Published in final edited form as:
*J Mol Biol*. 2007 May 18; 368(5): 1448–1457.

# The Alzheimer's peptides Aβ40 and 42 adopt distinct conformations in water: A combined MD / NMR study.

**Nikolaos G. Sgourakis**, **Yilin Yan**, **Scott McCallum**, **Chunyu Wang**, and **Angel E. Garcia***
*Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, NY 12180*

## Summary

The role of two peptides, Aβ40 and Aβ42 in the early pathogenesis of the Alzheimer's disease (AD) is frequently emphasized in the literature. It is known that Aβ42 is more prone to aggregation than Aβ40, even though they only differ in two (IA) amino acid residues at the C-terminal end. A direct comparison of the ensembles of conformations adopted by the monomers in solution has been limited by the inherent flexibility of the unfolded peptides. Here we characterize the conformations of Aβ40 and Aβ42 in water by using a combination of molecular dynamics (MD) and measured scalar $^{3}J_{\mathrm{HNH}\alpha}$ data from NMR experiments. We perform replica exchange MD (REMD) simulations and find that classical forcefields quantitatively reproduce the NMR data when the sampling is extended to the microseconds time scale. Using the quantitative agreement of the NMR data as a validation of the model, we proceed to compare the conformational ensembles of the Aβ40 and Aβ42 peptide monomers. Our analysis confirms the existence of structured regions within the otherwise flexible Aβ peptides. We find that the C-terminus of Aβ42 is more structured than that of Aβ40. The formation of a β-hairpin in the sequence **$^{31}$IIGL**MVG**GVVI**A involving short strands at residues 31−34 and 38−41 reduces the C-terminal flexibility of the Aβ42 peptide and may be responsible for the higher propensity of this peptide to form amyloids.

### Keywords

Alzheimer's disease; Amyloid-β peptides; conformational ensemble; replica exchange molecular dynamics; J-coupling constants

Two peptides have received tremendous interest in modern Alzheimer's disease (AD) research. Aβ40 and 42 are major products of the proteolytic cleavage of a multi-domain integral membrane type I protein, Amyloid-β Precursor Protein (APP), whose functions include cell adhesion, neuronal mobility and transcriptional regulation[1]. APP metabolism includes processing by a group of dedicated proteases, named secretases, in two known pathways to yield intracellular and extracellular fragments with a broad range of functions in synaptic transmission and neuronal plasticity[2]. During the amyloidogenic pathway, the action of β and subsequently γ secretase yields the Aβ peptides. The exact location of the transmembrane cleavage site for γ secretase results in a variability in the length of its Aβ product from 38 to 43 residues, however lengths of 40 and 42 are the dominant species. The physiological role of the Aβ peptides *in vivo* remains unclear. Aβ40 has been proposed to regulate the activity of K + channels[3] and also to modulate synaptic transmission[4], however the mechanisms upon it

*To whom correspondence should be addressed angel@rpi.edu

exerts these actions are unknown. Moreover, the Aβ peptides are notorious for their toxic effects to neurons[5] and for their profound ability to polymerize into insoluble fibrillar deposits, which are the main constituent of senile plaques observed in brains of AD patients. The extent to which amyloid formation has a causative role in AD is still controversial, however the neurotoxicity of these aggregated forms of Aβ is a fact and can be attributed to several proposed mechanisms[6].

The assembly of Aβ into fibrils occurs through a multi-step seed/nucleation process that involves key soluble oligomeric intermediates and protofibril states [7; 8]. Although the exact series of molecular events that lead to this transformation is unknown, it has been proposed that an early conformational switch towards an α-helical oligomeric structure is a prerequisite en route to fibril formation[9; 10]. The multi-pathway nature of the assembly process, as well as the obligatory α-helical intermediate, were also observed by means of molecular dynamics simulations of the oligomerization of Aβ (16−22) peptides[11].

A general interest towards elucidating the structure of Aβ fibrils has led to numerous structural models for fibrils of both peptides[12; 13; 14; 15], yet little is known about the structure of the oligomeric species and the mechanism of transitions among monomers, oligomers and fibrils. Understanding this process is essential in designing therapeutics that target amyloid formation at an early stage of the disease[16]. In this sense the conformational diversity of the monomers is of particular interest, as it is the starting point for modeling of any nucleation pathway. Different monomer conformations can provide the seed for pathologic oligomeric species as well as intermediates to fibril formation. For this purpose, several experimental studies on various lengths of the peptides have been performed, mainly in membrane mimicking environments (reviewed in[8]), where Aβ adopts a structure containing two α-helices joined by a loop. In water, which is the *in vivo* active form, it has been postulated that the Aβs are devoid of α or β structures, according to a collapsed-coil model that is based on NMR studies of an Aβ 26mer(10−35) [17]. The same study suggested that Aβ has a well-defined hydrophobic core with flexible terminal regions. However, due to inherent difficulties in sample preparation[18], limited experimental information is available on the range of conformations that are accessible to the full-length hydrated systems in atomic detail from three published studies[19; 20; 21].

Aβ42 is much more prone to aggregation and much more toxic to neurons than Aβ40[22; 23]. Apparently the two peptides form fibrils via different pathways since the dominant oligomeric species are larger for Aβ42 than Aβ40[24], although the structural basis for this phenotype is unknown. Thus, characterizing the structural differences in Aβ40 and Aβ42 monomers may provide insights into their differences in aggregation and toxicity.

In this study, we chose Molecular Dynamics simulations at the microsecond timescale as a method to explore the range of conformations that are accessible to the full length hydrated monomers of Aβ. The implementation of the Replica Exchange Molecular Dynamic (REMD) algorithm allows optimal sampling of the energy landscape. Although REMD efficiently samples the system's ensemble of conformations, the result can be biased by the selection of forcefields. That is, our results are valid within the context of the model. To overcome this limitation, we perform the simulations using various forcefields and validate our results with NMR $^3$J-coupling constants measured for this purpose. Our investigation of several forcefields concludes that OPLS[25] is efficient in reproducing our experimental measurements of $^3J_{\mathrm{HN}\alpha}$ for Aβ42.

Structural differences of the conformations of Aβ40 and 42 are evaluated and direct comparison with NMR data is made. Our results suggest that: 1) Aβ40 possesses unique structural features, namely a short N-terminal $3_{10}$ helix and a γ-hairpin spanning residues 12−18. 2) A flexible C-terminus domain of Aβ40 exists, at residues 25−40, that is partially stabilized by β-turns and

transient polyproline II (PPII) type structures. 3) The central hydrophobic cluster $^{17}$LVFF$^{21}$ A acts as an "intermolecular glue" within the monomer, through its interactions with several residues C-terminal to its position as well as with the N-proximal twin Histidine motif. 4) Aβ42 comprises a more diverse conformational ensemble, however its C-terminus is constrained to a β-hairpin structure. The formation of this beta structure may be responsible for Aβ42 higher propensity to form amyloids.

## Results

### Validation of the forcefield model with measured J-coupling constants

The Alzheimer's peptides encompass several characteristics of a disordered system[19; 26]. However, a recent study has argued that these flexible, collapsed peptides also possess order, diverging from the typical "random coil" behavior[27]. The first step to characterize the conformations adopted by the Aβ peptides is to assess the ability of various forcefields to capture the main features of these twilight systems. Therefore, we used experimentally determined J-coupling constants as a benchmark for the validity of the simulations.

