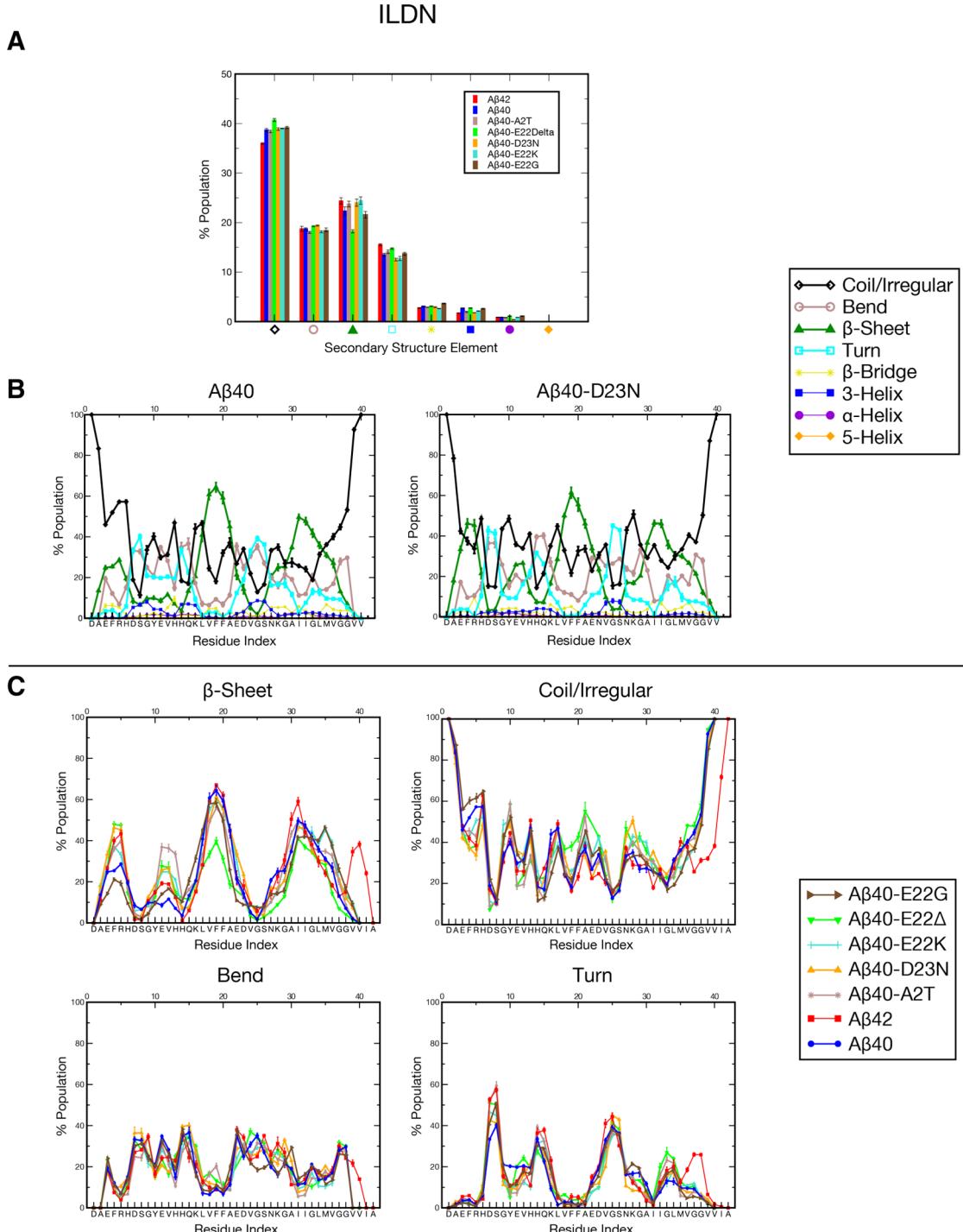
important discrepancies in the interactions constituting the main antidiagonal among the different variants of  $\text{A}\beta$ . A common theme among the FAD mutants of  $\text{A}\beta$  is that these often reduce contact probabilities between these regions on the diagonal compared to those of the wild type, but increase contact probabilities in diagonals flanking the center diagonal (manifesting as a blue–red–blue stripe pattern in the difference maps), effectively increasing the “line width” of the main antidiagonal. This suggests that all of the FAD mutants increase the polymorphism in possible binding partners involved in antiparallel binding. Further, D23N, in particular, preferably shifts the main diagonal toward the N-terminus by one or two residues, while the two E22 mutants favor more C-terminal binding diagonals. This corresponds to changes in the breadth of the bend regions as discussed in the previous subsection, with D23N possessing a narrower bend region and the two E22 mutants broadening the bend to include G29–A30.

$\text{A}\beta42$  and  $\text{A}\beta42\text{-M35ox(S)}$  also have a similar main antidiagonal, but its C-terminal partner is shortened compared to that of  $\text{A}\beta40$  such that the contacts are only highly populated between Q15–G25 and N27–V36. Further, the 42 residue variants have a new set of antiparallel contacts between I31–V36 and G37–I41, suggesting, in tandem with the secondary structure characterization above, that these regions are also prone to forming  $\beta$  hairpins. Q15–A21 also have a tendency to contact the G37–I41 region in both reduced and oxidized  $\text{A}\beta42$ . For  $\text{A}\beta42\text{-M35ox(S)}$ , these contacts have higher probability compared to  $\text{A}\beta42$  while the main antidiagonal is reduced in population, which suggest a rearrangement of  $\beta$  compatible regions. Lastly, like FAD mutants, the pairing of contacts is altered in  $\text{A}\beta42$ . However, the pattern of the difference contact maps reveals a register shift with a bias for  $\text{A}\beta42$  to sample a more N-terminal partner for the N27–V36 participants (a blue–red pattern in the difference contact map).

$\text{A}\beta42\text{-M35ox(S)}$ ’s map does not demonstrate this register shift; instead, its cross-central contacts are simply reduced in probability compared to  $\text{A}\beta40$ .

The extreme N-terminus, including E3–D7, is observed to have weak interactions with the whole peptide, even at the central region. These contacts are particularly well-populated for both the CHC and G29–V36. These long-range contacts are, in general, promoted by all of the FAD mutants, particularly E22K, which strongly promotes contacts to V18–E/K22 by most residues of the N-terminal domain (E3–L17, see the blue patch centered around residue 20 in the difference map for  $\text{A}\beta40\text{-E22K}$ ), and D23N, which promotes anti-diagonal contacts between the extreme N-terminus and F20–G25.  $\text{A}\beta40\text{-E22Δ}$ , in contrast, primarily enhances contacts between the N-terminus and other residues in the N-terminal region and CHC, while the wild type favors interactions with G29–I31. These features, in tandem with the increased breadth of binding partners across the main diagonal, represent increased long-range contact of the central region to N-terminal and C-terminal regions permitted by FAD mutants.

**3.1.3. Central Region Electrostatic Interactions.** Populations for electrostatic interactions involving the FAD-linked side chains of E22 and D23 are listed in Table 1. In the OPLS simulation of  $\text{A}\beta40$ , it is observed that the side chain carboxyl of D23 frequently populates hydrogen bonds with the backbone amides of subsequent residues V24–K28, with the bonds to G25 and S26’s amides being present in more than 60% of the frames of the ensemble. These bonds are, indeed, well-populated in all OPLS simulations, with the notable exception of  $\text{A}\beta40\text{-D23N}$ . In this mutant, these interactions are all greatly diminished, though local interactions between the polar asparagine O<sub>δ</sub> atom and the amides of V24 and G25 are still present (7% and 24%). Interactions involving more distant



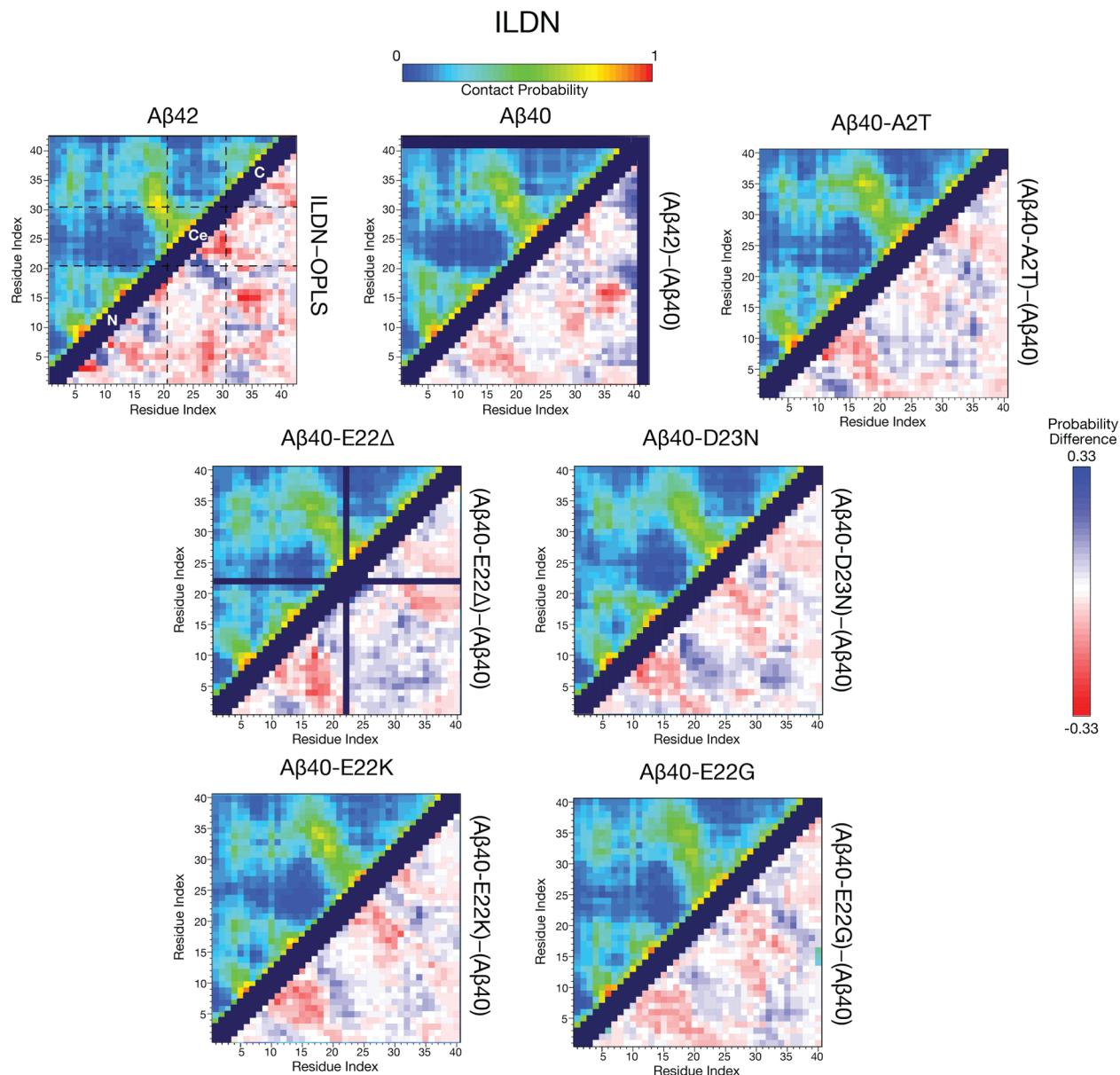
**Figure 3.** Percent secondary structure for  $\text{A}\beta$ , as annotated by DSSP,<sup>69,70</sup> averaged over the production ensemble, for simulations done with the ILDN condition (A) over all residues, and (B and C) as a function of residue index. All other details are the same as OPLS simulations in Figure 1.

residues, including S26–K28's amides and S26's side chain, are rare or not observed.

