Published in final edited form as:
Biochemistry. 2011 September 6; 50(35): 7612–7628. doi:10.1021/bi200732x.

Homogeneous and Heterogeneous Tertiary Structure Ensembles of Amyloid-*β* Peptides

K. Aurelia Ball[†], **Aaron H. Phillips**[‡], **Paul S. Nerenberg**[§], **Nicolas L. Fawzi**[§,⊥], **David E. Wemmer**[†,‡,∥], and **Teresa Head-Gordon**[*,†,§,∥]

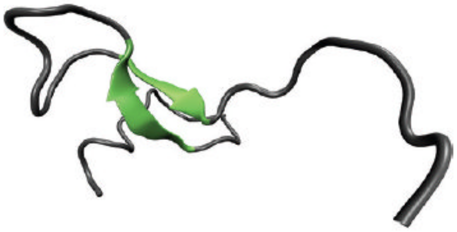
[†]Graduate Group in Biophysics, University of California, Berkeley, California 94720, United States

[‡]Department of Chemistry, University of California, Berkeley, California 94720, United States

[§]Department of Bioengineering, University of California, Berkeley, California 94720, United States

[∥]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States

Abstract



The interplay of modern molecular simulation and high-quality nuclear magnetic resonance (NMR) experiments has reached a fruitful stage for quantitative characterization of structural ensembles of disordered peptides. Amyloid-*β* 1–42 (A*β*42), the primary peptide associated with Alzheimer's disease, and fragments such as A*β*21–30 are both classified as intrinsically disordered peptides (IDPs). We use a variety of NMR observables to validate de novo molecular dynamics simulations in explicit water to characterize the tertiary structure ensemble of A*β*42 and A*β*21–30 from the perspective of their classification as IDPs. Unlike the A*β*21–30 fragment that conforms to expectations of an IDP that is primarily extended, we find that A*β*42 samples conformations reflecting all possible secondary structure categories and spans the range of IDP classifications from collapsed structured states to highly extended conformations, making it an IDP with a far more heterogeneous tertiary ensemble.

Intrinsically disordered proteins (IDPs), polypeptides that do not adopt a single or dominant tertiary conformation, are abundant in eukaryotic proteomes, comprising 25–30% of protein-encoding sequences found in the human genome.[1-4] Characterizing IDP structural diversity

---

[*]Corresponding Author tlhead-gordon@lbl.gov. Telephone: (510) 666-2744. Fax: (510) 642-5833.

[⊥]**Present Address**

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0520

Supporting Information

The Multiple-Reservoir Replica Exchange (MRRE) simulation technique is described in more detail. Medium-range A*β*42 NOEs that have been assigned from experiment alone are listed in Table S1. A*β*42 NOE false negatives and positives are listed in Tables S2 and S3. Convergence of the A*β*42 287 K ensemble is demonstrated in Figure S1. The calibration between A*β*42 experimental and calculated NOE intensities is shown in Figure S2. This material is available free of charge via the Internet at http://pubs.acs.org.

is important for understanding their role in functional signaling, regulation, and transcription, as well as their detrimental involvement in pathologies such as cancer and neurodegenerative diseases. Their characterization has presented new challenges in structural biology because traditional experimental techniques such as X-ray crystallography and electron microscopy cannot yield the same detailed structural information that has been paramount in our understanding of folded enzymes and well-formed protein complexes.[2]

Our current understanding of IDPs comes in part through contrasting their sequence and structural characteristics with those of folded proteins.[2] For example, IDP sequences are often significantly lower in hydrophobic residue content, have an over-representation of polar and/or charged amino acids, and have a low level of sequence complexity compared to globular, folded proteins.[3,5] IDPs often have little secondary structure compared to the (nondenatured) unfolded ensembles of structured protein at native or near native conditions that retain α-helical and β-strand segments. The structural diversity of IDPs suggests that conformational subpopulations interconvert at a rate much faster than the folded to unfolded transition time scale of a structurally ordered protein. While folded proteins can be usefully categorized into classes based on well-ordered secondary or tertiary structure, IDPs are classified by their degree of disorder, i.e., whether they predominately are collapsed semiordered ensembles (“structured” molten globules or pre-molten globule states), collapsed disordered ensembles, or extended disordered ensembles.[1,6] These classification schemes have been useful for IDPs with known function.

Amyloid-β 1–42 (Aβ42) is the primary peptide associated with Alzheimer’s disease (AD), generated in the amyloidogenic pathway by proteolytic cleavage of the transmembrane amyloid precursor protein (APP) by membrane-anchored β- and γ-secretases. While Aβ appears to interconvert between α-helical and random coil populations when it is part of the APP integral membrane protein,[7] the cleaved Aβ42 peptide and fragments such as Aβ21–30 have been categorized as IDPs.[8,9] Although both are classified as IDPs, it would be useful to know whether their structural ensembles are similar or dissimilar so that we might usefully interpret the physiological behavior of the full-length peptide or the physiological significance of the smaller Aβ21–30 fragment.

The interplay of modern molecular simulation and high-quality nuclear magnetic resonance (NMR) experiments has reached a fruitful stage for quantitative characterization of structural ensembles of disordered peptides, which we review here. We use a variety of NMR observables to validate de novo molecular dynamics simulations in explicit water to contrast the tertiary structure ensembles of Aβ21–30 and Aβ42. We show that Aβ42 samples conformations reflecting all possible secondary structure categories and spans the range of IDP classifications from collapsed structured states to highly extended conformations, while Aβ21–30 has a relatively homogeneous ensemble of extended structures, showing that the two amyloid-β peptides belong to very different IDP classes.