We evaluated the performance of the different forcefields in simulations of Aβ42, and selected the one that best reproduced the experimental J-coupling constants, to perform REMD simulations on Aβ40 and to characterize the ensemble of conformations for both peptides using the OPLS forcefield. Our choice of the OPLS forcefield for Aβ40 was justified by the agreement of calculated J-coupling constants with their experimental values (figure 1).

A variety of forcefields was used in combination with the TIP3P or SPC water models. The characteristics of each simulation and performance of forcefields is summarized in table 1. In general, we observe a moderate agreement with experimental data for most forcefields. The AMBER family of forcefields[28] reproduce well the experimental J-couplings, but underestimate the values of most J-coupling constants. The resulting conformations are rich in short helices, and under populate the β basin of the Ramachandran plot (data not shown). The PARM94 forcefield overestimates the amount of α-helical structure, while PARM96 results in structures that are rich in β-turns. In particular, we observed a type I β-turn centered at residues 39 and 40 in 30% of all structures at normal temperature. The radius of gyration shows a bimodal distribution with an average value of 12A at low temperature. Conformations produced by the modified forcefield[29] are devoid of regular elements of secondary structure, yet show some formation of short PPII-type helices towards the C-terminus of Aβ42. The GROMOS forcefield[30] also reproduces well the experimental J-coupling constants, with conformations that are mainly extended. However, for the OPLS forcefield we observed the best agreement with experimental results (figure 1).

The agreement with experimental data for Aβ40 with the OPLS forcefield was higher than for Aβ42, reaching a Pearson's correlation coefficient of 0.71 for the full-length trajectories (figure 2). Furthermore, the convergence of calculated J-coupling constants to their experimentally determined values was used to assess the simulation time required for the system to equilibrate. In figure 2 we see that for the simulations of Aβ40 with the OPLS forcefield, calculated constants reach good agreement with experimental data after ~20ns/replica. However, according to the same criterion, the time required for Aβ42 to equilibrate with the same forcefield was much longer (~60ns/replica).

### Clustering of low temperature structures

We performed cluster analysis to separate the pool of low temperature (276−305K) conformations produced by the OPLS REMD simulation into clusters of similar geometric

properties. In this way we are able to capture the main structural features of the systems by focusing on a few structures that are representative of their clusters.

Results from different clustering methods applied in this study were found to be in qualitative agreement. In particular, for Aβ40, the central structures obtained from the k-means method were mainly extended conformations with no features of secondary structure, for small values of k (∼10). However, when the number of clusters was set equal to the number of clusters obtained by the Daura method (k=257), the central structures of the most populated clusters were collapsed, and resembled the dominant centroids of the former method (figure 3). Furthermore, the total number of structures in the prevailing clusters was approximately the same, covering ∼30% of the total sample size (6200 structures with k-means and 6865 with Daura's algorithm in a total of 21120 structures).

Clustering with the Daura method for Aβ40 resulted in two dominant clusters. The two corresponding centroids shared common structural features and were separated by backbone RMSD of 4.47A. Also, the Daura method yielded 43 clusters of intermediate size that contained ∼45% of all conformations and 212 small clusters with the remaining conformations. Approximately 50% of all conformations were contained in the 10 larger clusters. Results with the same method and RMSD cutoff of 3A, applied to the ensemble of Aβ42 conformations were more heterogeneous, in the sense that structures were more evenly distributed among clusters. However, when the RMSD threshold was optimized to a value of 5.5A, a dominant cluster containing 21% of all conformations was observed. In this case, the 10 larger clusters contained 61% of all conformations while the majority of the 96 clusters where of small to intermediate size (<5%). Clustering of Aβ40 conformations was found to be less dependent on the choice of RMSD cutoff, since change of the threshold from 3 to 5.5A resulted in a dominant central structure with the same overall features as before.

## Analysis of structures within clusters: Aβ40

Analysis of central structures uncovers a consensus of structural motifs; all centroids have a short $3_{10}$ helix at their N-terminal region (residues 1 to 4), and a hairpin-like conformation towards the center of the peptide (residues 12 to 18) (figure 4). A γ-turn[31] involving residues 14, 15 and 16 and a (i, i+2) backbone hydrogen bond, favors the divergence of the adjacent strands, to conclude in a C-terminal domain that is rich in β-turns and PPII type conformations. In particular, β-turns are well defined for residues 22−25, 24−27, 26−29 and a short PPII-type helix for residues 34−36. These features are invariant in the two dominant centroids, covering 21% and 11% respectively of all of the low temperature conformations. We label these clusters as cluster-1 and cluster-2. The main difference between the two clusters resides in the conformational variability of G29; in the first centroid this residue adopts a conformation at the β basin of the Ramachandran −122;140, while in the second the dihedral angles are 114; −19, thus resulting in divergence of the backbone towards opposite directions. Therefore, structures in the prevailing cluster-1 have a right-handed direction of the C-terminal backbone, while those of cluster-2 have a left-handed. Given the relative sizes of the two dominant clusters (21% vs 11%), we may argue that that the right-handed directionality is more stable.

Secondary structure in the ordered regions is enforced mainly by hydrogen bonds, as well as short-range contacts. In particular, hydrogen bonds are formed between the backbone amide and carboxyl groups of residues 1 and 4 in the short $3_{10}$ helix, as well as between the pairs of residues 12−18 (probability 100%) and 14−16 along the γ-hairpin. These H-bonds are present in almost all members of the dominant clusters (>95% for the H-bond involving residues 1 and 4, 100% for the two H-bonds between residues 12, 18 and > 90% for the H-bonds involving residues 14 and 16). The turns observed are open, in the sense that their stabilization is not relied upon an intra turn H-bond. A salt bridge involving K28 with any of its two partners, E22 or D23 is transiently observed within cluster members (<10% within cluster-1), however the

side chains of these residues are oriented in opposite directions in most low temperature ensembles.

In addition, analysis of contact maps for non-neighboring atoms of all members in the prevailing clusters reveals several key residues that are central in stabilizing the structure, through their interaction despite a large sequence separation (figure 4). For instance, short-range contacts are formed with high probability between H13 and V39, thus favoring the formation of a 27-residue long loop. We also observe a persistent hydrogen bond (present in 94% of all members in the dominant clusters) between the C-terminal carboxyl group of V40 and the basic side chain of R5 that brings the termini of the peptide in close proximity. Hydrophobic contacts between the two Valine residues at the C-terminus with A2 and H6 also contribute in the formation of a circular structure. In fact, both dominant clusters display an average end-to-end distance of approximately 8A.

Contact map analysis also reveals a region within Aβ40 that is extensively involved in intermolecular interactions. The high probability contour spanning residues 16−22 is indicative of a region that controls the energy landscape of Aβ40. In particular, K16 is prone to interact with H13 (probability ~93%) and L17 forms contacts with high probability (>90%) with residues 12−14 as well as V39. Furthermore, a very strong interaction between V18 and V12 can be observed in our simulation data. We see that the twin Histidine-containing motif VHH is essential in the formation of these contacts. In addition, residues 19 to 22 participate in several contacts, with probability ranging from 60−90%, that span most of the molecule's length. Overall, we notice that this segment, and more specifically the central hydrophobic cluster $^{17}$LVFF$^{21}$A, is the most interacting region within Aβ40.

Furthermore, we observe that the N-terminal region (residues 1−18) of the peptide is restricted by the aforementioned structural motifs, as opposed to the flexible C-terminal region (residues 19−40). In addition, the J-coupling constants that were calculated for this region are very close to their experimental values (figure 1). Taken together, these results suggest that the N-terminal region samples a much more narrow range of conformations, while most of the backbone flexibility resides on the second half of the structure.