E22 forms a salt bridge with K28 across the central V24–K28 bend with a high probability of 23% over the wild type  $\text{A}\beta40$  ensemble. This salt bridge cannot be formed in the E22 $\Delta$  or E22K mutants, while it is rare in the D23N variant (8%). Meanwhile, salt bridges of K28 involving residue 23 are rare (5% or less in all simulations), suggesting this residue does not substitute as a salt bridge partner in any of the forms of  $\text{A}\beta$  (even in the E22 FAD mutants). Interestingly, while the E22–K28 salt bridge is present in  $\text{A}\beta42$ –M3Sox(S) in 20% of the

frames, roughly the same percentage as  $\text{A}\beta40$ , this salt bridge is somewhat weaker in the wild type  $\text{A}\beta42$  variant (10%). Therefore, loss of the E22–K28 salt bridge has an apparent correlation with aggregation propensity since it is weakened in the  $\text{A}\beta42$  and  $\text{A}\beta40$ –D23N simulations and absent for the E22 mutants of  $\text{A}\beta40$ .

For  $\text{A}\beta40$ , these E22 and D23 electrostatic interactions contribute to a feature observed in representative structures of the ensemble where V24–K28 amide groups point inward to interact with D23's side chain to form a bulge-like bend with outward pointing carbonyls, which we called the “crown motif”



**Figure 4.** (Upper Corner) Contact probability maps for C $\alpha$  atoms for the production ensemble of simulations of different A $\beta$  isoforms and mutants done with REMD under the ILDN condition. (Lower Corner) Difference between contact probability maps for selected pairs of A $\beta$  species with the ILDN condition. All other details are the same as OPLS simulations in Figure 2.

and illustrated previously.<sup>48</sup> Again, we saw that the A $\beta$ 40-D23N simulation severely reduces the bonds involving residue 23, even after considering interactions with the still polar O $\delta$  present in the asparagine variant. Further, none of its representative cluster centroid structures are observed to have “crown” features in the central turn, instead forming alternate bend structures that are shorter in breadth.<sup>28</sup> In contrast, neither of the E22 mutants can form the E22 to K28 salt bridge, yet they sample centroid structures with crown features.<sup>28</sup> All of this suggests the amino acid identity of D23 is critical to forming the crown, and only the D23N mutation strongly reduces the tendency of crown formation for OPLS simulations.

Lastly, we note that while A $\beta$ 40-E22K introduces a new cation in the central region, it does not interact with D23 (population <1%); among its possible partners, it interacts with D1 and D7 the most (population ~6% and ~4%, respectively).

These ionic interactions may drive the increased interaction of E22K's CHC with the N-terminus.

**3.2. ILDN.** **3.2.1. Secondary Structure.** Reported in Figure 3A is the overall secondary structure content based on DSSP<sup>69,70</sup> annotation for the production period of each A $\beta$  ILDN simulation. These A $\beta$  ensembles are even more similar in terms of secondary structure profile than the OPLS ensemble (compare to Figure 1). Ranking for  $\beta$  structure is A $\beta$ 40-E22K > A $\beta$ 42 > A $\beta$ 40-D23N > A $\beta$ 40-A2T > A $\beta$ 40  $\gg$  A $\beta$ 40-E22 $\Delta$ , with coil, bend, and turn structures making up the difference between these species. The most striking difference in secondary structure distributions between OPLS and ILDN conditions are an exchange of bend structure in OPLS for turn character in ILDN. Bend, turn, and irregular are each different types of loop structures in DSSP nomenclature; however, DSSP defines bends to possess high curvature (that is, a direction change around the annotated residue of more than 70°), and

DSSP turns also possess high curvature but also must possess helical-like hydrogen bonds (that is, a hydrogen bond from carbonyl oxygen of residue  $i$  to the amide proton of residue  $i+n$ , where  $n = \{3, 4, 5\}$ ), while regions annotated as DSSP irregular are loops that lack either of these features.<sup>69</sup> In our discussion, we will define a “loop”-like structure in this paper to encompass both bend and turn secondary structures.

**Figure 3B,C** depicts the secondary structure propensity as a function of residue index over each ILDN ensemble. Like OPLS (compare to **Figure 1B,C**), ILDN simulations are characterized by strong  $\beta$  character in the CHC and A30–L34 regions separated by a loop-like motif, with comparable population percentages between the force fields. Among the wild type simulations, the equivalent of that A30-L34 $\beta$  region is actually considerably wider in OPLS, possessing two maxima at I31–I32 and L34–M35 as opposed to ILDN’s single maximum at I31. The FAD mutations are less prone to perturbing these peaks under the ILDN condition, versus the large degree of polymorphism seen with OPLS. However,  $\text{A}\beta40\text{-E}22\Delta$  is an exception in that it strongly reduces the  $\beta$  sheet content of both regions, while both  $\text{A}\beta40\text{-E}22\text{K}$  and  $\text{A}\beta40\text{-E}22\text{G}$  expand  $\beta$  content to include M35 and V36. Further, the E22G mutation has a prominent effect on the  $\beta$  distribution of the peptide, primarily shifting the CHC  $\beta$  sheet peak N-terminally, while biasing the peak normally among residues A30–V36 toward the C-terminus. In other words, E11–Q15 and L34–V39 are more prone to be involved in  $\beta$  sheets in this variant, while F18–D23 and N27–I32 lose  $\beta$  structure. To compensate, the central region (as defined in the previous section) is apparently broadened in  $\text{A}\beta40\text{-E}22\text{G}$  so that in fringe regions (F19–D23 and G29–A30)  $\beta$  sheet character is replaced by coil/irregular character, while the center of the region (V24–S26) supplants bend character for an increase in  $\alpha$  helix annotations; interestingly, however, the distribution of turn is mostly unaffected. Finally,  $\text{A}\beta42$  actually promotes  $\beta$  character in these regions, and, like OPLS, samples a new maximum involving V39–I41 with an intervening loop region dominated by G37 and G38, suggesting the creation of new C-terminal  $\beta$  topology.

Some of the ILDN simulations also show considerably more  $\beta$  character in the N-terminus than OPLS, with substantial content at residues E3–H6 and Y10–H13. For ILDN, all of the point mutations of  $\text{A}\beta40$ , in particular, enhance  $\beta$  content in these regions and turn content in the intervening residues. In contrast,  $\text{A}\beta40$  is more likely to sample turn and bend in this region. This suggests that FAD mutants may promote the involvement of the N-terminus in new  $\beta$  topologies not observed in the wild type.

The D7–G9, A21–G29, and G37–G38 regions are also different between the force fields. Both OPLS and ILDN conditions characterize these regions in  $\text{A}\beta$  as looplike, but OPLS is characterized predominantly by bend structure while ILDN contains significantly more turn and irregular character in these regions. The increased heterogeneity of structures over the ensemble suggests that these regions are considerably more disordered in nature in the ILDN force field set. Unlike OPLS, in general, the profile of the central region of  $\text{A}\beta40$  is not strongly perturbed by mutations in ILDN, though subtle changes are present. For example, there is reduced turn content in A21–D/N23 and bend in E/K22 in all FAD mutants compared to wild type  $\text{A}\beta40$  and  $\text{A}\beta42$ , in exchange for increased  $\beta$  and coil character.  $\text{A}\beta40$  also samples the largest turn character in K28–A30 of the different species, while  $\text{A}\beta40$ –

D23N samples reduced turn content (but more bend and coil content) in this region. The effect is that the distribution of the turn character is narrower in FAD mutants, with the distribution for  $\text{A}\beta40\text{-D}23\text{N}$  having the most obvious difference. Lastly, like OPLS,  $\alpha$ -helical character is sparse in ILDN, except for a small tendency (<10%) to form 3-helices in D7–G9, H14–K16, and V24–N27.

**3.2.2. Tertiary Structure.**  $\text{Ca}$  contact maps (**Figure 4**, upper corner) and difference contact maps (**Figure 4**, lower corner) were calculated for all ILDN simulations, revealing biases in tertiary structure of the  $\text{A}\beta$  ensembles very similar to the OPLS condition in nature (compare to **Figure 2**). Common to both force field conditions are the three regions of the contact map, as described previously for OPLS: N-terminal (residues D1–F20), central (residues A21–A30), and C-terminal (residues I31–V40/A42), wherein the central region very rarely contacts the other two regions in wild type  $\text{A}\beta40$ . Changes in contact probabilities across the different mutants and isoforms of  $\text{A}\beta$  are subtle with the ILDN condition, even more so than those than those noted for OPLS, but grant important insight into the structural differences between the simulated ensembles.

Characteristic of ILDN simulations are prevalent antidiagonal contacts between K16–G25 and N27–I32. Residues G33–G37 also interact with the CHC region, but contacts that extend the antidiagonal farther to involve residues N-terminal of L17 or C-terminal of V36 are much sparser than in the OPLS condition. FAD mutants are also far less prone to perturb the register of binding across this antidiagonal among the ILDN simulations. Each of the point mutants of  $\text{A}\beta40$  studied here have a common effect of promoting a register shift for the main antidiagonal where more C-terminal binding partners are favored, except E22 $\Delta$ , which merely reduces the contact probability of this hairpin in general. This register shift, however, is much less prevalent than the shifts in diagonal binding discovered in OPLS simulations. These antidiagonal contacts also exist in  $\text{A}\beta42$  but are curtailed to L17–G25 and N27–I32, with G33–V36 also interacting with L17–A21 (but not extending the antidiagonal to more N-terminal residues). There is no register shift in the contacts of the main antidiagonal compared to  $\text{A}\beta40$ . Analogous to OPLS,  $\text{A}\beta42$  is prone to sampling a second set of antidiagonal contacts between I31–V36 and G37–I41, though these are somewhat more robust in the OPLS condition. Also prevalent are contacts between the CHC and the extreme C-terminus, with higher probability than OPLS. Subtle differences aside, the two force fields share a strong commonality in that they both involve the extreme terminus in a second hairpin that can join with the central hairpin to form a  $\beta$  meander (that is, 2 subsequent antiparallel  $\beta$  hairpins). Sampling of this extra C-terminal hairpin may be the possible basis for increased rigidity observed at the C-terminus in monomeric  $\text{A}\beta42$ ,<sup>72</sup> while a similar structural feature is observed in the intrapeptide model for neurotoxic oligomers.<sup>73</sup> Thus, this motif may be implicated in the increased toxicity and faster aggregation of the  $\text{A}\beta42$  alloform.