METHODS

NMR Experiments

The peptide preparation and NMR experiments with Aβ21–30 have been reported in a previous publication.[10] Here we report on the experimental details of Aβ42. Recombinant Aβ42 peptide was purchased from a commercial source (rPeptide, Athens, GA) and was reported to be >97% pure. The peptide was lyophilized from trifluoroacetic acid (TFA). The alkaline pretreatment of Aβ and the preparation of LMW Aβ by filtration protocols outlined by Teplow[11] were used to prepare a monomeric solution of Aβ42. The lyophilized peptide was dissolved in 2 mM NaOH to produce a peptide concentration of 0.21 mg/mL with a pH of >10.5. The peptide was then sonicated for 1 min in a bath sonicator and lyophilized. It was then resuspended in 20 mM sodium phosphate buffer (pH 7.2) and 0.01% (w/v) sodium azide. This protocol ensures that when the lyophilized peptide is dissolved in buffer it will not pass through its pI of ~5.31. Before NMR data were collected, the sample was filtered with a 0.22 μm filter to remove any fibril seeds and brought to a concentration of ~600 μM at pH 7.2.

NMR data were collected at 287 K at the NMR Facility at the University of California on Bruker Avance II 800 and 900 MHz spectrometers, the latter equipped with a Bruker cryogenic probe. The data were processed using NMRPipe, and peaks were assigned and analyzed using CARA.[12] Chemical shifts were assigned using two-dimensional (2D) ^{1}H–^{1}H total correlation spectroscopy (TOCSY)[13,14] and nuclear Over-hauser effect spectroscopy (NOESY).[15-17] The TOCSY spectra were recorded in a 90% H_2O (10% ^2H_2O) buffer solution with a mixing time of 60 ms and in a 99% ^2H_2O buffer solution with a mixing time of 80 ms. The NOESY spectra were recorded in a 90% H_2O (10% ^2H_2O) buffer solution with mixing times of 100 and 200 ms and in a 99% ^2H_2O buffer solution with a mixing time of 200 ms. In the direct dimension, 1024–2048 points were collected, while 256 complex points were collected in the indirect dimension for all spectra. The sweep width was 12 ppm in each dimension with 16–48 scans. One-dimensional (1D) spectra were recorded at time points before and after the experiments to measure the degree to which the monomeric peptide signal decreased because of aggregation. We also ensured that our NOESY data resulted only from a monomeric form of the peptide by collecting a NOESY spectrum immediately after dissolving the peptide and another NOESY spectrum ~48 h after dissolving the sample. We saw no difference between the two spectra except for a decrease in peak intensity due to aggregation, which eliminates the possibility that some NOEs could be a result of oligomers forming during peptide incubation. Peaks were identified at particular chemical shifts in the 200 ms mixing time NOESY spectra in both H_2O and ^2H_2O, and those peaks that could be unambiguously assigned to two specific hydrogen atoms were used to calibrate the NOE peak intensities derived from simulation. The majority of peaks could not be clearly assigned to a unique pair of hydrogen atoms because of spectral overlap. These peaks were instead assigned a list of potential hydrogen pairs consistent with the observed peak frequencies. This list consists of atoms within a 0.04–0.08 ppm range around each observed peak.

De Novo Molecular Dynamics Simulations

The de novo molecular dynamics simulations of Aβ21–30 have been reported in a previous publication.[10] Here we report on the simulation details of Aβ42. The peptide was modeled using the Amber ff99SB force field[18] and aqueous solvent represented by the TIP4P-Ew water model.[19] We chose the ff99SB force field and TIP4P-Ew water model combination because it was used in the previous Aβ21–30 study,[10] and subsequent studies support its clear superiority relative to other biomolecular simulation force fields.[10,20-23]

To calculate equilibrium ensemble averages in the NVT ensemble, we used an Andersen thermostat, a leapfrog integrator with a 1.0 fs time step, and periodic boundary conditions. Particle mesh Ewald was used for calculating long-range electrostatic forces, and a cutoff of 9.0 Å was used for the real space electrostatics and Lennard-Jones forces throughout the study. The initial structure for Aβ42 was built in an extended form using the *LEaP* module that is part of the AMBER package. Each structure was solvated in a water box such that there were 10 Å of water surrounding the peptide on all sides (15142 water molecules, box volume of 580 nm^3) and three Na^+ ions to neutralize the charge of the peptide. The structure was minimized and equilibrated with a constant volume while the temperature was increased to 300 K and then equilibrated for 2 ns under a constant pressure of 1 bar, maintained with a Berendsen barostat, to achieve the correct density. Then, a 2 ns, 498 K simulation was run on the extended Aβ42 peptide to obtain a more collapsed starting structure for equilibrium ensemble simulations. Two different collapsed structures were chosen to start two independent replica exchange simulations. These systems were prepared by first removing the solvent, and then the previous solvation and equilibration steps were used to create a final box that contained 6251 water molecules (193 nm^3 volume). The *sander* module of AMBER and the Multiple-Reservoir Replica Exchange (MRRE) method presented by Ruscio et al.[24] were used to achieve a Boltzmann weighted ensemble of 100000 Aβ42 structures at the experimental temperature of 287 K, from a 100 ns equilibrated simulation. Details of the MRRE calculations for Aβ42 are reported in the Supporting Information.