## Comparison with Aβ42

The conformational ensemble for Aβ42 was much more diverse that that of Aβ40, resulting in several alternative central structures (figure 5). In general, the observed diversity of conformations can be attributed to the different topologies of the backbone trail. In the central structure of the largest cluster, the absence of any helical structure at the N-terminus of the peptide, as observed for Aβ40, is striking. Instead, the conformations demonstrate a disordered N-terminal tail, spanning residues 1 to 7. The γ-hairpin observed at residues 12−18 of Aβ40 is also absent. In Aβ42 the same region forms a loop involving a β turn at residues 12 to 15.

Furthermore, we observed a collapsed central region at residues 8−29 that is rich in loops and tight turns. A hydrophobic pocket formed by S8 and G9 accommodates a cross-over of the backbone trail through a chain-to-chain contact with E22 and D23, to result in an extended loop at residues 8−23. Short-range contacts among these dipeptide regions are formed with high probability (>99%), as illustrated in the corresponding contact map. These regions were observed to have perfect complementarity of molecular surfaces, an important factor for the formation of a tight Van der Waals lock[32]. At the contact map we can also observe a high probability of contact formation between the partners D23 and S26, which demonstrates the persistence of a tight turn centered at these residues. In fact, this turn is promoted by the interaction of the side chain carbonyl of D23 and backbone amide group of S26, (asx turn), as classified in[33]. In addition, a network of hydrogen bonds involving the backbone groups of

G25, K28 G29 is crucial in stabilizing a type IV β turn at positions 25−28 or alternatively an $I\alpha_{RS}$ turn at positions 25−29.

Interestingly, as opposed to the N-terminus of the peptide, the C-terminus of Aβ42 is more structured than that of Aβ40. The formation of a β-hairpin in the sequence $^{31}$**IIGL**MVG**GVVI**$^{42}$A involving short strands at residues 31−34 and 38−41 reduces the C-terminal flexibility of the peptide. A type $VI_B$ β-turn, centered at residues 35 and 36 is important for the alignment of the strands. This turn is mainly open, in the sense that the hydrogen bond between L34 and G37 is transiently formed (∼10%). Backbone hydrogen bonds between the partners A30;A42, I32;V40 and L34;G38 are frequently observed and provide additional stabilization of the hairpin structure.

Conformations within the smaller clusters are less ordered. A common structural theme among the clusters is the coiling of the backbone chain back to itself to form a stable loop that is enforced by backbone hydrogen bonds involving a variety of partners in the adjacent strands. The length of this C-terminal loop might vary from 10−13 residues among different centroids, with a β-turn forming at alternative positions in the region 33−37. Interactions between the two termini are also commonly observed, however the aminoacid partners differ among the clusters; in the dominant cluster distal interactions are observed between the regions 10−13 and 35−38, however in the following clusters these contacts involve the partners 5;39 and 8;34 respectively. In conclusion, as opposed to Aβ40, most of the structure of Aβ42 resides in its C-terminus.

## Discussion

Our analysis confirms the existence of structured regions within the otherwise flexible Aβ peptides. Brief, regular elements of secondary structure are observed for both Aβ 40 and 42. Several experimental results support this view[20; 21], however, in an NMR study of the Aβ10−35 congener Zhang et al. argue that hydrogen bonding is not a significant force in shaping the conformational landscape of Aβ, and emphasize on the importance of Van der Waals and electrostatic interactions. Perhaps hydrogen bonding interactions become more important in the full-length system.

Furthermore, our results are in both quantitative agreement with experimentally determined J-coupling constants, as well as qualitative agreement with structural models that are based on solution studies of the Aβ. In particular, our simulations of Aβ 40 yield structural features that have been observed in NMR studies: The classification of the conformations in two dominant clusters and their characteristics closely resembles the two families of clockwise and counterclockwise structures observed in[17]. Interestingly, both families were consistent with NMR constraints; however our results show that the right-handed conformation is significantly more stable. In the same study, the authors suggest the existence of β-turns at positions 22−25 and 24−27 as confirmed in the low temperature MD conformations. However, the authors also observed a β-turn at positions 27−30 that was not observed in our simulations. The region 22−30 seems to be prone to turn formation, as suggested by a NMR study on the protease resistant segment Aβ21−30[34], and corroborated by MD simulations on the same fragment[35]. Furthermore, our observation of a flexible C-terminal domain is in line with CD investigations of Aβ40 structure[26]. Having confirmed these observations, our analysis aims at providing intuition on detailed structural features of the system.

Another confirmation of our simulation data can be obtained from comparison with the 3D profile method[36]. In that study, Eisenberg and coworkers implemented a scanning/threading method to predict amyloidogenic segments in protein sequences. When applied to the sequence of Aβ42, their method identifies two regions of low energy in the fibril state: Residues 12−23

and the C-terminal segment, $^{37}$GGVVIA. The first position coincides with the short γ-hairpin and central hydrophobic cluster that we observed in the low temperature ensemble of Aβ40, while the second was within the beta turn region described above. However, the absence of a high contact density contour for the central hydrophobic cluster in Aβ42 signifies the role of structural context in determining the function of this motif.

Other theoretical works on Aβ also compare their results with experimental measurements. In a previously published study, Baumketner and coworkers perform a direct comparison of the conformations produced by REMD simulations with ion-mobility measurements[37]. The effect of solvation is evaluated and cluster analysis is performed with satisfactory results. However, their analysis is limited by the short timescale of their simulations (20 nsec/replica, for 20 replicas, resulting in a dataset of 355 conformations, versus 100ns/replica, for 52 replicas, and 21,120 conformations in our ensemble). In the current study we show that, based on comparison with the experimental J-coupling constants, the same system requires at least 60 ns to equilibrate. In another theoretical work, Massi and Straub compute several experimental observables including the hydrodynamic radius, proton chemical shifts and order parameters for the fragment 10−35[38]. In the former calculation however, the authors implement the Lipari-Szabo model free formalism that relies upon the assumption of separability between globular and intermolecular motions[39]. This is clearly not the case for Aβ, as it is an intrinsically flexible system[17].

The microsecond timescale of our simulations, in combination with the Replica Exchange method, is an important factor in exploring the accessible regions of Aβ's energy landscape. In addition, to the best of our knowledge, this is the first comparative simulation study of the two major full-length variants of Aβ in water. In a pioneering solution NMR study, Riek and coworkers suggest the structural similarity between the ensemble conformations of the two monomers, at the time resolution of their experiments[20]. In fact, the only difference they report is the larger NOE values for the C-terminus of Aβ42 versus 40 that is an indication of lower flexibility for that region. Although Riek conducted experiments on M35-oxidized Aβ42, the same observation has been confirmed for the M35-reduced peptide in a previous study by our group[21]. In line with the above measurements, we observe a significantly more structured C-terminus for Aβ42. Furthermore, we attribute the increased stability to the coiling of the backbone chain to itself and subsequent interactions involving hydrogen bond partners at the segment 31−41. In addition, the type II β-turn observed in the members of the larger cluster at position 34−37 signifies the critical position of M35, as it is found in a PPII conformation at the center of the turn.