Like OPLS, the N-terminus also plays an active role in contacting the rest of the peptide for the ILDN condition, particularly at the CHC and the I31–V36 region. The N-terminus is less prone to contact the central region (particularly in wild type  $\text{A}\beta$ ), while it is more prone to form antidiagonals between A/T2–D7 and S8–H14 and between V12–K16 and the CHC. The point mutants of  $\text{A}\beta40$  studied here have a strong effect on the interactivity of this region. In common, all

mutants reduced contact between D1–Y10 and the CHC, but increased contact between the extreme N-terminal residues and the following regions: Y10–Q15 (strongest increase in A2T), F20–G24 (strongest increase in D23N and E22K), and I31–M35 (strongest increase in D23N and E22Δ). This suggests that these mutants are more likely to form additional  $\beta$  topologies involving the extreme N-terminus, in contrast to local, turn-like structures in wild type  $\text{A}\beta$ 40. Further, the increased contact of the N-terminus to F20–G24, in tandem with the increased likelihood to sample more C-terminal binding partners across the main diagonal, represents ways in which these point mutants promote long-range contacts to the normally isolated central region.

**3.2.3. Central Region Electrostatic Interactions.** Table 2 illustrates the populations of electrostatic interactions involving

**Table 2. Central Region Electrostatic Interactions of  $\text{A}\beta$  Involving the Side Chains of E22 and D23 for the ILDN Simulation Set<sup>a</sup>**

| donor–acceptor atoms                  | % for ILDN $\text{A}\beta$ 42      | % for ILDN $\text{A}\beta$ 40      | % for ILDN $\text{A}\beta$ 40-A2T  |
|---------------------------------------|------------------------------------|------------------------------------|------------------------------------|
| Hydrogen Bonds                        |                                    |                                    |                                    |
| V24NH–D/N23O $_{\delta}$              | 9                                  | 8                                  | 7                                  |
| G25NH–D/N23O $_{\delta}$              | 9                                  | 10                                 | 10                                 |
| S26NH–D/N23O $_{\delta}$              | 4                                  | 5                                  | 3                                  |
| S26H $_{\gamma}$ –D/N23O $_{\delta}$  | 3                                  | 6                                  | 4                                  |
| N27NH–D/N23O $_{\delta}$              | 4                                  | 3                                  | 1                                  |
| K28NH–D/N23O $_{\delta}$              | 2                                  | 1                                  | 1                                  |
| Salt Bridges                          |                                    |                                    |                                    |
| K28N $_{\zeta}$ –E/K22O $_{\epsilon}$ | 4                                  | 5                                  | 5                                  |
| K28N $_{\zeta}$ –D/N23O $_{\delta}$   | 4                                  | 5                                  | 4                                  |
| donor–acceptor atoms                  | % for ILDN $\text{A}\beta$ 40-E22Δ | % for ILDN $\text{A}\beta$ 40-D23N | % for ILDN $\text{A}\beta$ 40-E22K |
| Hydrogen Bonds                        |                                    |                                    |                                    |
| V24NH–D/N23O $_{\delta}$              | 12                                 | 13                                 | 11                                 |
| G25NH–D/N23O $_{\delta}$              | 11                                 | 11                                 | 11                                 |
| S26NH–D/N23O $_{\delta}$              | 4                                  | 2                                  | 4                                  |
| S26H $_{\gamma}$ –D/N23O $_{\delta}$  | 5                                  | 0                                  | 6                                  |
| N27NH–D/N23O $_{\delta}$              | 5                                  | 1                                  | 2                                  |
| K28NH–D/N23O $_{\delta}$              | 5                                  | 1                                  | 2                                  |
| Salt Bridges                          |                                    |                                    |                                    |
| K28N $_{\zeta}$ –E/K22O $_{\epsilon}$ | N/A                                | 2                                  | N/A                                |
| K28N $_{\zeta}$ –D/N23O $_{\delta}$   | 3                                  | 2                                  | 2                                  |
| donor–acceptor atoms                  | % for ILDN $\text{A}\beta$ 40-E22G |                                    |                                    |
| Hydrogen Bonds                        |                                    |                                    |                                    |
| V24NH–D/N23O $_{\delta}$              | 7                                  |                                    |                                    |
| G25NH–D/N23O $_{\delta}$              | 11                                 |                                    |                                    |
| S26NH–D/N23O $_{\delta}$              | 7                                  |                                    |                                    |
| S26H $_{\gamma}$ –D/N23O $_{\delta}$  | 9                                  |                                    |                                    |
| N27NH–D/N23O $_{\delta}$              | 1                                  |                                    |                                    |
| K28NH–D/N23O $_{\delta}$              | 0                                  |                                    |                                    |
| Salt Bridges                          |                                    |                                    |                                    |
| K28N $_{\zeta}$ –E/K22O $_{\epsilon}$ |                                    | N/A                                |                                    |
| K28N $_{\zeta}$ –D/N23O $_{\delta}$   |                                    | 4                                  |                                    |

<sup>a</sup>All other details are the same as OPLS simulations in Table 1.

FAD-linked residue 22 and 23's side chains in each of the ILDN simulations. Compared to the OPLS simulations where these were highly prevalent, ILDN simulations have a much smaller likelihood for D23 to interact with V24–K28 through buried side chain to backbone hydrogen bonds and for E22 to interact with K28 through a cross-region salt bridge. These

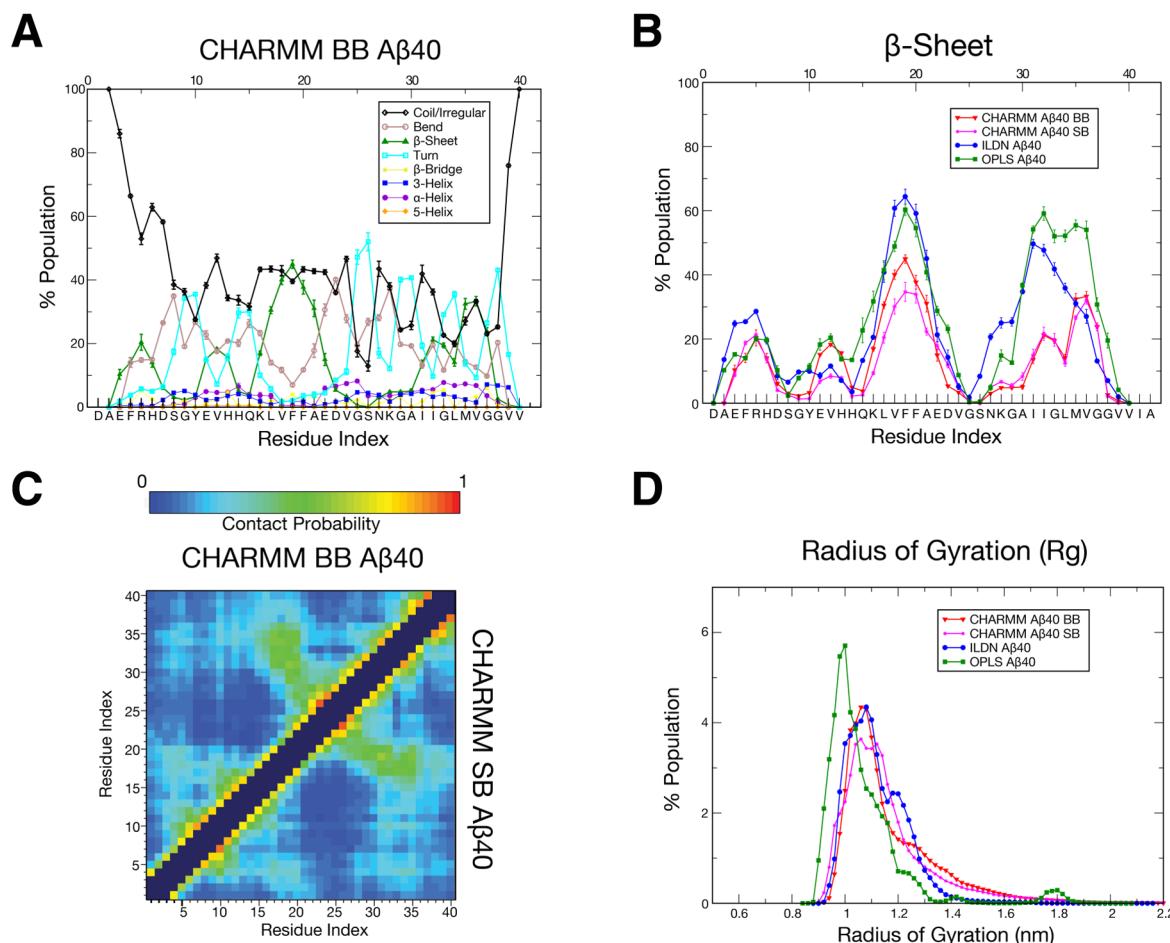
interactions do not appear to be replaced by interactions involving other parts of the peptide; instead, these residues' side chains interact with solvent with a much higher proclivity. Further, with the S26 and N27 relatively free from interacting with the D23 side chain, ILDN  $\text{A}\beta$ , unlike OPLS, is much more prone to sample helical-like backbone hydrogen bonds at the central region turn involving (in donor amide/acceptor carbonyl nomenclature) S26/D (or N) 23, N27/D (or N) 23, and N27/V24 (populated 17%, 12%, and 20% of the ILDN  $\text{A}\beta$ 40 ensemble, but only 3%, 3%, and 0% in OPLS  $\text{A}\beta$ 40, respectively). The presence of these new bonds may therefore explain the turn in secondary structure annotation in this region as described in the Secondary Structure subsection above, and why it is absent in OPLS.