To calculate time correlation data on the 287 K reservoir, we used selected structures from this ensemble to run micro-canonical (NVE) ensemble simulations using the *pmemd* module of AMBER10. Fifty structures were chosen from each 287 K Aβ42 ensemble (after equilibration for 10 ns), spaced 1 ns apart (ensuring complete decorrelation through swaps between replicas), and equilibrated at a constant temperature of 287 K for 100 ps before a 20 ns simulation was run for each initial structure. In total, the Aβ42 constant energy simulation time with starting conformations equilibrated at 287 K was 2 μs.

Calculation of NMR Observables

For both Aβ21–30 and Aβ42, we evaluated NMR chemical shifts and scalar coupling constants,[25-27] NOE peak intensities for not only backbone amide proton contacts[8] but also the full set of hydrogen contacts that were measured by us, using the methods described by Fawzi et al.[10] For Aβ42, we also evaluated the residual dipolar couplings.[25,27,28]

Chemical Shifts—We used SHIFTS[29] to calculate chemical shifts for H_α and amide hydrogen atoms and carbon chemical shifts. These values were compared directly to the chemical shift (in parts per million) assigned to each hydrogen and carbon atom from the NMR spectra reported here, without any normalization. The C_α and C_β chemical shifts were also calculated and compared to experimental values from ref 8.

Scalar Couplings—The *ptraj* module of AMBER was used to calculate the ϕ angles for each residue of Aβ21–30 or Aβ42 at every snapshot in the 287 K ensemble, and the J coupling constant, $J(\phi)$, was evaluated for each snapshot using the Karplus equation[30]

$$J(\phi)=A\cos^2(\phi-60^\circ)+B\cos(\phi-60^\circ)+C \quad (1)$$

with coefficients A (6.51), B (−1.76), and C (1.60) corresponding to the original parameter set of Vuister and Bax.[31] We note that the ϕ values used to calculate the J coupling constants correspond to the dihedral angle defined by C^{i-1}, N^i, $C_\alpha{}^i$, and C^i, although the J splitting seen experimentally is a result of coupling between two hydrogen atoms, related by the dihedral angle defined by HN^i, N^i, $C_\alpha{}^i$, and $H_\alpha{}^i$. These two dihedral angles are geometrically related by a phase factor of ~60° as per eq 1. This discrepancy was not properly identified in a previous publication,[10] and we show the corrected calculation in Results. To obtain an overall metric of how well the calculated coupling constants correspond to the experimental data and compare this to other simulated systems, we calculate the figure of merit

$$\chi^2=\frac{1}{N}\sum_{i=1}^{N}\frac{(\langle J_i\rangle_{\text{sim}}-J_{i,\text{expt}})^2}{\sigma^2} \quad (2)$$

where N is the number of J coupling constants measured, $\langle J_i\rangle_{\text{sim}}$ is the ith calculated J coupling constant averaged over all structures in the simulated ensemble, $J_{i,\text{expt}}$ is the ith experimental J coupling constant, and σ^2 is the uncertainty in $\langle J_i \rangle_{\mathrm{sim}}$, which we expect to be dominated by the uncertainty in the Karplus parameters.

ROESY and NOESY Intensities—We used the method of Peter et al.[32] to calculate the spectral density function (which is related to the NOE or ROE peak intensity) from the short NVE simulations. Specifically, we used *ptraj* to analyze the NVE simulations. The *ptraj* output is the normalized correlation function for each pair of hydrogen atoms of the Aβ peptides (the normalization option is convenient as the values of the correlation function, output as decimal numbers with limited field width by *ptraj*, are then all available to high precision, which is especially critical for distant ^{1}H–^{1}H pairs). These are calculated according to

$$C(\tau)=\left\langle \frac{1}{r^6(t)} \right\rangle^{-1} \left\langle \frac{P_2(\cos \chi_{t,t+\tau})}{r^3(t) r^3(t+\tau)} \right\rangle \quad (3a)$$

where $r(t)$ is the vector between each hydrogen atom pair at time t, $\chi_{t,t+\tau}$ is the angle between the $r(t)$ and $r(t+\tau)$ vectors, and P_2 is the second-order Legendre polynomial. We multiply these normalized correlation functions by $\langle r^{-6}(t) \rangle$, averaged over the individual trajectory, to obtain the expression

$$\left\langle \frac{P_2(\cos \chi_{t,t+\tau})}{r^3(t) r^3(t+\tau)} \right\rangle \quad (3b)$$

for each of the 100 independent 20 ns NVE simulations and average these values together and then renormalize the average correlation function by dividing by $\langle r^{-6}(t) \rangle$, averaged over all NVE simulations (equivalent to dividing the entire correlation function by the value at time zero). Each correlation function is then fit over a 5 ns range for τ to a multiexponential form

$$C(\tau) \approx \sum_{i=1}^{N} a_i \mathrm{e}^{-t/\tau_i} \quad (4)$$

where $N = 1, 2, 3,$ or 4 and

$$\sum_{i=1}^{N} a_i = 1$$

using the MATLAB function *lsqcurvefit*, which fits curves in a least-squares sense (The MathWorks, Natick, MA). The value for N is determined using Mallow's C_p statistic[33]

$$\widehat{R}(S)=\widehat{R}_{\mathrm{tr}}(S)+2k\widehat{\sigma}^2 \quad (5)$$

where S is the set of N covariates, R_{tr}, the training error, is the square difference between the correlation values and the fit exponential function, k is the number of regressors $(2N-1)$, and σ^2 is the variance under the full model, which we estimate as 0.001 on the basis of a selected number of spin vectors. The value of N that yields the lowest value for Mallow's C_p statistic is used as the fit, which is a better criterion for model selection than simply the training error, which is biased toward higher values of N.