Finally, the importance of a central hydrophobic region (residues 17−21) for the fibrillogenic activity of the 40-residue length monomer can also be inferred from our simulation data: In the relevant contact map we observe that this hydrophobic cluster is highly prone to interact with other regions of the molecule (high density contour). Contacts are formed along the entire sequence of Aβ40 with high probability, especially with the C-terminal domain (residues 25 −40) (figure 4). This observation is in line with the high order parameter and large number of dipolar coupling measurements for the same segment of Aβ40[17]. We can anticipate that when the concentration of the monomer becomes critical, it is this region that promotes the formation of intermolecular contacts, driven by the exclusion of water in the interface between monomers. In fact, this pentapeptide fragment of Aβ has been found to act as a binding sequence that inhibits fibril formation, through its interaction with the full length monomer[40]. The same study signifies the importance of L17 for this interaction, although its counterparts on the full-length monomer were not identified. Our analysis corroborates these findings and suggests the interacting partners. Given the pharmacological interest of peptide ligands as potential drug candidates against amyloids[16], our results provide insight to the mechanism of their action.

We hope our simulation results will motivate experimental studies aiming in designing more efficient inhibitors of fibril formation, through the information gained at the monomer level.

## Materials

All Molecular Dynamics simulations were performed with the GROMACS[41] package that was modified to accommodate the Replica Exchange scheme. The REMD method[42] consists of several identical copies of the system, or replicas, being simulated in parallel over a range of temperatures. To optimize the temperature spacing of the replicas, several short constant temperature simulations were performed for each system at different temperatures The histograms of potential energy obtained from these short trajectories were then used to define the temperatures of the replicas, such that the probability of exchange is constant throughout the range of temperatures[43]. The range of temperatures for each simulation is shown at table 1. Analysis was performed only on conformations sampled by all replicas at temperatures in the range 276−305K. The forcefields used in this study are summarized in table 1 and better characterized in[29]. For the GROMOS simulations we used the SPC water model[44]. The PARM94 and all other AMBER-derived forcefields, as well as the OPLS forcefield were combined with the TIP3P water model[45].

The following procedure was used to construct each system: First, we run a 10ns MD simulation of the peptide *in vacuo*, at high temperature (~700K) starting from a completely extended conformation. The collapsed peptide was then solvated in a cubic box, whose dimensions were adjusted to accommodate 3629 water molecules. The solvated system was then equilibrated at constant pressure (1 atm) for 2 ns with a short integration time step of 0.1fs. Finally, REMD simulations at constant volume were run for various times, depending on the choice of forcefield (table 1). At this stage, the application of the LINCS[46] and SETTLE[47] algorithms to constrain the bond lengths in the peptides and water molecules respectively allowed a relatively large integration step of 2 fs. The system was coupled to a Nose-Hoover[48] heat bath to maintain a constant temperature between exchanges. Simulations were run on 52 CPUs of a Linux-cluster at Rensselaer.

Several clustering algorithms were implemented[49], based on the matrix of RMSD as a measure of the distance between two peptide conformations. Ideally, we wanted to use a clustering scheme that classifies the pool of conformations into a few, large clusters with global, representative structures. The algorithm described by Daura[50] meets these qualifications, since it is based on a winner-takes-all criterion, according to which a structure and all its neighbors (a cluster) are removed from the pool of conformations in an order defined by the size of the cluster. Thus, we are able to characterize each cluster by one central structure (centroid), since all structures in that cluster are within a cutoff in RMSD. Since the total numbers of clusters as well as their representative centroids depend on the choice of this threshold, in this study we used the total number of structures contained within the 10 larger structures as a criterion to optimize the cutoff. The final RMSD threshold (3A for Aβ40 and 5.5A for Aβ42) was chosen to be the inflection point in the plot of coverage (i.e. the number of conformations contained within the 10 largest clusters) vs RMSD cutoff.

Contact maps were generated according to the following algorithm: For any two residues i,j that are separated in sequence by more than two residues, a contact is considered formed if any atom of residue i is within a cutoff from any atom of residue j. The general features of the contact maps where found to be insensitive of the exact cutoff value (results in figures 4,5 are displayed for a cutoff of 4.5A). Backbone hydrogen bonds were considered formed if the distance between then non-bonding atoms was less than 3.5 A, and salt bridges if charged atoms were within 4 A. Tight turns were characterized according to the criteria reviewed in[51].

The experimentally determined J-coupling constants were used to assess the validity of the simulations (table 1). We used solution NMR to measure the J-coupling constants between the amide protons of the peptide bonds and the corresponding α protons, for several residues of both peptides (21 in Aβ42 and 24 in Aβ40). All NMR experiments were performed on a 800 MHz spectrometer equipped with a cryoprobe at the NMR facility in the Center for Biotechnology and Interdisciplinary Studies at Rensselaer. Sample preparations and experimental conditions were as described in[21]. Experimental results were in turn used to compare with calculated $^{3}J_{\mathrm{HNH}\alpha}$ constants from our simulation data. Calculations were made by applying the Karplus equation with fitted coefficients[52; 53] on the peptide dihedral angles for each frame and taking the ensemble average. Comparison with the experimental J-couplings was also used to define the boundary between the equilibration and production phases for the OPLS simulations (figure 2), as summarized in table 1.

To have a statistical estimate of the simulation error of the calculations, we used block averages[54]. According to this method, a sample size of 10 ns/replica was found to be sufficient to estimate the error bars of the calculated J-coupling constants. We used 4 samples of 10 ns each for the OPLS trajectories, and three samples spanning the second half of the simulation for other forcefields. As a measure of agreement with experimental data we used the Pearson's Correlation Coefficient, as well as the average correlation index (C.I.), i.e the absolute deviation from experimental values over the sum of errors. A correlation index of one or less is considered to be a good indication of agreement with experimental data.

#### Acknowledgements

The authors acknowledge financial support by Rensselaer, NSF (MCB 0543769 and DMR 0117792), and the NIH Molecular Libraries Roadmap Initiative Grant 1P20HG003899−01. CW gratefully acknowledges funding from the Alzheimer's Association and the James D. Watson Young Investigator program of NYSTAR.

## References

1. Reinhard C, Hebert SS, De Strooper B. The amyloid-beta precursor protein: integrating structure with biological function. Embo J 2005;24:3996–4006. [PubMed: 16252002]
2. Turner PR, O'Connor K, Tate WP, Abraham WC. Roles of amyloid precursor protein and its fragments in regulating neural activity, plasticity and memory. Prog Neurobiol 2003;70:1–32. [PubMed: 12927332]
3. Ramsden M, Plant LD, Webster NJ, Vaughan PF, Henderson Z, Pearson HA. Differential effects of unaggregated and aggregated amyloid beta protein (1−40) on K(+) channel currents in primary cultures of rat cerebellar granule and cortical neurones. J Neurochem 2001;79:699–712. [PubMed: 11701773]
4. Kamenetz F, Tomita T, Hsieh H, Seabrook G, Borchelt D, Iwatsubo T, Sisodia S, Malinow R. APP processing and synaptic function. Neuron 2003;37:925–37. [PubMed: 12670422]
5. Walsh DT, Montero RM, Bresciani LG, Jen AY, Leclercq PD, Saunders D, AN EL-A, Gbadamoshi L, Gentleman SM, Jen LS. Amyloid-beta peptide is toxic to neurons in vivo via indirect mechanisms. Neurobiol Dis 2002;10:20–7. [PubMed: 12079400]
6. Soto C. Unfolding the role of protein misfolding in neurodegenerative diseases. Nat Rev Neurosci 2003;4:49–60. [PubMed: 12511861]
7. Walsh DM, Hartley DM, Kusumoto Y, Fezoui Y, Condron MM, Lomakin A, Benedek GB, Selkoe DJ, Teplow DB. Amyloid beta-protein fibrillogenesis. Structure and biological activity of protofibrillar intermediates. J Biol Chem 1999;274:25945–52. [PubMed: 10464339]
8. Morgan C, Colombres M, Nunez MT, Inestrosa NC. Structure and function of amyloid in Alzheimer's disease. Prog Neurobiol 2004;74:323–49. [PubMed: 15649580]
9. Narayanan S, Reif B. Characterization of chemical exchange between soluble and aggregated states of beta-amyloid by solution-state NMR upon variation of salt conditions. Biochemistry 2005;44:1444–52. [PubMed: 15683229]