That said, these OPLS "crown motif" interactions are not completely abolished even in these ILDN simulations; indeed, local interactions between D/N23's side chain O $_{\delta}$  and V24 and G25's amide are populated 7–13% in all simulations, while interactions involving S26's amide and side chain are present to 3–6% except in  $\text{A}\beta$ 40-D23N. Further, salt bridges with E22 or D23 as the anion and K28 as the cation are present in wild type  $\text{A}\beta$  simulations up to 5%. FAD mutants reduce the prevalence of these salt bridges to varying degrees. All three mutants of E22 trivially eliminate the E22–K28 interaction, but the D23–K28 interaction is also reduced in each case (to 3% in E22K, 2% in E22Δ, and 4% in E22G), while  $\text{A}\beta$ 40-D23N demonstrates reduced E22–K28 salt bridge occupancy, as well as low populations for the interaction with between K28's H $_{\zeta}$  and N23's O $_{\delta}$  (2% in each case). Unlike OPLS, residue 22 of the E22K variant interacts with D23 in 5% of the ensemble; no other significantly populated anionic partners exist for this lysine. In general, we conclude that the reduction of the "crown motif"-affiliated central region electrostatic interactions permits the increased helical hydrogen bonding and greater secondary structure diversity seen in all simulations of the ILDN condition at the central region.

**3.3. CHARMM.** The CHARMM combination is the last force field compared here: only  $\text{A}\beta$ 40 was simulated with this condition, using two different box sizes and replica exchange temperature sets (CHARMM BB [big box] and CHARMM SB [small box], as defined in the Methods section).

DSSP<sup>69,70</sup> annotations of secondary structure for the two CHARMM simulations as a function of residue are presented in Figure 5A,B. Both CHARMM simulations do produce a profile for the first half of the peptide that is similar to the other force fields, particularly, a CHC that samples substantial  $\beta$  character (up to ~45% at the apex at F19 for CHARMM BB, which is less than the ~60% for both OPLS and ILDN, but still substantial), some  $\beta$  bias in E3–H6 and E11–H13, and turn character particularly in G9–Y10 and H14–Q15. CHARMM SB formed substantially less  $\beta$  character in these CHC and E11–H13 peaks, but otherwise possesses very similar secondary structure to the BB variant throughout the peptide. For both CHARMM simulations, the central motif is also prone to forming irregular, turn, and coil structures, with G25–S26 substantially biased toward turn like the ILDN simulation (though V24 is also characterized by turn for ILDN, but not CHARMM). Where CHARMM simulations primarily differ from OPLS and ILDN is the characterization of more C-terminal regions. In addition to the G25–S26 turn, G29–A30, G33–L34, and G37–G38 are substantially more populated with turn character than either OPLS and ILDN, which both tend to populate these regions with more bend (particularly for

## CHARMM



**Figure 5.** Analysis of A $\beta$ 40, simulated with REMD using the CHARMM condition. (A) Percent composition of secondary structure as determined by DSSP<sup>69,70</sup> for CHARMM A $\beta$ 40 BB, averaged over the production ensemble. Compare to Figures 1 and 4. (B) Percent composition of  $\beta$  sheet alone for A $\beta$ 40 simulated over multiple force fields. The sequence of wild type A $\beta$ 42 is on the lower x-axis, while residue numbers are on the upper x-axis. (C) Map illustrating the probability of contact over the production ensemble for CHARMM A $\beta$ 40 BB in the upper corner and CHARMM A $\beta$ 40 SB in the lower corner. Here, two residues are defined to be in contact if their C $\alpha$ –C $\alpha$  distance is  $\leq 8 \text{ \AA}$ .<sup>71</sup> Compare with Figures 2 and 4. (D) Histogram for radii of gyration ( $R_g$ ) for A $\beta$ 40 simulated with multiple force fields, using 0.02-nm-wide bins.

G29–A30 and G37–G38) or  $\beta$  character (G33–L34). The new turn regions in CHARMM are predominantly interspersed with irregular/coil character. By DSSP definition, the presence of strongly populated turn regions in the C-terminus imply the presence of helical-like hydrogen bonds, and indeed, CHARMM populates the following bonds well (in donor amide/acceptor carbonyl nomenclature): S26/D23 (populated 36% in CHARMM BB), A30/N27 (29%), L34/I31 (22%), M35/I31 (12%), and G38/M35 (30%). Further, the spacing of the apex of these turn regions occur in an  $i$  to  $i + 4$  pattern, suggesting a bias to helical-like sampling. However, the presence of canonical DSSP  $\alpha$ -helices or 3-helices is sparing, with no residue sampling more than 10% of either character in the BB simulation. Lastly, we note that the pattern for these four C-terminal turns align to three GxxG motifs in the A $\beta$  sequence, with turns starting at G25, G29, G33, and G37. The exception to all of these trends in the C-terminus is a substantial peak of  $\beta$  character, which is strongest in M35–G37 at  $\sim 30\%$ . This C-terminal peak is very different from that of OPLS and ILDN, however, which both strongly characterize A30–L34 with  $\beta$  character, reaching  $\sim 50$ –60% at I31–I32 (these residues are <20%  $\beta$  character in both CHARMM

simulations, in contrast). The similarity of contact maps but differences in secondary structure biases suggest that the nature of backbone hydrogen bonding is substantially different in the CHARMM simulations.

Interestingly, despite the differences in secondary structure, Figure 5C shows that both CHARMM BB and CHARMM SB simulations produce an ensemble-averaged contact probability map similar in profile to A $\beta$ 40 simulated with other force fields (compare with Figure 2 and Figure 4), all possessing contacts between the central hydrophobic cluster (CHC, L17–A21) and C-terminus and a central region that is isolated from rest of peptide. This lends further support to the ensemble characterization of A $\beta$  made so far with the other force fields. However, unlike OPLS and ILDN, these contacts only weakly follow an antiparallel pattern, with a stronger bias for the CHC to interact with more C-terminal residues (L34–V36). Next, we observe that CHARMM SB is characterized with stronger contacts between the N-terminal and C-terminal regions than CHARMM BB (possibly indicating a crowding effect due to the smaller box). Nonetheless, either CHARMM simulation is less prone to form intrapeptide contacts overall relative to those

simulated with other force fields; in terms of probability of intrapeptide contacts, the trend is OPLS > ILDN > CHARMM.

These characterizations are in line with the observation that either CHARMM simulation samples a more extended ensemble, as seen when we compare histograms for radii of gyration over the ensemble for  $\text{A}\beta 40$  across force fields (Figure S5). Note that there is a small population with  $R_g$  1.65–1.90 nm for OPLS  $\text{A}\beta 40$  that are known to be biased by periodic boundary contacts (PBC), as described previously by Rosenman et al.<sup>48</sup> The average values for  $R_g$  over each of the  $\text{A}\beta 40$  ensembles are  $1.07 \pm 0.16$  nm for OPLS ( $1.05 \pm 0.10$  nm ignoring the PBC-biased frames described above),  $1.12 \pm 0.11$  nm for ILDN,  $1.17 \pm 0.18$  nm for CHARMM BB, and  $1.16 \pm 0.18$  nm for CHARMM SB; though we caution that these differences are not significant, the CHARMM > ILDN > OPLS trend in  $\langle R_g \rangle$  does inversely complement the ranking for intrapeptide contact probabilities above. The histograms of  $R_g$  further reveal that both CHARMM simulations have a positively skewed  $R_g$  distribution with a long right tail, indicating that CHARMM samples configurations with  $R_g$  between 1.35 and 1.70 nm that both OPLS and ILDN sample sparingly. This was the expected result in choosing this field, and it does suggest that use of the CHARMM force field may help avoid a potential bias toward overestimating ensemble compaction. Lastly, we do observe that box size did have a subtle effect on  $R_g$ , with CHARMM SB sampling a slightly broader distribution, particularly at low  $R_g$ .

Lastly, Table 3 reveals that, like ILDN, CHARMM  $\text{A}\beta 40$  weakly populates the “crown motif” interactions involving the

**Table 3. Central Region Electrostatic Interactions of  $\text{A}\beta$  Involving the Side Chains of E22 and D23 for the CHARMM Simulation Set<sup>a</sup>**

| donor–acceptor atoms                 | % in CHARMM $\text{A}\beta 40$ BB ensemble | % in CHARMM $\text{A}\beta 40$ SB ensemble |
|--------------------------------------|--|--|
| Hydrogen Bonds                       |  |  |
| V24NH–D23O <sub>δ</sub>              | 8  | 6  |
| G25NH–D23O <sub>δ</sub>              | 7  | 5  |
| S26NH–D23O <sub>δ</sub>              | 5  | 2  |
| S26H <sub>γ</sub> –D23O <sub>δ</sub> | 1  | 1  |
| N27NH–D23O <sub>δ</sub>              | 5  | 1  |
| K28NH–D23O <sub>δ</sub>              | 3  | 0  |
| Salt Bridges                         |  |  |
| K28N <sub>ε</sub> –E22O <sub>ε</sub> | 6  | 7  |
| K28N <sub>ε</sub> –D23O <sub>δ</sub> | 7  | 5  |

<sup>a</sup>All other details are the same as OPLS simulations in Table 1.

side chains of E22 and D23 relative to OPLS. Similar to ILDN, only local interactions between D23’s side chain and V24 and G25’s amide (5–8%) and salt bridges between E22 or D23 and K28 (5–7%) are sampled substantially. CHARMM forms a similar G25–S26 turn in the central region to ILDN and shares a strong propensity to sample the S26/D23 backbone hydrogen bond (36%, and 37% in the BB and SB simulations, respectively). Recall that, in contrast, this interaction is poorly sampled by OPLS, very likely because S26’s amide is locked in interactions with the D23 side chain in that simulation.

#### 4. DISCUSSION

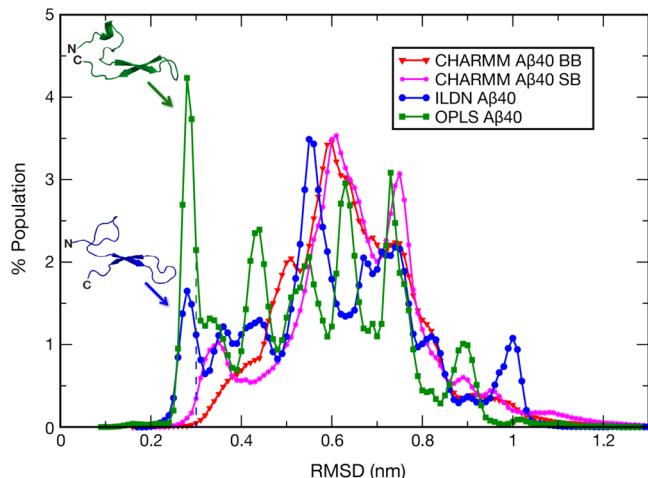
Here, we have provided a description of  $\text{A}\beta$  peptides through REMD simulations with three different all-atom force field and water model combinations (OPLS, ILDN, and CHARMM, as

defined in the Methods section), while keeping all other parameters constant. Nasica-Labouze et al.<sup>74</sup> have recently reviewed a number of computational simulations of monomeric  $\text{A}\beta$ , concluding that, while the most recent and exhaustive all-atom simulations have bared some similarity in secondary and tertiary structural biases, there are still many inconsistencies in the observations made through these studies. We suspect the choice of force field may be one of the primary sources of divergence between these investigations. While simulation with state of the art force fields has shown significant reproducibility in determining the folded state of ordered proteins, large divergence has been observed in simulations of intrinsically disordered proteins and even the unfolded states of globular proteins.<sup>40,41</sup> These cases lack unambiguous energy minima, so even subtle changes in the energy function can have profound effects on the final thermodynamic distribution of states. With this in mind, we suggest that the corroboration of results through cross-force field investigations be a possible consideration before making conclusions about intrinsically disordered systems through *in silico* simulation.