The resulting time correlation function can be Fourier-transformed

$$J(\omega)=\int_{-\infty}^{\infty}C(\tau)\,\mathrm{e}^{i\omega\tau}\mathrm{d}\tau \quad (6)$$

to yield

$$J(\omega)=\frac{1}{5}\sum_{i=1}^{N}\left(a_i\frac{2\tau_i}{1+\omega^2\tau_i^2}\right) \quad (7)$$

following the convention for Fourier transforms of exponentials. The correlation time constants, τ_i, have a practical upper bound imposed by the rotation of the entire peptide in solution. No vectors may remain correlated at time scales longer than the rotational correlation time of the entire molecule in the slowest dimension. In cases where the average correlation function for the vector between two hydrogen atoms is fit with a time constant of >15 ns (~5% of cases), we replace the long time constant with a time constant of 15 ns, which is a reasonable value for the longest possible rotational time scale, based on the Debye–Stokes theory for calculating the rotational correlation time of a spherical molecule

$$\tau=\frac{4\pi a^3\eta}{3kT} \quad (8)$$

where a is the molecule’s radius, η is the viscosity of the solvent (in this case water), k is the Boltzmann constant, and T is the temperature (287 K). The Aβ42 radius of gyration in our simulated ensemble is 13.0 ± 4.5 Å, which in the approximation of a spherical molecule corresponds to τ values of 0.8–7.4 ns, less than half of our 15 ns limit for the rotational correlation time. This limit on the correlation time has a physical basis, because global tumbling, independent of any local angle or distance fluctuations, causes a complete loss of correlation in $C(\tau)$ as defined in eq 3a. We note that the traditional method for disordered peptides of using the same fast motional correction for all spin pairs was not sufficient to accurately predict the Aβ21–30 ROESY peaks, so we were obliged to calculate correlation times and spectral densities from NVE simulations for both peptides.

For Aβ21–30, we predicted the ROESY spectra from our structural ensemble and dynamical trajectories by calculating the intensity

$$I(t_{\mathrm{mix}})=\mathrm{X}e^{-\Lambda t_{\mathrm{mix}}}\mathrm{X}^{-1}I(0) \quad (9)$$

where **X** and **Λ** are the eigenvectors and eigenvalues, respectively, of the full relaxation matrix, **R**, composed of the diagonal elements

$$\rho_{ii}=\sum_{j=1,\neq i}^{n}\frac{1}{10}K^2\times\left[\frac{3}{2}J_{ij}(2\omega_0)+\frac{9}{4}J_{ij}(\omega_0)+\frac{5}{4}J_{ij}(0)\right] \quad (10a)$$

and off-diagonal elements

$$\sigma_{ij}=\frac{1}{10}K^2\left[\frac{3}{2}J_{ij}(\omega_0)+J_{ij}(0)\right] \quad (10b)$$

where ρ is the direct dipolar relaxation rate and σ is the cross-relaxation rate for all proton pairs, and the constant factor K is defined as

$$K=\frac{\mu_0}{4\pi r_{\mathrm{eff}}^3}\hbar\gamma_{\mathrm{a}}\gamma_{\mathrm{b}} \quad (11a)$$

where μ_0 is the permeability of free space, $\hbar$ is Planck's constant, and γ_a and γ_b are the gyromagnetic ratios for the nuclei of interest. The effective r vector

$$r_{\text{eff}}=\left\langle \frac{1}{r^6(t)} \right\rangle^{-1/6} \quad (11b)$$

is the average of the r^{-6} values, which has then been converted back into units of distance. For Aβ42, we predicted the NOESY spectra from our structural ensemble and dynamical trajectories by calculating the intensity from a different relaxation matrix **R**, composed of diagonal elements

$$\rho_{ii}=\sum_{j=1,\neq i}^{n} \frac{1}{2}K^2 \times \left[3J_{ij}(2\omega_0)+\frac{3}{3}J_{ij}(\omega_0)+\frac{1}{2}J_{ij}(0) \right] \quad (12a)$$

and off-diagonal elements

$$\sigma_{ij}=\frac{1}{2}K^2 \left[3J_{ij}(\omega_0) - \frac{1}{2}J_{ij}(0) \right] \quad (12b)$$

We simulated all hydrogen atoms explicitly for each methyl group and hence calculated all pair correlation functions, including neighboring methylene and methyl group protons, and we ignored water proton coordinates, which is the standard assumption in the NMR experiment.

We solved for the magnetization matrix at the mixing times that were used in both H_2O and ^2H_2O ROESY and NOESY experiments, removing all exchangeable hydrogen atoms to simulate ^2H_2O buffer conditions. For the H_2O experiment, we multiplied all amine intensities by 0.9 to account for the presence of 10% ^2H_2O and removed all basic amine and hydroxyl groups, which would not appear because of broadening caused by exchange with solvent. We then added all of the degenerate intensities from methyl groups and intraresidue spins that are indistinguishable in the NMR spectra.

Finally, the constant relating the $t = 0$ matrix, **M**(0), of unity to the experimental cross-peak intensity scale was approximated by fitting a line of best fit (restrained to cross through the origin) to a plot of known experimental intensities (those that were unambiguously assigned) versus their corresponding calculated intensities. For this plot, the experimental intensity values are an average of the peak intensities measured on either side of the diagonal, which compensates for baseline differences on either side of the diagonal. The slope of this line can then be used as a conversion factor for all calculated intensities. All reported intensities (experimental and simulated) are also normalized by an estimate for the smallest experimentally identifiable peak intensity, so that an intensity of 1 indicates the smallest intensity that should be visible experimentally. After NOE peaks were calculated from the simulated ensemble, the intensities were compared with the experimental intensity and expressed as multiples of the smallest experimentally identifiable peak. This generated a distribution of simulated NOE peak intensities corresponding to every possible atomic contact. For these relative intensities, only peaks with intensities above 1 are expected to be experimentally visible above noise.