10. Kirkitadze MD, Condron MM, Teplow DB. Identification and characterization of key kinetic intermediates in amyloid beta-protein fibrillogenesis. J Mol Biol 2001;312:1103–19. [PubMed: 11580253]

11. Klimov DK, Thirumalai D. Dissecting the assembly of Abeta16−22 amyloid peptides into antiparallel beta sheets. Structure 2003;11:295–307. [PubMed: 12623017]

12. Luhrs T, Ritter C, Adrian M, Riek-Loher 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;102:17342–7. [PubMed: 16293696]

13. Morimoto A, Irie K, Murakami K, Masuda Y, Ohigashi H, Nagao M, Fukuda H, Shimizu T, Shirasawa T. Analysis of the secondary structure of beta-amyloid (Abeta42) fibrils by systematic proline replacement. J Biol Chem 2004;279:52781–8. [PubMed: 15459202]

14. Petkova AT, Ishii Y, Balbach JJ, Antzutkin ON, Leapman RD, Delaglio F, Tycko R. A structural model for Alzheimer's beta - amyloid fibrils based on experimental constraints from solid state NMR. Proc Natl Acad Sci U S A 2002;99:16742–7. [PubMed: 12481027]

15. Williams AD, Portelius E, Kheterpal I, Guo JT, Cook KD, Xu Y, Wetzel R. Mapping abeta amyloid fibril secondary structure using scanning proline mutagenesis. J Mol Biol 2004;335:833–42. [PubMed: 14687578]

16. Soto C, Sigurdsson EM, Morelli L, Kumar RA, Castano EM, Frangione B. Beta-sheet breaker peptides inhibit fibrillogenesis in a rat brain model of amyloidosis: implications for Alzheimer's therapy. Nat Med 1998;4:822–6. [PubMed: 9662374]

17. Zhang S, Iwata K, Lachenmann MJ, Peng JW, Li S, Stimson ER, Lu Y, Felix AM, Maggio JE, Lee JP. The Alzheimer's peptide a beta adopts a collapsed coil structure in water. J Struct Biol 2000;130:130–41. [PubMed: 10940221]

18. Jarrett JT, Lansbury PT Jr. Seeding “one-dimensional crystallization” of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie? Cell 1993;73:1055–8. [PubMed: 8513491]

19. Hou L, Shao H, Zhang Y, Li H, Menon NK, Neuhaus EB, Brewer JM, Byeon IJ, Ray DG, Vitek MP, Iwashita T, Makula RA, Przybyla AB, Zagorski MG. Solution NMR studies of the A beta(1−40) and 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. [PubMed: 14971932]

20. Riek R, Guntert P, Dobeli H, Wipf B, Wuthrich K. NMR studies in aqueous solution fail to identify significant conformational differences between the monomeric forms of two Alzheimer peptides with widely different plaque-competence, A beta(1−40)(ox) and A beta(1−42)(ox). Eur J Biochem 2001;268:5930–6. [PubMed: 11722581]

21. Yan Y, Wang C. Abeta42 is More Rigid than Abeta40 at the C Terminus: Implications for Abeta Aggregation and Toxicity. J Mol Biol. 2006

22. Barrow CJ, Yasuda A, Kenny PT, Zagorski MG. Solution conformations and aggregational properties of synthetic amyloid beta-peptides of Alzheimer's disease. Analysis of circular dichroism spectra. J Mol Biol 1992;225:1075–93. [PubMed: 1613791]

23. Jarrett JT, Berger EP, Lansbury PT Jr. The C-terminus of the beta protein is critical in amyloidogenesis. Ann N Y Acad Sci 1993;695:144–8. [PubMed: 8239273]

24. Bitan G, Kirkitadze MD, Lomakin A, Vollers SS, Benedek GB, Teplow DB. Amyloid beta -protein (Abeta) assembly: Abeta 40 and Abeta 42 oligomerize through distinct pathways. Proc Natl Acad Sci USA 2003;100:330–5. [PubMed: 12506200]

25. Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL. 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.

26. Danielsson J, Jarvet J, Damberg P, Graslund A. The Alzheimer beta-peptide shows temperature-dependent transitions between left-handed 3-helix, beta-strand and random coil secondary structures. FEBS J 2005;272:3938–49. [PubMed: 16045764]

27. Schweitzer-Stenner R, Measey T, Hagarman A, Eker F, Griebenow K. Salmon calcitonin and amyloid beta: two peptides with amyloidogenic capacity adopt different conformational manifolds in their unfolded states. Biochemistry 2006;45:2810–9. [PubMed: 16503636]

28. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A 2nd Generation Force-Field for the Simulation of Proteins, Nucleic-Acids, and Organic-Molecules. J Amer Chem Soc 1995;117:5179–5197.
29. Garcia AE, Sanbonmatsu KY. Alpha-helical stabilization by side chain shielding of backbone hydrogen bonds. Proc Natl Acad Sci U S A 2002;99:2782–7. [PubMed: 11867710]
30. Scott WRP, Hunenberger PH, Tironi IG, Mark AE, Billeter SR, Fennen J, Torda AE, Huber T, Kruger P, van Gunsteren WF. The GROMOS biomolecular simulation program package. J Phys Chem A 1999;103:3596–3607.
31. Nemethy G, Printz MP. Gamma-Turn, a Possible Folded Conformation of Polypeptide Chain - Comparison with Beta-Turn. Macromolecules 1972;5:755.
32. Berezovsky IN, Trifonov EN. Van der Waals locks: loop-n-lock structure of globular proteins. J Mol Biol 2001;307:1419–26. [PubMed: 11292352]
33. Duddy WJ, Nissink JW, Allen FH, Milner-White EJ. Mimicry by asx- and ST-turns of the four main types of beta-turn in proteins. Protein Sci 2004;13:3051–5. [PubMed: 15459339]
34. Lazo ND, Grant MA, Condron MC, Rigby AC, Teplow DB. On the nucleation of amyloid beta-protein monomer folding. Protein Sci 2005;14:1581–96. [PubMed: 15930005]
35. Baumketner A, Bernstein SL, Wyttenbach T, Lazo ND, Teplow DB, Bowers MT, Shea JE. Structure of the 21−30 fragment of amyloid beta-protein. Protein Sci 2006;15:1239–47. [PubMed: 16731963]
36. Thompson MJ, Sievers SA, Karanicolas J, Ivanova MI, Baker D, Eisenberg D. The 3D profile method for identifying fibril-forming segments of proteins. Proc Natl Acad Sci U S A 2006;103:4074–8. [PubMed: 16537487]
37. Baumketner A, Bernstein SL, Wyttenbach T, Bitan G, Teplow DB, Bowers MT, Shea JE. Amyloid beta-protein monomer structure: a computational and experimental study. Protein Sci 2006;15:420–8. [PubMed: 16501222]
38. Massi F, Peng JW, Lee JP, Straub JE. Simulation study of the structure and dynamics of the Alzheimer's amyloid peptide congener in solution. Biophys J 2001;80:31–44. [PubMed: 11159381]
39. Lipari G, Szabo A. J Am Chem Soc 1982:4546–4559.
40. Tjernberg LO, Naslund J, Lindqvist F, Johansson J, Karlstrom AR, Thyberg J, Terenius L, Nordstedt C. Arrest of beta-amyloid fibril formation by a pentapeptide ligand. J Biol Chem 1996;271:8545–8. [PubMed: 8621479]
41. Lindahl E, Hess B, van der Spoel D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. J Mol Mod 2001;7:306–317.
42. Sugita Y, Okamoto Y. Replica-exchange molecular dynamics method for protein folding. Chem Phys Lett 1999;314:141–151.
43. Garcia AE, Herce H, Paschek D. Simulations of Temperature and Pressure Unfolding of Peptides and Proteins with Replica Exchange Molecular Dynamics. Annu Rep Comput Chem 2006;2:83–95.
44. Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J. Interaction models for water in relation to protein hydration. Nature 1969;224:175–177. [PubMed: 5343518]
45. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of Simple Potential Functions for Simulating Liquid Water. J Chem Phys 1983;79:926–935.
46. Hess B, Bekker H, Berendsen HJC, Fraaije J. LINCS: A linear constraint solver for molecular simulations. J Comp Chem 1997;18:1463–1472.
47. Miyamoto S, Kollman PA. Settle - an Analytical Version of the Shake and Rattle Algorithm for Rigid Water Models. J Comp Chem 1992;13:952–962.
48. Nose S. Constant Temperature Molecular-Dynamics Methods. Prog Theor Phys Suppl 1991:1–46.
49. Hartigan, J. Clustering Algorithms. John Wliey & Sons; New York: 1975.
50. Daura X, Suter R, van Gunsteren WF. Validation of molecular simulation by comparison with experiment: Rotational reorientation of tryptophan in water. J Chem Phys 1999;110:3049–3055.
51. Chou KC. Prediction of tight turns and their types in proteins. Anal Biochem 2000;286:1–16. [PubMed: 11038267]
52. Karplus M, Anderson DH. Valence-Bond Interpretation of Electron-Coupled Nuclear Spin Interactions - Application to Methane. J Chem Phys 1959;30:6–10.