#### 4.1. Convergence Among Simulations of the Three Force Fields.

Interestingly, the simulations conducted with the OPLS, ILDN, and CHARMM parameters are characterized by very similar  $\text{C}\alpha$  contact maps. Moreover, both OPLS and ILDN simulations feature  $\beta$  hairpin topologies sampled between the CHC and the A30–G33 regions that are prevalent in the ensemble, while the A21–G29 region is characterized by an intervening loop structure that rarely contacts any of the residues in the protein except those at the extreme N-terminus. Further, both OPLS and ILDN suggest  $\text{A}\beta 42$ ’s extra two residues induce the formation of a second hairpin involving V39–I41, with a loop at G37–G38, and characterized by a prevalent L34–I41 contact. These commonalities suggest a force-field-independent characterization of the  $\text{A}\beta$  ensemble.

This shared description of the  $\text{A}\beta$  ensemble bears some interesting similarities to experimental data. For example, the secondary structure trends shared by OPLS and ILDN are also consistent with NMR data on the  $\text{A}\beta$  monomer: i.e.,  $\beta$ -rich regions match regions of positive chemical shift indices,<sup>75</sup> high  $^3\text{J}_{\text{HNHA}}$  couplings,<sup>28,31,48,76</sup> and slow solvent exchange rates<sup>28,77,78</sup> relative to the rest of the peptide. Further, the common topologies as suggested by our simulation contact maps are reminiscent of the intramolecular models of higher order aggregates including fibrils,<sup>79–81</sup> and in the case of the  $\beta$  meander suggested for in  $\text{A}\beta 42$ , a toxic, pentameric oligomer believed to be a paranucleus specific to  $\text{A}\beta 42$  aggregation.<sup>73</sup> Next, as illustrated in Figure 6, both OPLS and ILDN  $\text{A}\beta 40$  simulations sample a substantial population of configurations with  $\text{C}\alpha$  RMSD  $\leq 0.3$  nm to an Affibody-trapped  $\text{A}\beta 40$  monomer structure solved by solution NMR:<sup>82</sup> 13.5% of the ensemble for OPLS and 6.6% for ILDN (but <0.1% for both CHARMM simulations). This NMR structure is an antiparallel  $\beta$  hairpin from residues L17–V36, between the central hydrophobic cluster (CHC) and C-terminus. This aligns with the central topology consistent with the tertiary structure biases seen in the  $\text{A}\beta$  monomer simulation. The Hoyer et al. hairpin features a bonding register with (in donor amide/acceptor carbonyl nomenclature) V18/M35 and F20/G33 backbone hydrogen bonds. These bonds are also well-populated in the OPLS and ILDN  $\text{A}\beta 40$  simulations (19% and 22% in OPLS, 12% and 12% in ILDN, but only 2% and 1% in CHARMM BB). Due to the similarity of this structure to intrapeptide models of higher order aggregates, the fact that these



**Figure 6.** Histogram of  $C\alpha$  RMSDs of  $\text{A}\beta40$  REMD simulation frames to chain C, model 1 from an NMR structure of  $\text{A}\beta40$  monomer trapped by the  $Z_{\text{A}\beta3}$  Affibody protein<sup>82</sup> (PDB: 2OTK), using 0.01-nm-wide bins. Only residues 16–35 for both the reference structure and the simulation were considered for this analysis. The top centroid structure after single linkage clustering (0.2 nm cutoff) of all structures within a cutoff of  $\text{RMSD} \leq 0.3$  nm for both OPLS and ILDN  $\text{A}\beta40$  is illustrated. The dashed lines illustrate the location of this 0.3 nm cutoff on the histograms.

configurations can expose hydrophobic side chains to solvent, and the observation of an antiparallel  $\beta$  signature for early  $\text{A}\beta$  oligomers but not for mature fibrils,<sup>83,84</sup> a “Venetian blinds” mechanism of aggregation involving a rotation of stacked intramolecular antiparallel  $\beta$  strands into an intermolecular parallel  $\beta$  sheets has been proposed.<sup>82,85</sup> Our simulations provide additional evidence that CHC to C-terminal antiparallel  $\beta$  topologies like this could represent transiently occupied structures that unbound  $\text{A}\beta$  is prone to sample, and that can interact with other peptides and, thus, offer support for this hypothesis at the monomer level. Similar  $\beta$  topologies have also been observed through some of the more recent, extensive, and experimentally refined or validated simulations of  $\text{A}\beta40$  and  $\text{A}\beta42$  produced by other groups.<sup>29,31,37–39</sup>

**4.2. Differences between Simulations of the Three Force Fields.** This said, it is important to discuss the key differences between the three simulation sets. OPLS simulations significantly favor more collapsed states, and intramolecular contacts are populated with higher probability. Moreover, the local interactions that prevail between each set are considerably different. First, while both OPLS and ILDN form loop interactions in key regions (D7–G9, A21–G29, and G37–G38 in  $\text{A}\beta42$ ), OPLS is prone to form very a rigid bend structure at these positions, stabilized by strongly populated charge–charge interactions. Other types of loop structures, like turn and coil, are rare in wild type  $\text{A}\beta$ . Meanwhile, ILDN and CHARMM simulations demonstrate much more flexibility in these regions, with a greater tendency to sample a mixture of bend and turn secondary structures, the latter of which is characterized by medium range, helical backbone hydrogen bonds. The most prominent example of this disparity is the central region of  $\text{A}\beta$ . As described previously, for OPLS, this region is dominated by structures that established what we called the “crown motif”: a bulge of outwardly pointing carbonyl groups stabilized by hydrogen bonds formed by D23’s side chain with the inwardly drawn backbone amides of the downstream 4 residues, and a salt bridge between E22 and K28.

In contrast, these interactions are not strongly populated in ILDN or CHARMM simulations; instead, E22 and D23’s side chains interact with solvent much more frequently. Consequentially, the backbone amides of the central region are not locked into interactions with the D23 side chain carboxyl, while the backbone carbonyls are not forced to point outward to accommodate the crown structure. Thus, these central region amides and carbonyls are freer to form helical hydrogen bonds, which are especially well-populated between D23/S26 in both ILDN and CHARMM’s  $\text{A}\beta40$ . Experimentally, it has been previously found that there is an  $i$  to  $(i + 3)$   $d\alpha\text{N}$  nuclear Overhauser effect (NOE) constraint between D23 and S26 in  $\text{A}\beta40$ , the longest range unambiguous backbone NOE observed in monomeric  $\text{A}\beta$  to our knowledge.<sup>75</sup> Additionally, the D7–Y10 and F20–A30 regions are also the only regions observed to have  $i$  to  $(i + 2)$  backbone NOEs.<sup>75</sup> The presence of helical hydrogen bonds characterizing these regions in ILDN and CHARMM, rather than the strictly bend-characterized regions of OPLS, may be more consistent with these findings.

Another major difference between simulations with the different force fields is the role the N-terminus plays in  $\text{A}\beta$  structure. The N-terminus is conspicuously absent or poorly resolved in both the  $\text{A}\beta40$  trapped monomer solution NMR structure and solid state NMR structures of  $\text{A}\beta$  fibrils,<sup>26,79–82,86</sup> suggesting that it remains disordered even when the rest of the peptide is stabilized in  $\text{A}\beta$  aggregates or in a complex with other molecules. Further, the N-terminal sequence, like the central region but unlike the rest of the peptide, contains many charges, and is thus most IDP-like in sequence. It is not surprising that the three force field combinations considered here, which emphasize electrostatics in fundamentally different ways (see **Introduction**), would also disagree on the structure of this region. OPLS characterizes the N-terminus with a mixture of structures, but primarily coil and bend character, in contrast with the highly structured C-terminus. The CHARMM N-terminus is also similarly characterized primarily with turn, bend, and coil. However, for ILDN, while coil, bend, and turn are prevalent,  $\beta$  sheet character is also highly populated. This occurs particularly at E3–R5 and Y10–H13, and particularly in all ensembles except wild type  $\text{A}\beta40$ . Very often, these regions directly form local  $\beta$  sheets, but more global  $\beta$  pairing does occur with other regions of the protein. Finally, the extreme N-terminus is prone to contacting the rest of the protein, particularly OPLS simulations, with interactions occurring even with the central region.

**4.3. Effects of FAD Mutations and A2T.** The ensembles of FAD mutants more explicitly describe the atomistic effects of these disease-causing substitutions on the monomer form. A common theme among the OPLS simulations is that each of the FAD mutants alters the central region’s charge–charge interactions and structural biases, which in turn permit different antiparallel contact patterns seen in the contact maps in **Figure 2**. The way each mutant does this is quite different among the different species. In general, mutants of the E22 reduce but do not remove the prevalence of the D23 side chain to backbone interactions (i.e., populations of key interactions are still 50% or greater in both mutants), but trivially nullify the E22–K28 salt bridge, with no anion in the central region replacing this interaction. For OPLS, the primary effect of the E22 $\Delta$  modification is a broader distribution of bend character in the central region, with reduced bend character in apex residues V24 and G25 but increased bend character in C-terminal boundary residues G29 and A30. In E22K, ionic and  $\beta$  sheet

interactions before the bend with the N-terminus are encouraged, while the bend character is generally reduced.  $\text{A}\beta 40\text{-D23N}$ , on the other hand, has the most profound effect on central region electrostatics observed among the OPLS set. Substitution of the less charged side chain greatly reduces both E22 and D23 characteristic interactions, and disrupts sampling of bulge “crown” features in favor of narrower bends. The distribution of bend secondary structure is both reduced and narrowed, while residues flanking the bend (F20–D/N23 and K28–I31) are greatly increased in  $\beta$  character. Consequentially, all OPLS FAD mutants broaden the distribution of C-terminal partners in the CHC to I31–V36 main antidiagonal, but the two E22 mutants are particularly bias bonding toward more C-terminal registers, while D23N prefers registers more N-terminal to that of wild type. As discussed in more depth in the **Introduction**, certain FAD mutants such as E22 $\Delta$  and D23N are affiliated with novel aggregation morphologies.<sup>22–26</sup> If transient sampling of the central hairpin in the monomeric form truly acts as a seed for  $\text{A}\beta$  aggregation, as we have suggested above, then we can speculate that the increased polymorphism in the types of cross-bend interaction patterns seen through OPLS REMD simulations may permit the adoption of unique aggregates and new, faster aggregation pathways.