Because of the complexity of the noise distribution in the Aβ42 spectra relative to the Aβ21–30 spectra, uncertainties in Aβ42 peak intensity due to noise were estimated from simulation (again expressed relative to the smallest identifiable peak). We used a factor of 0.44 for the H_2O spectrum and 0.50 for the ^2H_2O spectrum, and this noise was assumed to have a normal distribution (although there are other nonrandom noise features in the NMR data). On the basis of simulated NOE intensities for the H_2O spectrum, we were able to estimate the number of NOE peaks below 1 that should statistically be made visible because of noise increasing their intensities above 1, and the number of NOE peaks above 1 that noise should make less than 1, rendering them unobservable. These statistical estimates were made by binning the simulated peaks by intensity. For each bin below 1, a normal distribution with a standard deviation σ of 0.44 was used to determine the probability that random experimental noise would increase the lowest intensity in the bin to a value above 1. This probability is given by

$$P_i = 1 - \Phi\left(\frac{1 - W_i}{\sigma}\right) \quad (13)$$

where Φ is the cumulative distribution function of the normal distribution with a standard deviation of 1, centered at zero, and W_i is the intensity of the weakest NOE peak in bin i. The number of peaks that are increased above 1 due to noise is then given by the sum over all bins below 1

$$N_{\mathrm{fn}} = \sum_{i}^{N_{\mathrm{b}}} P_i n_i \quad (14)$$

where n_i is the number of peaks in bin i. This number N_{fn} is a reasonable estimate for the number of false negatives that should statistically be observed if all true NOE peaks are simulated with the correct intensities. An analogous procedure was performed for the bins above 1 to determine the number of intensities that should statistically drop below 1 because of noise, i.e., the number of false positives expected if all simulated peaks are exactly correct. We used a bin size of $\sigma/2 = 0.22$. The distribution of peak intensities was obtained from the simulated H_2O NOE spectra because experimental NOE intensities below 1 are not visible by definition; however, the intensity distribution should be quite similar to that of the experiment because simulated NOE intensities were scaled to fit experiment and normalized by the lower bound for experimentally visible peak intensity. No peaks in high-noise regions of the experimental spectra (noise from experimental artifacts such as t1 noise) were used in this procedure because these regions do not have the same normally distributed noise with a σ of 0.44.

Residual Dipolar Couplings—We used PALES[34] to calculate residual dipolar coupling (RDC) values by residue for each structure in our simulated Aβ42 ensemble. The program computes the RDC by using steric properties of the molecule to generate an alignment orientation. Then, the angle between the internuclear vector and the external magnetic field was used to calculate the RDC for each snapshot. We calculate the RDC for the backbone amide bond vector for each Aβ42 residue. The average value over the ensemble for each residue is then compared to the Aβ42 experimental data from Yan et al.[27] The simulation values are multiplied by a constant to put them on the same scale as the experimental data, which is determined by experimental peptide concentration.

Ensemble Structural Analysis

Structural analysis of the de novo molecular dynamics simulations of Aβ21–30 and the Aβ42 287 K ensembles was performed using *ptraj*, perl scripts, and MATLAB (The MathWorks) scripts. *ptraj* was used to identify regions of secondary structure in the peptide backbone using the DSSP criteria.[35] We also used *ptraj* to calculate radii of gyration and to identify the most commonly formed hydrogen bonds and salt bridges. We used a cutoff of

3.5 Å between heavy atoms for identifying hydrogen bonds and a 4.0 Å cutoff for salt bridges with a 60° angle cutoff for both.

RESULTS

Chemical Shifts

We compare the Aβ21–30 and Aβ42 experimental and simulated chemical shifts for H_α and H_N (Figure 1), as well as C_α and C_β (Figure 2), for each residue. The experimental carbon chemical shift data for Aβ42 were taken from ref 8; otherwise, the experimental data were generated in the Wemmer lab. Residue specific random coil chemical shifts are subtracted from both experimental and simulated chemical shifts.[36] The agreement is very good between measured and calculated C_α, C_β, H_α, and most H_{SC} chemical shifts, while the calculated amide hydrogen chemical shift values are less good, especially for the Aβ21–30 peptide. This is consistent with previous results showing that SHIFTS and other chemical shift calculators have difficulty predicting amide hydrogen values because of the high sensitivity of chemical shifts to hydrogen bonding geometry.[29,37] In summary, the chemical shift comparisons serve as one indication that our simulations have sampled the correct conformational landscape measured in the experimental ensemble.

In general, the chemical shifts by residue assume values near the average shift for that residue type from the BMRB database:[38] 98% of the hydrogen chemical shift root-meansquare (rms) differences are within one standard deviation, and these trends do not vary greatly across peptide sequence, indicating that the ensemble does not show a strong preference for α-helical or β-sheet conformations. In fact, the rms difference between experimental chemical shifts and average chemical shifts is on the same order of magnitude as the difference between the experimental and simulated chemical shifts. Furthermore, we note that average chemical shifts by residue do not differ greatly from random coil chemical shifts[36] (Figure 1a, inset). This distinction is important when comparing the structural ensembles of Aβ21–30 and Aβ42. We show later that while the tertiary ensemble of the Aβ21–30 fragment peptide is largely a random coil, Aβ42 samples both extended and collapsed states containing significant amounts of β-turn, α-helical, or β-sheet secondary structure, and therefore, its average chemical shift value should not be interpreted as a random coil ensemble. Therefore, the chemical shift data do not discriminate between the two IDP peptide ensembles.