53. Vuister GW, Bax A. Quantitative J Correlation - a New Approach for Measuring Homonuclear 3-Bond J(H(N)H(Alpha) Coupling-Constants in N-15-Enriched Proteins. J Am Chem Soc 1993;115:7772–7777.

54. Flyvbjerg H, Petersen HG. Error-Estimates on Averages of Correlated Data. J Chem Phys 1989;91:461–466.

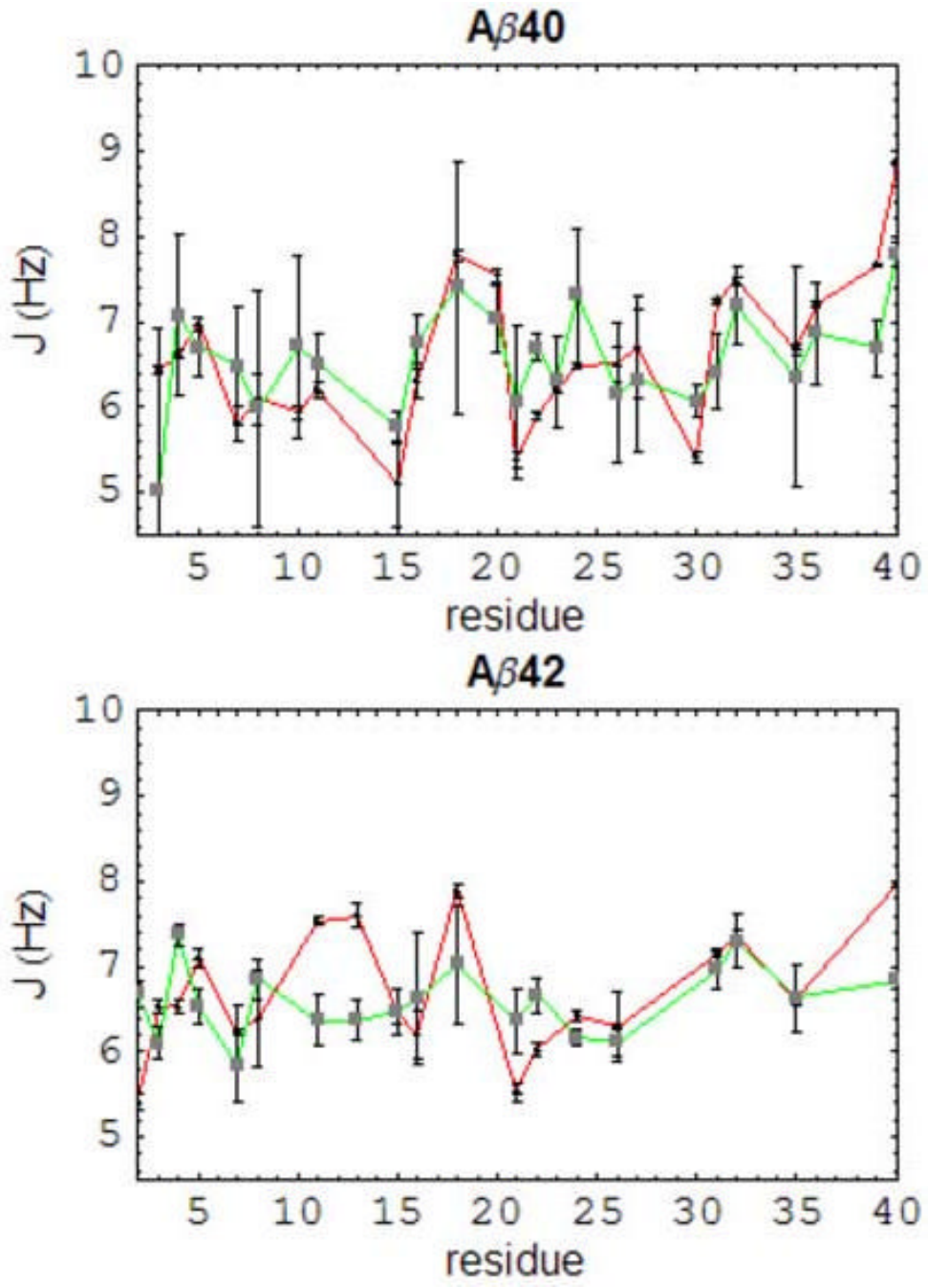


**Figure 1.**

**Comparison with experimental data.** Measured $^{3}J_{\mathrm{HNH}\alpha}$ constants (triangles) are compared with their calculated values (squares) from the REMD low temperature trajectory segments (275−305K). Simulation errors are the differences between the larger and lower values among results computed for 4 samples of length 10nsec/replica. Results are presented for the simulations under the best performing forcefield, OPLS. Glycine residues are not included, since the experimental values of the $^{3}J_{\mathrm{HNH}\alpha}$ constants for these residues are ambiguous.