However, while OPLS featured strongly populated electrostatic roles for E22 and D23 that stabilize the “crown motif” structure, the physicochemical influences of E22 and D23 are not as clear in the ILDN and CHARMM sets. These residues do form some of the same salt bridges and hydrogen bonds, but the populations of these interactions are greatly reduced in population compared to OPLS, in exchange for more frequent interactions with solvent. Consequentially, the FAD mutants of  $\text{A}\beta$  that we studied, which target these two residues, have considerably different effects on the  $\text{A}\beta$  ensemble for each force field. OPLS saw well-defined changes in the central region bend structure and shifts in antiparallel interaction patterns associated with changes in the population of electrostatic interactions, but the effects of these substitutions are not so prevalent among the ILDN set. It is still true that FAD mutations simulated with ILDN still change the frequencies of interactions in the central region and the structure of the central region; the distribution of the turn region is narrowed while residues 22 and 23 in these systems are more prone to adopting the  $\beta$  structure. However, these changes are relatively subtle compared to the OPLS systems. Ultimately, the depopulation of the E22 and D23 interactions in both ILDN and CHARMM compared to OPLS may suggest that the “crown motif” may be an artifact specific to the OPLS force field.

More obvious among the differences common to the ILDN simulated point mutants of  $\text{A}\beta$  versus wild type is that the N-terminus is more active in forming both local and global  $\beta$  hairpins in the former, as opposed to more localized turn and bend structure in the latter. This behavior was expected for the A2T variant as threonine is a residue far more inclined to adopt  $\beta$  structure than alanine,<sup>87,88</sup> but the mechanism behind this change for FAD mutants of the central region is not as clear. Consider, however, that both the N-terminal region and the central region in question are net negatively charged. These FAD mutations share in common that they reduce the amount of net negative charges in the central region (or in certain mutants like H6R or D7N, the N-terminus itself), so one could speculate that, by mutating these residues, we are removing

repulsive interactions from the system that would otherwise force the N-terminus to avoid interactions with these regions. This causal hypothesis is challenging to test, however. Interestingly,  $\text{A}\beta 40\text{-E22}\Delta$  was also the only variant to greatly reduce the  $\beta$  content at the CHC and the C-terminus. This was seen in OPLS simulations as well, although there, the presence of more C-terminal character (M35 and V36) made up for the loss of  $\beta$  character in residues like I31 and I32, which is not true for the ILDN variant.  $\text{A}\beta 40\text{-E22G}$  also showed an interesting change by shifting the distributions of  $\beta$  character for the CHC and the C-terminus apart in sequence, and expanding the breadth of the intervening central region. These changes aside, however, FAD mutants were observed to be far less prone to changing the distribution of  $\beta$  sheets for ILDN compared to OPLS, whereas heightened N-terminal  $\beta$  character was a common trend seen among simulations with this force field.

The increased interactivity of the N-terminus in  $\beta$  pairing seen in ILDN simulations of these mutants leads to the possibility of more diverse  $\beta$  topologies. One could further speculate that the presence of these newly allowed  $\beta$  combinations in the monomer ensemble may allow for the seeding of different aggregates that could either account for the novel higher order forms and/or faster aggregation pathways that define the phenotypes of these mutants. This hypothesis is a different alternative to the one we can propose for OPLS FAD mutants, which suggest the phenotype of FAD mutants may arise from the broader range of registers adopted by the central hairpin seen in OPLS simulations. Since neither of these hypotheses are based on observations that are common to both force field combinations, we cannot endorse one over the other at this time.

Further, since  $\text{A}\beta 40\text{-A2T}$  contributes a similar effect as the FAD mutants on the involvement of the N-terminus in the ILDN set, it is also unclear how (or even if) this mutant may confer changes in monomeric structure that lead to its neuroprotective phenotype (though the similarity in behavior would be in line with the finding that this mutant actually enhances  $\text{A}\beta 40$  aggregation<sup>89</sup>). This said, recently, REMD simulations up to 175 ns/replica with the OPLS-AA/TIP3P force field comparing this A2T variant with wild type as well as the A2 V FAD mutant in the context of  $\text{A}\beta 42$  were published<sup>90</sup> that may clarify the involvement of the N-terminus in  $\text{A}\beta$  aggregation. Like the findings described here for our  $\text{A}\beta 40\text{-A2T}$  simulation, it was observed that this point mutation encourages the N-terminus to engage distant regions of the peptide. However, in the context of  $\text{A}\beta 42$  in this Das et al.<sup>90</sup> study, it was observed that this behavior by A2T discouraged the formation of double  $\beta$  hairpin conformations involving the CHC and C-terminus (matching those prone to be adopted by wild type  $\text{A}\beta 42$  in the OPLS and ILDN simulations described here) in favor of single hairpins at the C-terminus, while the A2V mutant encouraged sampling of these meander structures. As described above, these types of double hairpins are unique to  $\text{A}\beta 42$  in our simulations with OPLS and ILDN and bear strong resemblance to a solid state NMR model of neurotoxic  $\text{A}\beta 42$  oligomers.<sup>73</sup>

**4.4. Ensemble Compaction.** Besides being yet another force field to test the generalizability of our results, the CHARMM simulations of  $\text{A}\beta 40$  were also conducted to investigate a recently observed trend that *de novo* simulations of other IDPs using commonly used force fields, like the ones run here, have been observed to underestimate the radii of gyration as determined by small-angle X-ray scattering (SAXS).<sup>41,42,61</sup>

Of the force fields investigated previously, CHARMM22\* was the most accurate in reproducing SAXS values of these other IDPs.<sup>41,42</sup> We observe that simulations of A $\beta$ 40 monomers with this force field preserve the tertiary structural biases seen in the contact maps of both OPLS and ILDN simulations of this peptide (though the secondary structure is substantially different). Further, CHARMM promotes a more extended ensemble than either OPLS or ILDN, with a bias toward populating frames with higher radii of gyration. If the force fields used for OPLS and ILDN are indeed biased to collapsed states, CHARMM may be a better conformational representation of the A $\beta$  ensemble. However, it is difficult to make a conclusion because experimental shape-dependent measurements for monomeric A $\beta$  under physiological conditions are unavailable to our knowledge (presumably because of the aggregation-prone nature of the peptide). We propose to test our methodology by applying it to other peptides that have available size measurements. If these simulations should fail to adequately reproduce experimental data, other approaches to correct for this bias include the following: strengthening short-range protein–water pair interactions,<sup>43,91</sup> modifying the weight of other nonbonded interactions in pre-existing force fields to de-emphasize the nonspecific protein–protein interactions that promote collapsed states, or employing experimental-based constraints on either the simulation or the analyzed ensemble.

## 5. CONCLUSIONS

The similarities between the three simulations make it very challenging to find direct evidence to discount one model over the other on the basis of the available experimental data. Hence, we value those properties that are consistent among these different investigations. Fortunately, the descriptions that fit this criterion are sufficient to develop a generalized depiction of wild type A $\beta$  monomers. The agreement between these different techniques is an important and nontrivial finding: we have previously observed that REMD simulations of a small peptide from  $\alpha$ -synuclein using the very same OPLS and ILDN force fields employed in our study of A $\beta$  here nevertheless produce very different ensembles for that system.<sup>28</sup> We describe A $\beta$  peptides here, in summary, as disordered proteins that are nonetheless biased to forming  $\beta$  topologies that bear resemblance to the intrapeptide models of higher order aggregates. As such, we suspect that the shared conformations presented here for A $\beta$  may act as the seeds for the formation of these aggregates and, thus, potential targets for drug discovery in Alzheimer's disease.

## AUTHOR INFORMATION

### Corresponding Author

\*Phone: (505) 667-3883. E-mail: [agarcia@lanl.gov](mailto:agarcia@lanl.gov).

### Notes

The authors declare no competing financial interest.

## Biographies



David J. Rosenman earned his B.S. with Honors in Biology from the California Institute of Technology in 2009, and he graduated with a Ph.D. in Biology at the Rensselaer Polytechnic Institute in 2015. His graduate research was performed under the mentorship of Prof. Angel E. García and Prof. Chunyu Wang and focused on the combined use of molecular dynamics and nuclear magnetic resonance techniques to characterize the conformational ensembles of the A $\beta$  peptides implicated in Alzheimer's disease. He is now a postdoctoral researcher at the University of Delaware under the supervision of Prof. Abraham M. Lenhoff; his current project investigates the self-interaction properties of monoclonal antibodies.



Chunyu Wang is Associate Professor of Biology at Rensselaer Polytechnic Institute. He obtained his M.D. from Peking Union Medical College in 1996 and his Ph.D. in biochemistry from Cornell University in 2000. He mainly uses solution NMR techniques to study the structure, dynamics, and interactions of amyloid  $\beta$  peptide, in combination with molecular dynamics simulations. His other interests include intein-mediated protein splicing and related phenomena, such as hedgehog autoprocessing.



Angel E. García is the Director of the Center for Non Linear Studies (CNLS) at Los Alamos National Laboratory and Professor of Physics at Rensselaer Polytechnic Institute. He obtained a Ph.D. in theoretical physics from Cornell University. His main research interest is the understanding of the physicochemical principles that determine biomolecular folding, stability, and dynamics and the application of these principles to build computer models that describe these processes. A main effort in his laboratory has been the characterization of the ensemble of conformations adopted by the  $\text{A}\beta$  peptides in the monomeric state. Other interests include the study of the effect of hydrostatic pressure on protein and RNA stability.

## ACKNOWLEDGMENTS

Portions of this publication were previously submitted in a thesis in partial fulfillment for the requirements of the Ph.D. for D.J.R. (Rensselaer Polytechnic Institute, August 2015), cited as ref 28 in this publication. More information on the simulations discussed here, including analyses of ensemble convergence, representative cluster centroid structures, radii of gyration, solvent accessible surface area, and quantitative comparisons to experimental NMR observables (chemical shifts, *J*-couplings, etc.) are described in detail in the aforementioned thesis document. This work was supported by the National Institutes of Health—National Institute of General Medical Sciences Biomolecular Science and Engineering Training Grant (5 T32 GM067545-08), the Weissman Family Fellowship, and the National Science Foundation (MCB-1050966). All simulations were conducted using resources provided by the Extreme Science and Engineering Discovery Environment (XSEDE) Grant TG-MCB130178.