J Couplings

We next consider the comparison of experimental and our simulated $^3J_{H_NH^\alpha}$ scalar coupling constants for Aβ21–30 (Figure 3a) and Aβ42 (Figure 3b), using the established Karplus parameters from ref 31. We initially compare our simulations to the experimental data reported in our previous study on Aβ21–30[10] and those of Yan and co-workers for Aβ42.[27] Overall, the average J coupling constants do not reflect any well-defined secondary structure for either IDP.

We note that our original calculated $^3J_{H^NH^\alpha}$ constants for Aβ21–30 used an incorrect definition of the ϕ angle,[10] and Figure 3a displays the simulation data using the correct definition of the ϕ angle in eq 1. We see that the calculated J coupling constants (blue symbols) do not agree as well with the experimental data (red symbols) for either Aβ42 or Aβ21–30, thereby implying that our ensemble of structures does not sample the same dihedral angle distribution as that of the peptide in the NMR sample and thereby yields average J coupling values that differ significantly from experimental values.

However, we have determined that for Aβ21–30 most of the disagreement is likely due to experimental error in the original studies.[10] We went back to the experimental data and reprocessed it to remove inappropriate truncation of, and apply a cosine window function to, the time domain data in the original processing, yielding values of $^3J_{\mathrm{H^N H^\alpha}}$ significantly larger than the original splittings we reported. To conclusively rule out dangers of interpreting apparent splittings in spectra with either large peak widths or truncation artifacts, we also performed line shape analysis. The (new) directly measured splittings were within ±0.2 Hz of the $^3J_{\mathrm{H^N H^\alpha}}$ obtained by least-squares nonlinear fitting of eq 6.5 in ref 39 to a 1D slice through the peak maxima. These $^3J_{\mathrm{H^N H^\alpha}}$ values are also significantly higher on average than the values we previously reported. These new experimental values are shown in Figures 3a (green symbols) and are shifted upward by ~0.5–1.0 Hz relative to the original data. We also see much improved agreement between simulation and experiment, providing further support of the accuracy of our simulated ensemble for Aβ21–30. We point out that the high coupling constants for the glycine residues (Figure 3a, boxed) are sums of the coupling constants for both glycine H_α atoms, which are measured as a single splitting experimentally.

For Aβ42, there are experimental differences in the $^3J_{\mathrm{H^N H^\alpha}}$ scalar coupling constants reported in two independent NMR studies.[8,27] The earlier 2004 NMR study reported scalar coupling values of 7.0–8.0 Hz for most residues, with some residues exhibiting values as high as 9.0 Hz.[8] The later NMR study[27] determined an average scalar coupling value of ~6.9 Hz across all residues, with a maximum of 8.14 Hz. However, the later NMR study that derived J coupling values from a three-dimensional $\mathrm{H_N H_\alpha}$ experiment did not take into account relaxation effects;[40] if these were included, then the scalar couplings would be shifted upward by ~10%[31] relative to that reported in Figure 3b. In rough terms, accounting for the relaxation effect would provide excellent qualitative agreement between the data of Yan et al.[27] and the earlier NMR study and our simulation of $^3J_{\mathrm{H^N H^\alpha}}$ scalar couplings for Aβ42 that determines an average scalar coupling value of ~7.6 Hz across all residues. We attribute any remainder of J coupling discrepancies to insufficient parametrization of the intrinsic backbone dihedral angle preferences that oversample the extended β-region of the Ramachandran map $\phi = -120°$; $\psi = 145°$) relative to the slightly more collapsed polyproline II (PP II) conformation $\phi = -75°$; $\psi = 145°$). In recent work by Nerenberg and Head-Gordon, using the same force field combination of ff99SB and TIP4P-Ew, we found that agreement of the $^3J_{\mathrm{H^N H^\alpha}}$ coupling constant with experiment is complete for a glycine tripeptide [$\chi^2 = 0.00(37)$].[22] However, in the same study, we found that agreement between experiment and simulation diminishes steeply for alanine and valine tripeptides ($\chi^2 = 1.44$ and 1.01, respectively), and the discrepancy is amplified as peptide length increases, a known shortcoming of current force fields when they are used to simulate unfolded proteins.[22,23]

Residual Dipolar Couplings

Table 1 compares the experimental RDC values[27] of Aβ42 to those calculated from our simulated ensemble using PALES.[34] We compare all available RDCs for 33 residues reported in ref 27 unlike previous work[28] that compared only their simulated RDC data against 19–22 of the 33 experimental RDC values. The rms difference between simulation and experiment is 2.12 Hz; however, this error is dominated by Lys16, which has a very large experimental RDC value compared to those of the other residues. If Lys16 is excluded, our rms difference is lowered to 1.73 Hz, which is on the order of the experimental and simulation error bars (Figure 4). In our opinion, it is difficult to interpret the structural content of the experimental RDCs for an IDP,[25,27,28] and therefore, we primarily present the RDCs to provide further experimental validation of the simulated ensemble.