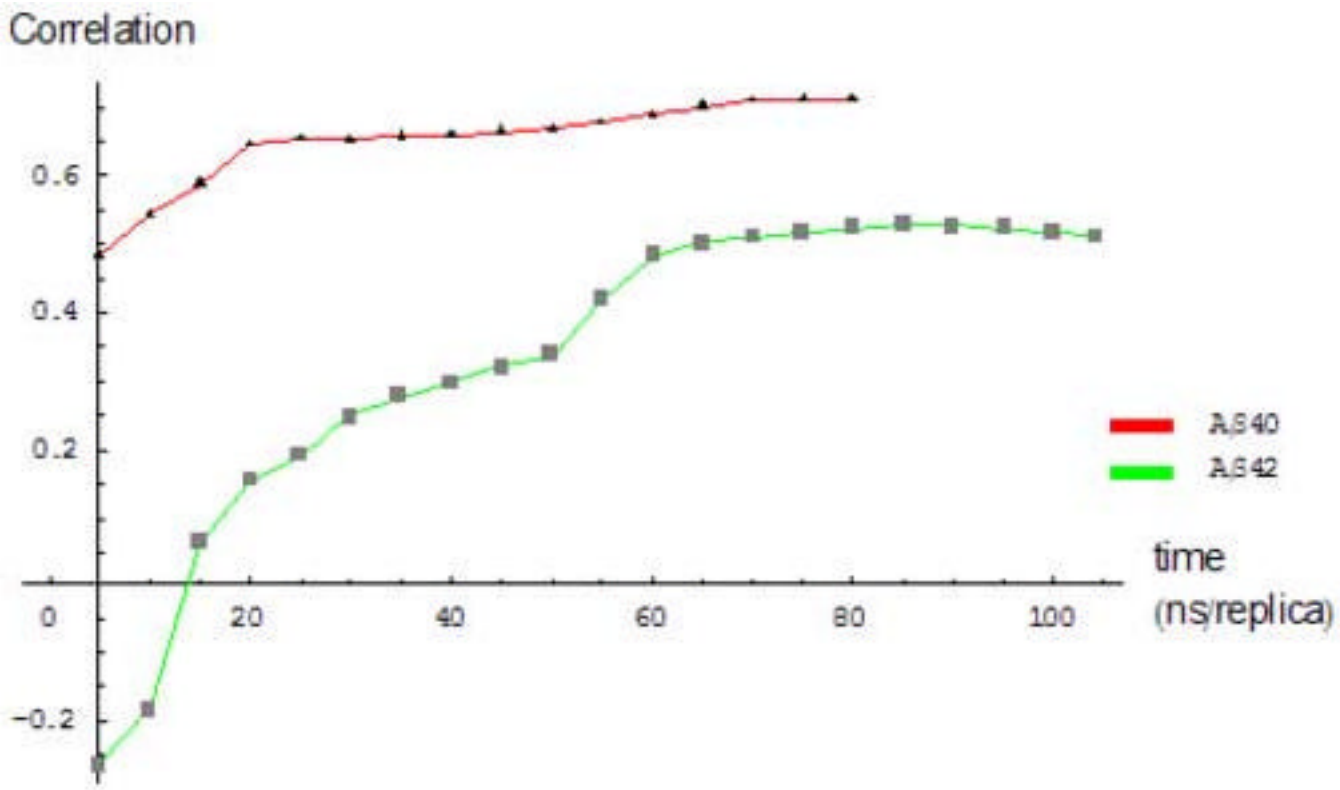


**Figure 2.**

**Convergence of calculated $^{3}J_{HNH\alpha}$ constants towards their measured values.** Pearson's correlation coefficient with experimental data is presented as a function of the total simulation time for Aβ40 (triangles) and 42 (squares). Cumulative average values of the J-coupling constants were computed every 5ns/replica, as described in materials.

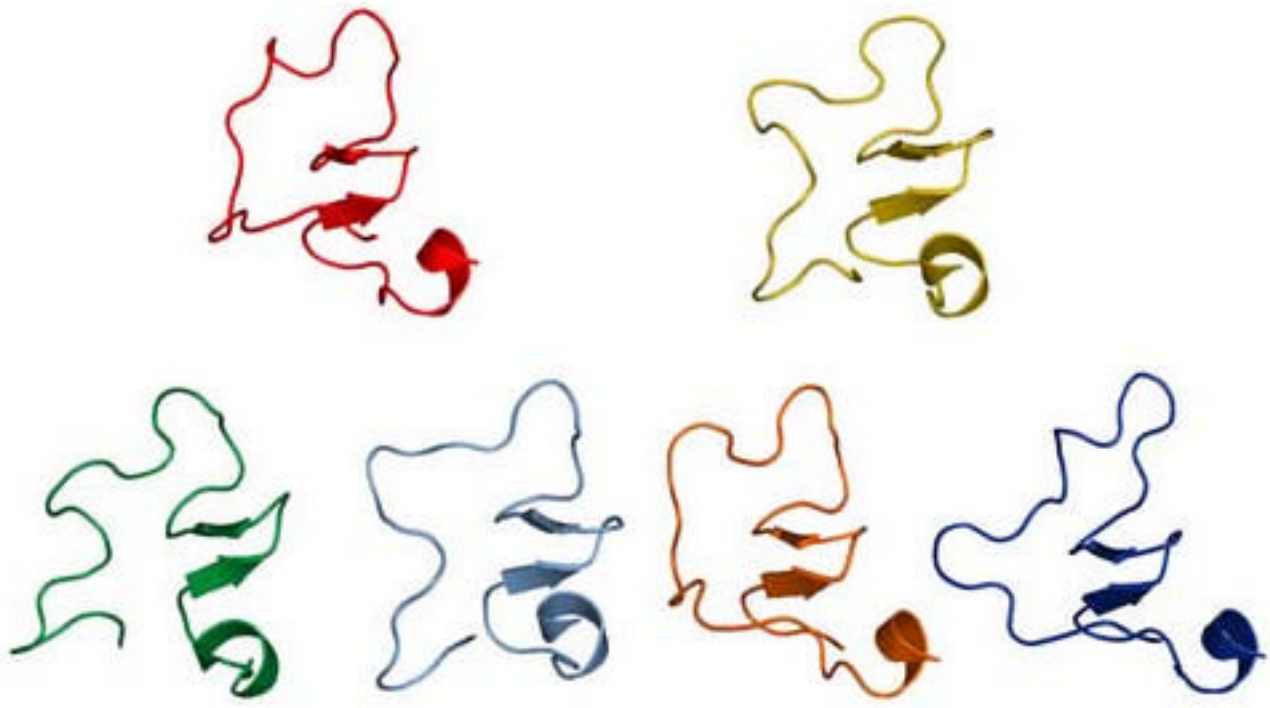

**Figure 3.**

**Representative conformations of the Aβ40 trajectories as selected with different clustering methods**. Central structures from the Daura clustering method (1st row of structures) are structurally comparable to those produced by the k-means algorithm (2nd row), when k was set equal to the number of clusters obtained with the first method. These central structures are representative of their corresponding clusters. In each row, the structures are presented from left to right in an order that corresponds to the relative size of the cluster they represent, as described in text. The original dataset consists of 21120 conformations, obtained form the low temperature REMD trajectory segments. The position of the N and C-termini can be traced through the direction of the cartoon arrows that correspond to β-sheets.