## ABBREVIATIONS

$\text{A}\beta$ , amyloid  $\beta$ ; AD, Alzheimer's disease; FAD, familial Alzheimer's disease; REMD, replica exchange molecular dynamics; CHC, central hydrophobic cluster

## REFERENCES

- (1) Uversky, V. N.; Gillespie, J. R.; Fink, A. L. Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins: Struct., Funct., Genet.* **2000**, *41*, 415–427.
- (2) Hård, T.; Lendel, C. Inhibition of amyloid formation. *J. Mol. Biol.* **2012**, *421*, 441–465.
- (3) Knowles, T. P. J.; Vendruscolo, M.; Dobson, C. M. The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* **2014**, *15*, 384–396.
- (4) Hardy, J.; Selkoe, D. J. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* **2002**, *297*, 353–356.
- (5) Batsch, N.; Mittelman, M. World Alzheimer Report 2012: Overcoming the Stigma of Dementia. *London Alzheimers Dis. Int.* **2012**, *75*.
- (6) Bitan, G.; Vollmers, S. S.; Teplow, D. B. Elucidation of primary structure elements controlling early amyloid  $\beta$ -protein oligomerization. *J. Biol. Chem.* **2003**, *278*, 34882–34889.
- (7) Bitan, G.; Kirkpatrick, M. D.; Lomakin, A.; Vollmers, S. S.; Benedek, G. B.; Teplow, D. B. Amyloid  $\beta$ -protein ( $\text{A}\beta$ ) assembly:  $\text{A}\beta40$  and  $\text{A}\beta42$  oligomerize through distinct pathways. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 330–335.
- (8) Selkoe, D. J.; Podlisny, M. B. Deciphering the genetic basis of Alzheimer's disease. *Annu. Rev. Genomics Hum. Genet.* **2002**, *3*, 67–99.
- (9) Takami, M.; Nagashima, Y.; Sano, Y.; Ishihara, S.; Morishima-Kawashima, M.; Funamoto, S.; Ihara, Y.  $\gamma$ -Secretase: successive tripeptide and tetrapeptide release from the transmembrane domain of  $\beta$ -carboxyl terminal fragment. *J. Neurosci.* **2009**, *29*, 13042–13052.
- (10) Wolfe, M. S.; Citron, M.; Diehl, T. S.; Xia, W.; Donkor, I. O.; Selkoe, D. J. A substrate-based difluoro ketone selectively inhibits Alzheimer's  $\gamma$ -secretase activity. *J. Med. Chem.* **1998**, *41*, 6–9.
- (11) Haass, C.; Selkoe, D. J. Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid  $\beta$ -peptide. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 101–112.
- (12) Findeis, M. A. The role of amyloid  $\beta$  peptide 42 in Alzheimer's disease. *Pharmacol. Ther.* **2007**, *116*, 266–286.
- (13) Thal, D. R.; Griffin, W. S.; de Vos, R. A.; Ghebremedhin, E. Cerebral amyloid angiopathy and its relationship to Alzheimer's disease. *Acta Neuropathol.* **2008**, *115*, 599–609.
- (14) Tomidokoro, Y.; Rostagno, A.; Neubert, T. A.; Lu, Y.; Rebeck, G. W.; Frangione, B.; Greenberg, S. M.; Ghiso, J. Iowa variant of familial Alzheimer's disease: accumulation of posttranslationally modified  $\text{A}\beta\text{D23N}$  in parenchymal and cerebrovascular amyloid deposits. *Am. J. Pathol.* **2010**, *176*, 1841–1854.
- (15) Alzheimer's Association.. 2014 Alzheimer's disease facts and figures. *Alzheimer's Dementia* **2014**, *10*, 47–92.
- (16) Van Broeckhoven, C.; Haan, J.; Bakker, E.; Hardy, J. A.; Van Hul, W.; Wehnert, A.; Vegter-Van der Vlis, M.; Roos, R. A. Amyloid  $\beta$  protein precursor gene and hereditary cerebral hemorrhage with amyloidosis (Dutch). *Science* **1990**, *248*, 1120–1122.
- (17) Levy, E.; Carman, M. D.; Fernandez-Madrid, I. J.; Power, M. D.; Lieberburg, I.; van Duinen, S. G.; Bots, G. T.; Luyendijk, W.; Frangione, B. Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemorrhage, Dutch type. *Science* **1990**, *248*, 1124–1126.
- (18) Miravalle, L.; Tokuda, T.; Chiarle, R.; Giaccone, G.; Bugiani, O.; Tagliavini, F.; Frangione, B.; Ghiso, J. Substitutions at codon 22 of Alzheimer's  $\text{A}\beta$  peptide induce diverse conformational changes and apoptotic effects in human cerebral endothelial cells. *J. Biol. Chem.* **2000**, *275*, 27110–27116.
- (19) Grabowski, T. J.; Cho, H. S.; Vonsattel, J. P.; Rebeck, G. W.; Greenberg, S. M. Novel amyloid precursor protein mutation in an Iowa family with dementia and severe cerebral amyloid angiopathy. *Ann. Neurol.* **2001**, *49*, 697–705.
- (20) Ozawa, K.; Tomiyama, T.; Maat-Schieman, M. L.; Roos, R. A.; Mori, H. Enhanced  $\text{A}\beta40$  deposition was associated with increased  $\text{A}\beta42-43$  in cerebral vasculature with Dutch-type hereditary cerebral hemorrhage with amyloidosis (HCHWA-D). *Ann. N. Y. Acad. Sci.* **2002**, *977*, 149–154.
- (21) Grant, M. A.; Lazo, N. D.; Lomakin, A.; Condron, M. M.; Arai, H.; Yamin, G.; Rigby, A. C.; Teplow, D. B. Familial Alzheimer's disease mutations alter the stability of the amyloid  $\beta$ -protein monomer folding nucleus. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 16522–16527.

- (22) Tycko, R.; Sciarretta, K. L.; Orgel, J. P. R. O.; Meredith, S. C. Evidence for novel  $\beta$ -sheet structures in Iowa mutant  $\beta$ -amyloid fibrils. *Biochemistry* **2009**, *48*, 6072–6084.
- (23) Cloe, A. L.; Orgel, J. P.; Sachleben, J. R.; Tycko, R.; Meredith, S. C. The Japanese mutant  $A\beta$  ( $\Delta E22-A\beta(1-39)$ ) forms fibrils instantaneously, with low-thioflavin T fluorescence: seeding of wild-type  $A\beta(1-40)$  into atypical fibrils by  $\Delta E22-A\beta(1-39)$ . *Biochemistry* **2011**, *50*, 2026–2039.
- (24) Ovchinnikova, O. Y.; Finder, V. H.; Vodopivec, I.; Nitsch, R. M.; Glockshuber, R. The Osaka FAD mutation E22 $\Delta$  leads to the formation of a previously unknown type of amyloid  $\beta$  fibrils and modulates  $A\beta$  neurotoxicity. *J. Mol. Biol.* **2011**, *408*, 780–791.
- (25) Inayathullah, M.; Teplow, D. B. Structural dynamics of the  $\Delta E22$  (Osaka) familial Alzheimer's disease-linked amyloid  $\beta$ -protein. *Amyloid* **2011**, *18*, 98–107.
- (26) Qiang, W.; Yau, W. M.; Luo, Y.; Mattson, M. P.; Tycko, R. Antiparallel  $\beta$ -sheet architecture in Iowa-mutant  $\beta$ -amyloid fibrils. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 4443–4448.
- (27) Jonsson, T.; Atwal, J. K.; Steinberg, S.; Snaedal, J.; Jonsson, P. V.; Björnsson, S.; Stefansson, H.; Sulem, P.; Gudbjartsson, D.; Maloney, J.; et al. A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. *Nature* **2012**, *488*, 96–99.
- (28) Rosenman, D. J. *Biophysical characterization of intrinsically disordered peptides through molecular dynamics simulations and solution nuclear magnetic resonance spectroscopy*. Ph.D. Thesis, Rensselaer Polytechnic Institute, Troy, NY, 2015.
- (29) Sgourakis, N. G.; Yan, Y.; McCallum, S. A.; Wang, C.; Garcia, A. E. The Alzheimer's peptides  $A\beta40$  and 42 adopt distinct conformations in water: a combined MD/NMR study. *J. Mol. Biol.* **2007**, *368*, 1448–1457.
- (30) Mitternacht, S.; Staneva, I.; Härd, T.; Irbäck, A. Comparing the folding free-energy landscapes of  $A\beta42$  variants with different aggregation properties. *Proteins: Struct., Funct., Genet.* **2010**, *78*, 2600–2608.
- (31) Sgourakis, N. G.; Merced-Serrano, M.; Boutsidis, C.; Drineas, P.; Du, Z.; Wang, C.; Garcia, A. E. Atomic-level characterization of the ensemble of the  $A\beta(1-42)$  monomer in water using unbiased molecular dynamics simulations and spectral algorithms. *J. Mol. Biol.* **2011**, *403*, 570–583.
- (32) Côté, S.; Derreumaux, P.; Mousseau, N. Distinct morphologies for amyloid  $\beta$  protein monomer:  $A\beta_{1-40}$ ,  $A\beta_{1-42}$ , and  $A\beta_{1-40}(D23N)$ . *J. Chem. Theory Comput.* **2011**, *7*, 2584–2592.
- (33) Ball, K. A.; Phillips, A. H.; Nerenberg, P. S.; Fawzi, N. L.; Wemmer, D. E.; Head-Gordon, T. Homogeneous and heterogeneous tertiary structure ensembles of amyloid- $\beta$  peptides. *Biochemistry* **2011**, *50*, 7612–7628.
- (34) Lin, Y. S.; Bowman, G. R.; Beauchamp, K. A.; Pande, V. S. Investigating how peptide length and a pathogenic mutation modify the structural ensemble of amyloid  $\beta$  monomer. *Biophys. J.* **2012**, *102*, 315–324.
- (35) Lin, Y. S.; Pande, V. S. Effects of familial mutations on the monomer structure of  $A\beta42$ . *Biophys. J.* **2012**, *103*, L47–49.
- (36) Granata, D. *Characterizing structure and free energy landscape of proteins by NMR-guided metadynamics*. Ph.D. Thesis, SISSA, Trieste, Italy, 2013.
- (37) Ball, K. A.; Phillips, A. H.; Wemmer, D. E.; Head-Gordon, T. Differences in  $\beta$ -strand populations of monomeric  $A\beta40$  and  $A\beta42$ . *Biophys. J.* **2013**, *104*, 2714–2724.
- (38) Zhu, M.; De Simone, A.; Schenk, D.; Toth, G.; Dobson, C. M.; Vendruscolo, M. Identification of small-molecule binding pockets in the soluble monomeric form of the  $A\beta42$  peptide. *J. Chem. Phys.* **2013**, *139*, 035101.
- (39) Song, W.; Wang, Y.; Colletier, J.-P.; Yang, H.; Xu, Y. Varied probability of staying collapsed/extended at the conformational equilibrium of monomeric  $A\beta40$  and  $A\beta42$ . *Sci. Rep.* **2015**, *5*, 11024.
- (40) Piana, S.; Lindorff-Larsen, K.; Shaw, D. E. How robust are protein folding simulations with respect to force field parameterization? *Biophys. J.* **2011**, *100*, L47–L49.
- (41) Rauscher, S.; Gapsys, V.; Volkhardt, A.; Blau, C.; de Groot, B. L.; Grubmüller, H. Structural ensembles of intrinsically disordered proteins depend strongly on force field. *Biophys. J.* **2014**, *106*, 271a.
- (42) Rauscher, S.; Gapsys, V.; de Groot, B.; Grubmüller, H. Structural ensembles of intrinsically disordered proteins using molecular dynamics simulation. *Biophys. J.* **2015**, *108*, 14a.
- (43) Henriques, J.; Cagnell, C.; Skepö, M. Molecular dynamics simulations of intrinsically disordered proteins: force field evaluation and comparison with experiment. *J. Chem. Theory Comput.* **2015**, *11*, 3420–3431.
- (44) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B* **2001**, *105*, 6474–6487.
- (45) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–926.
- (46) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Struct., Funct., Genet.* **2010**, *78*, 1950–1958.
- (47) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120*, 9665–9678.
- (48) Rosenman, D. J.; Connors, C. R.; Chen, W.; Wang, C.; Garcia, A. E.  $A\beta$  monomers transiently sample oligomer and fibril-like configurations: ensemble characterization using a combined MD/NMR approach. *J. Mol. Biol.* **2013**, *425*, 3338–3359.
- (49) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Genet.* **2006**, *65*, 712–725.
- (50) Best, R. B.; Hummer, G. Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. *J. Phys. Chem. B* **2009**, *113*, 9004–9015.
- (51) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; et al. Atomic-level characterization of the structural dynamics of proteins. *Science* **2010**, *330*, 341–346.
- (52) Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Shaw, D. E. How fast-folding proteins fold. *Science* **2011**, *334*, 517–520.
- (53) Piana, S.; Lindorff-Larsen, K.; Shaw, D. E. Protein folding kinetics and thermodynamics from atomistic simulation. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 17845–17850.
- (54) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **1984**, *106*, 765–784.
- (55) Jorgensen, W. L.; Tirado-Rives, J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.
- (56) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
- (57) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
- (58) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; et al. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **2003**, *24*, 1999–2012.