ROESY and NOESY Peaks

In our original high-field NMR study of Aβ21–30,[10] the 2D ROESY experiments yielded a set of 155 assigned ROESY cross-peaks, of which 28 were weak medium-range ROE interactions. The medium-range ROEs comprised several $i-i+2$ and $i-i+3$ interactions and two extremely weak $i-i+4$ interactions; no longer-range ROE cross-peaks were observed, and no strong patterns of α-helical or β-sheet contacts were evident. In the case of Aβ21–30, all experimental peaks could be assigned. The role of simulation for Aβ21–30 was to ascertain whether the medium-range interactions were occurring simultaneously, as would be typical in structure calculations on folded proteins that assume that weak NOE or ROE cross-peaks correspond to large (~4.5-6.0 Å) upper distance restraints on a single well-defined structural population. In fact, we found that the peptide ensemble involves significant disorder, and hence, it is inappropriate to use the standard structure determination methods.

We determined very good agreement between the simulated and experimentally observed ROE cross-peaks, indicating that our ff99SB–TIP4P-Ew simulations provide a well-validated estimate of the ensemble of structures interrogated by the experiments, allowing them to be used to describe the full structural ensemble diversity. We found that the conformational ensemble of Aβ21–30 was dominated by an unstructured population that lacked any consensus secondary structure or hydrogen bond interactions. The remaining minority population involved an ~14% population of β-turn structure centered at Val 24 and Gly 25, which in turn populated contacts between Asp 23 and Ser 26. The simulations also indicate that the Asp 23–Lys 28 salt bridge, important to the fibril structure,[41] was formed in ~7% of the ensemble. Nonetheless, while Aβ21–30 exhibits some residual structure, it is largely an extended random coil peptide. Experimental and simulation data suggest that the size of the structured populations may increase when the Aβ peptide is extended to larger fragments,[42-45] and that is what we consider next for Aβ42.

Figure 5 shows a region of the experimental ^1H_2O NOESY spectrum of Aβ42 that emphasizes the difficulties in obtaining unambiguous tertiary structure assignments for longer IDPs with spectra that are congested [giving multiple possible peaks assignments (Figure 5a)] and have appreciable variation in noise (Figure 5b). We define an experimental NOE peak as "assignable" on the basis of the definition that there is one dominant short-range contact that must contribute the majority toward the peak intensity, and that its intensity is above experimental noise uncertainties. Therefore, because of the degeneracy of chemical shifts, only 196 of the 705 NOESY peaks seen in the H_2O and ^2H_2O NMR spectra could be assigned from the experimental data alone. Of these assignable peaks, 11 were found to be nonsequential and non-intraresidue, and these peaks had experimental intensities significantly lower than those of many other assigned peaks. Therefore, 509 experimental peaks, 165 of which are $i-i+2$ or greater, have multiple possible assignments for Aβ42, far different than that seen for Aβ21–30.

To assign more of the NOEs, we performed de novo calculations of the NOE intensities using our MD simulation as described in Methods, in which we use the experimentally assigned peaks to place the simulation intensities on the absolute scale of the NMR experiment (see Figure S2 of the Supporting Information). We also determine a scaling factor based on the smallest peak that can be identified clearly above the noise in the experimental spectra and express experimental and simulated intensities as a multiple of this threshold. Only peaks simulated with an intensity above 1 are considered visible above noise. With this scaling factor determined from the ^1H_2O and ^2H_2O experimental data, 176 of the 196 assigned experimental peaks were identified as being visible from the simulation data. None of the 11 long-range assigned peaks were predicted by simulation to be above 1, but 10 of them were predicted to be above 0.1. A factor of 10 in intensity corresponds to a factor of $10^{1/6}$ (1.5) difference in distance; thus, the predicted distances are just outside the range of giving a detectable peak. We note that 10 of these peaks involved $i-i+2$ and $i-i+3$ contacts, with only one peak that involved an $i-i+4$ contact.

The remaining 509 unassigned experimental NOEs were compared to the simulation data by summing the calculated NOE intensities for all proton pairs that had the same chemical shift of the experimentally observed peak. Of these 509 unassigned peaks, the simulation predicted 355 peaks, of which 54 had no sequential or intraresidue contact. Only 223 of these 355 contacts were dominated by a single atomic contact included, while 132 of these simulated peaks had two or more contacts contributing with similar intensities. Therefore, there were 174 false negatives (not predicted to be visible), 122 of which could not result from an intraresidue or sequential contact. In addition to the 531 simulated peaks that were consistent with experimentally observed peaks, there were 383 false positive peaks (calculated from simulations that were not seen in experiment). These data are summarized in Table 2.

In total, the simulations yielded 174 false negatives and 383 false positives when compared to the NMR experiment. However, unlike our previous experiment on more concentrated samples of Aβ21–30, the noise across the NOESY spectra for Aβ42 was quite variable. We estimate that the average experimental uncertainty in the intensities due to noise is ±0.44 of the threshold for the H_2O spectrum and ±0.50 for the ^2H_2O spectrum. However, there are regions where noise moves far outside these experimental ranges because of t1 noise, modulations from the water signal, and/or dense regions of overlapping large peaks that make peak intensities difficult to determine (see Figure 5). This required a more painstaking analysis of different regions of the spectra to determine our confidence in what are genuine false positives and false negatives in the NOESY assignments. For example, 82 false positive peaks were in regions of the spectra well above the baseline intensity uncertainties, which would prevent these 82 peaks from being observed experimentally. Of the false negatives, 62 of these peaks were observed in the high-noise regions of the experimental spectra, which could cause a weak peak to appear stronger.