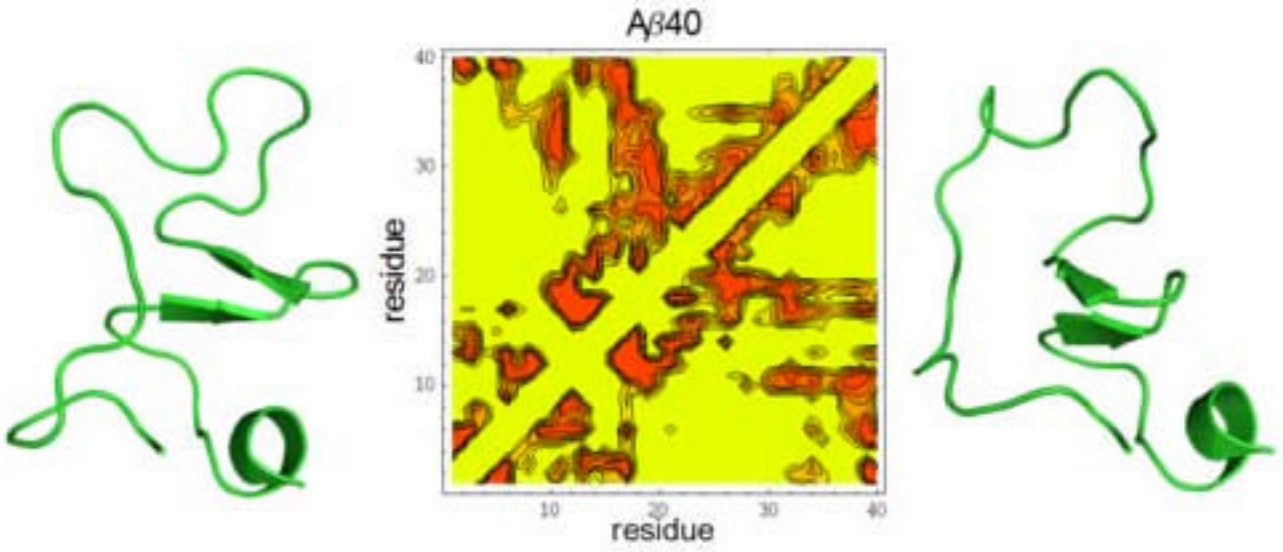


**Figure 4.**

**Representative conformations within the ensemble of Aβ40.** Central structures from the two dominant clusters are displayed next to the contact maps corresponding to all members of these clusters. They consist of collapsed structures with a short helical region in the N-terminus of the peptide, a short γ-hairpin towards the center of the structure and a disordered C-terminus that can obtain either a right-handed (left) or left-handed conformation (right) of the backbone trail. The sizes of the clusters displayed here cover 21% and 11% of the total population respectively. Contours are based on a log scale of counts, while red color denotes a high probability of contact formation.

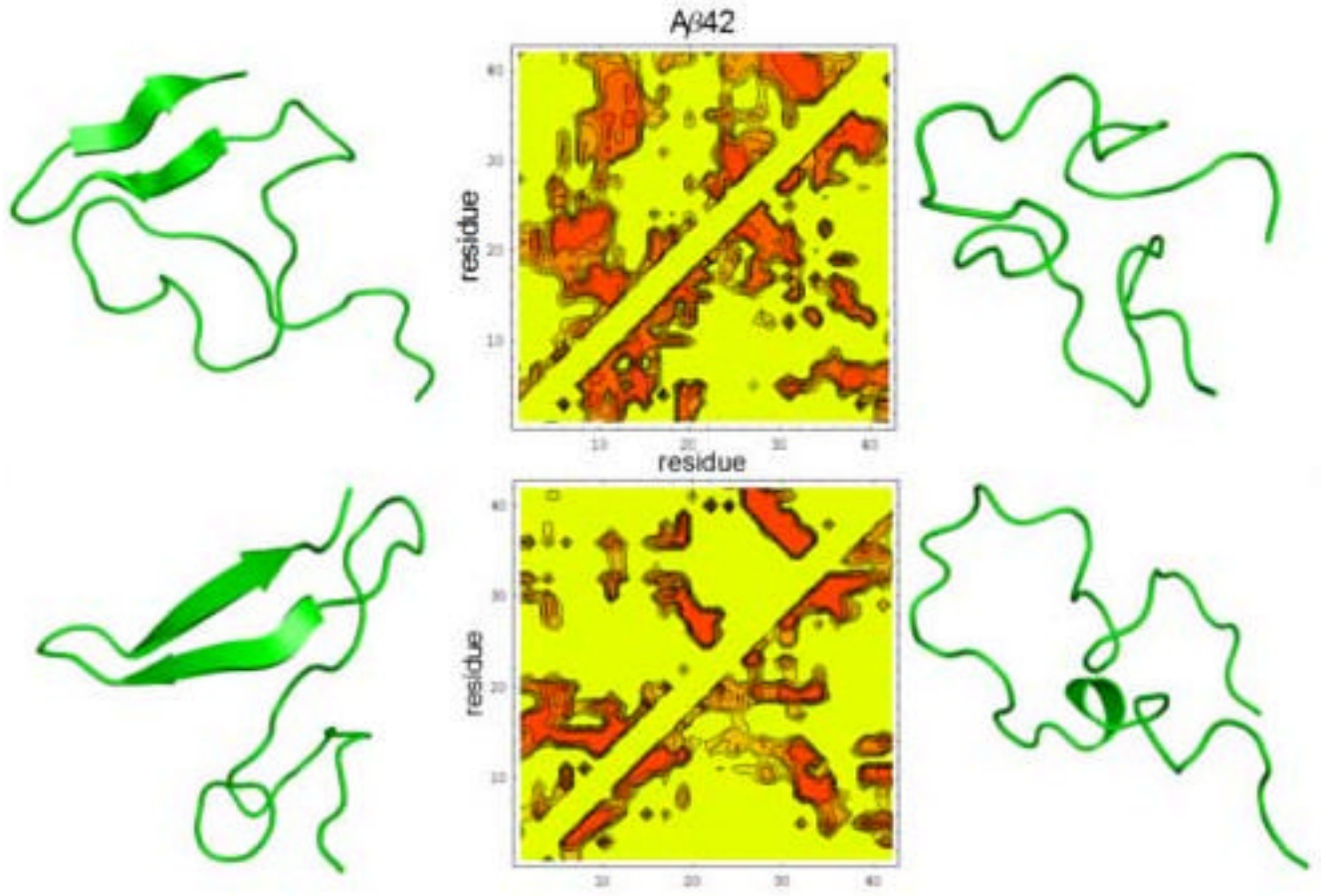


**Figure 5.**

**Structural characterization of the ensemble of Aβ42.** Central structures are displayed next to the contact maps of their corresponding clusters. Each half of the contact maps describes a different cluster, as in figure 4. We observe a diversity of topologies of the backbone of the peptide; however the C-terminus is usually trapped in a β-hairpin or alternatively an extended loop. The ensemble is more diverse that that of Aβ40. Clusters are presented in a decreasing size order, from the upper left to the lower right of the figure. Their sizes cover 21%, 5.9%, 5.8% and 5.6% respectively on a total sample of 21120 conformations. In all images, the conformations are presented with the N-terminus down and the C-terminus up.

**Table 1**

**Simulation performance and details.** Forcefields were evaluated according to comparison with measured $^3J_{HNH\alpha}$ of Aβ42. For OPLS, values are presented for Aβ42/Aβ40. The Pearson's Correlation Coefficient and correlation index were used as a measure of agreement with experimental data. We observe low/moderate consistency for all forcefields, with the exception of OPLS, that efficiently reproduces the measured values. Times are the sum of all REMD simulations.

| | total time (μs) | Production time (μs) | T range (K) | <correlation index> | PCC |
|---|---|---|---|---|---|
| PARM94 | 2.34 | 0.78 | 270−654 | 7.3 | 0.38 |
| PARM96 | 1.872 | 0.78 | 270−654 | 7.3 | 0.27 |
| MOD-PARM | 1.872 | 0.78 | 270−654 | 7.7 | 0.17 |
| GROMOS | 1.716 | 0.78 | 260−574 | 5.8 | −0.01 |
| OPLS | 5.4/4.264 | 2.08/2.08 | 265−624 | 2.2/1.2 | 0.43/0.66 |