- (59) Jorgensen, W. L.; Tirado-Rives, J. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6665–6670.
- (60) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1993**, *97*, 10269–10280.
- (61) Palazzesi, F.; Prakash, M. K.; Bonomi, M.; Barducci, A. Accuracy of current all-atom force-fields in modeling protein disordered states. *J. Chem. Theory Comput.* **2015**, *11*, 2–7.
- (62) Lindorff-Larsen, K.; Trbovic, N.; Maragakis, P.; Piana, S.; Shaw, D. E. Structure and dynamics of an unfolded protein examined by molecular dynamics simulation. *J. Am. Chem. Soc.* **2012**, *134*, 3787–3791.
- (63) Piana, S.; Klepeis, J. L.; Shaw, D. E. Assessing the accuracy of physical models used in protein-folding simulations: quantitative evidence from long molecular dynamics simulations. *Curr. Opin. Struct. Biol.* **2014**, *24*, 98–105.
- (64) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- (65) Hukushima, K.; Nemoto, K. Exchange Monte Carlo method and application to spin glass simulations. *J. Phys. Soc. Jpn.* **1996**, *65*, 1604–1608.
- (66) Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *314*, 141–151.
- (67) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Genet.* **2006**, *65*, 712–725.
- (68) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.
- (69) Kabsch, W.; Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983**, *22*, 2577–2637.
- (70) Joosten, R. P.; te Beek, T. A. H.; Krieger, E.; Hekkelman, M. L.; Hooft, R. W. W.; Schneider, R.; Sander, C.; Vriend, G. A series of PDB related databases for everyday needs. *Nucleic Acids Res.* **2011**, *39*, D411–D419.
- (71) Voelz, V. A.; Shell, M. S.; Dill, K. A. Predicting peptide structures in native proteins from physical simulations of fragments. *PLoS Comput. Biol.* **2009**, *5*, e1000281.
- (72) Yan, Y.; Wang, C.  $\text{A}\beta$ 42 is more rigid than  $\text{A}\beta$ 40 at the C terminus: implications for  $\text{A}\beta$  aggregation and toxicity. *J. Mol. Biol.* **2006**, *364*, 853–862.
- (73) Ahmed, M.; Davis, J.; Aucoin, D.; Sato, T.; Ahuja, S.; Aimoto, S.; Elliott, J. I.; Van Nostrand, W. E.; Smith, S. O. Structural conversion of neurotoxic amyloid- $\beta_{1-42}$  oligomers to fibrils. *Nat. Struct. Mol. Biol.* **2010**, *17*, 561–567.
- (74) Nasica-Labouze, J.; Nguyen, P. H.; Sterpone, F.; Berthoumieu, O.; Buchete, N.-V.; Coté, S.; De Simone, A.; Doig, A. J.; Faller, P.; Garcia, A.; et al. Amyloid  $\beta$  protein and Alzheimer's disease: when computer simulations complement experimental studies. *Chem. Rev.* **2015**, *115*, 3518–3563.
- (75) Hou, L.; Shao, H.; Zhang, Y.; Li, H.; Menon, N. K.; Neuhaus, E. B.; Brewer, J. M.; Byeon, I.-J. L.; Ray, D. G.; Vitek, M. P.; et al. Solution NMR studies of the  $\text{A}\beta(1-40)$  and  $\text{A}\beta(1-42)$  peptides establish that the Met35 oxidation state affects the mechanism of amyloid formation. *J. Am. Chem. Soc.* **2004**, *126*, 1992–2005.
- (76) Yan, Y.; McCallum, S. A.; Wang, C. M35 oxidation induces  $\text{A}\beta$ 40-like structural and dynamical changes in  $\text{A}\beta$ 42. *J. Am. Chem. Soc.* **2008**, *130*, 5394–5395.
- (77) Rezaei-Ghaleh, N.; Andreetto, E.; Yan, L.-M.; Kapurniotu, A.; Zweckstetter, M. Interaction between amyloid  $\beta$  peptide and an aggregation blocker peptide mimicking islet amyloid polypeptide. *PLoS One* **2011**, *6*, e20289.
- (78) Abelein, A.; Bolognesi, B.; Dobson, C. M.; Gräslund, A.; Lendel, C. Hydrophobicity and conformational change as mechanistic determinants for nonspecific modulators of amyloid  $\beta$  self-assembly. *Biochemistry* **2012**, *51*, 126–137.
- (79) Petkova, A. T.; Yau, W.-M.; Tycko, R. Experimental constraints on quaternary structure in Alzheimer's  $\beta$ -amyloid fibrils. *Biochemistry* **2006**, *45*, 498–512.
- (80) Paravastu, A. K.; Leapman, R. D.; Yau, W.-M.; Tycko, R. Molecular structural basis for polymorphism in Alzheimer's  $\beta$ -amyloid fibrils. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 18349–18354.
- (81) Bertini, I.; Gonnelli, L.; Luchinat, C.; Mao, J.; Nesi, A. A new structural model of  $\text{A}\beta$ 40 fibrils. *J. Am. Chem. Soc.* **2011**, *133*, 16013–16022.
- (82) Hoyer, W.; Gronwall, C.; Jonsson, A.; Stahl, S.; Hard, T. Stabilization of a  $\beta$ -hairpin in monomeric Alzheimer's amyloid- $\beta$  peptide inhibits amyloid formation. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 5099–5104.
- (83) Habicht, G.; Haupt, C.; Friedrich, R. P.; Hortschansky, P.; Sachse, C.; Meinhardt, J.; Wielgmann, K.; Gellermann, G. P.; Brodhun, M.; Gotz, J.; et al. Directed selection of a conformational antibody domain that prevents mature amyloid fibril formation by stabilizing  $\text{A}\beta$  protofibrils. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 19232–19237.
- (84) Cerf, E.; Sarroukh, R.; Tamamizu-Kato, S.; Breydo, L.; Derclaye, S.; Dufrêne, Y. F.; Narayanaswami, V.; Goormaghtigh, E.; Ruysschaert, J. M.; Raussens, V. Antiparallel  $\beta$ -sheet: a signature structure of the oligomeric amyloid  $\beta$ -peptide. *Biochem. J.* **2009**, *421*, 415–423.
- (85) Härd, T. Protein engineering to stabilize soluble amyloid  $\beta$ -protein aggregates for structural and functional studies. *FEBS J.* **2011**, *278*, 3884–3892.
- (86) Luhrs, T.; Ritter, C.; Adrian, M.; Riek-Lohr, D.; Bohrmann, B.; Dobeli, H.; Schubert, D.; Riek, R. 3D structure of Alzheimer's amyloid- $\beta(1-42)$  fibrils. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 17342–17347.
- (87) Chou, P. Y.; Fasman, G. D. Prediction of protein conformation. *Biochemistry* **1974**, *13*, 222–245.
- (88) Prevelige, P., Jr.; Fasman, G. D. Chou-Fasman prediction of the secondary structure of proteins. In *Prediction of Protein Structure and the Principles of Protein Conformation*; Fasman, G. D., Ed.; Plenum Press: New York, NY, 1989; pp 391–416.
- (89) Meinhardt, J.; Tartaglia, G. G.; Pawar, A.; Christopeit, T.; Hortschansky, P.; Schroeckh, V.; Dobson, C. M.; Vendruscolo, M.; Fandrich, M. Similarities in the thermodynamics and kinetics of aggregation of disease-related  $\text{A}\beta(1-40)$  peptides. *Protein Sci.* **2007**, *16*, 1214–1222.
- (90) Das, P.; Murray, B.; Belfort, G. Alzheimer's protective A2T mutation changes the conformational landscape of the  $\text{A}\beta$ 1–42 monomer differently than does the A2V mutation. *Biophys. J.* **2015**, *108*, 738–747.
- (91) Best, R. B.; Zheng, W.; Mittal, J. Balanced protein–water interactions improve properties of disordered proteins and non-specific protein association. *J. Chem. Theory Comput.* **2014**, *10*, 5113–5124.