If peaks that could not be seen due to high noise were disregarded, there were 301 remaining false positives, 115 of which contain only residue $i-i+2$ or longer contacts, and 112 remaining false negatives, of which 75 are long-ranged. However, for many hydrogen–hydrogen pairs with NOEs observed but not predicted (false negatives), NOEs between other proton pairs within the same residue pair were not observed but were predicted (false positives). Thus, although the simulations sometimes miss the correct interproton contacts, it does largely succeed in predicting correct residue contacts. Considering predictions between the proper residue pairs as correct removes 223 false positives and 66 false negatives (Table 2). Finally, the 78 remaining false positives were analyzed to see whether their intensities were dominated by a specific set of contacts: 16 of these peaks, although predicted to be above noise level, did not contain any dominant atomic contact of significant intensity but instead were composed of many atomic contacts whose individual intensities were below half of the noise level. Considering that multiple cross-peaks would probably be spread over a range of shifts wider than that used in constructing the simulated spectrum, these are not serious experimental disagreements.

The remaining 46 false negatives (of 705 observed peaks) and 62 false positives (of 914 simulated peaks) are weak as estimated by the experimental intensity uncertainty inherent in the H_2O (±0.44) and ^2H_2O spectra (±0.50), although we report them in the Supporting Information. If the noise is assumed to be normally distributed, on the basis of our distribution of experimental intensities, we would expect to see ~69 false negatives, compared to the 46 that we determined. Given the experimental variation in noise, our level of false negatives is not statistically significant. Similarly, if the noise is assumed to be normally distributed, on the basis of our distribution of simulated peak intensities, we would expect to see ~29 false positive peaks, whereas we see 60. Even so, a vast majority of the false positives are short-range and do not grossly influence the backbone structural ensemble. These data are also summarized in Table 2.

Peptide Structural Ensemble

Given the overall good agreement of the calculated NMR observables from de novo molecular dynamics with various NMR measurements, we proceed to a structural analysis of the simulated ensemble. Figure 6 shows the radius of gyration (R_g) distribution of Aβ42, which emphasizes that the peptide adopts conformations that vary extensively in the degree of compactness. It is interesting to compare this to a random coil polymer, in which the original Flory model[46] states that R_g scales with the number of residues, n

$$R_g = R_0 n^{\nu} \quad (15)$$

where R_0 is an estimated persistence length and ν depends on the quality of the solvent. In a good (denaturing) solvent $\nu \sim 0.598$, and assuming a persistence length (R_0) of 1.33,[47] a 42-amino acid peptide's average radius of gyration, $\langle R_g \rangle$, would be estimated to be 12.4 ± 1.0 Å, consistent with what we find for the Aβ42 peptide in which $\langle R_g \rangle$ = 13.0 ± 4.5 Å. However, the large standard deviation for $\langle R_g \rangle$ and the nature of its distribution (Figure 6) emphasize that the Aβ42 peptide samples many conformations that are even more compact or more extended than expected from a good solvent model or estimates of the persistence lengths of denatured proteins. For example, if we assume a larger persistence length of 2.1,[48] we obtain an $\langle R_g \rangle$ value of 19.6 Å in a good solvent or an $\langle R_g \rangle$ of 7.2 Å in a poor (waterlike) solvent ($\nu \sim 0.33$), which better correlates with Aβ42's skewed R_g distribution with a peak around 9–10 Å but with a long tail that samples conformations with an R_g as large as ~30 Å.

We also find that the more compact portion of the Aβ42 ensemble (~90% of the ensemble) is dominated by conformations with some type of secondary structure, hydrogen bonding, or ion pairing structural features. Figure 7 shows the probability of observing a given DSSP secondary structure category for each amino acid; DSSP defines secondary structure classes by distinct hydrogen bonding patterns and geometry (primarily dihedral angles), with helices and β-sheets requiring the cooperative organization of repeats based on localized turns and bridges.[35] We find that only 0.2% of the Aβ42 ensemble is completely lacking any identifiable secondary structure! This is far different than our previous study of Aβ21–30 in which ~65% of the structural ensemble had no identifiable stabilizing secondary structure or consensus hydrogen bonds.[26]

To better characterize the Aβ42 peptide conformations that contain these turns, we focused on six pairs of residues that define the following turns. The first turn is defined as residues 7 and 8 (Turn7–8), the second turn as residues 14 and 15 (Turn14–15), the third turn as residues 18 and 19 (Turn18–19), the fourth turn as residues 24 and 25 (Turn24–25), the fifth turn as residues 26 and 27 (Turn26–27), and the final turn as residues 34 and 35 (Turn34–35). We also combine the α-helix structure and 3_{10}-helix classifications into one "helix" classification in our secondary structure analysis because the same sequence region has the same tendency to form either helical structure. Residues 14–19 often adopt helical structure, as do residues 22–28, while residues 30–35 do so to a lesser extent. However, for residues 14–19 and 22–28, in 26% and 10% of the ensemble, respectively, only part of the helix is formed while the other portion of the helical region forms one of the turns mentioned above. For this reason, we break these regions into five pairs of residues that often form helical structure: residues 14 and 15 (Helix14–15), residues 17 and 18 (Helix17–18), residues 23 and 24 (Helix23–24), residues 26 and 27 (Helix26–27), and residues 32 and 33 (Helix32–33). These residues simply label the central regions of the helix, because DSSP requires that helices (as well as turns and sheets) be supported by hydrogen bonds involving additional residues. Finally, antiparallel β-sheet secondary structure is also observed in the ensemble. Antiparallel β-sheet structure is observed for residues 3–6 (Beta3–6), residues 10–13 (Beta10–13), residues 16–21 (Beta16–21), and finally residues 29–36 (Beta29–36). We do not display parallel β-sheet structure because this structure was seen at only very low levels, and when the particular conformations were examined more closely, they were found not to be true parallel β-sheets but rather 180° turns.

Table 3 summarizes the percentage of time each of the above secondary structure categories appears in the Aβ42 conformational ensemble, the percentage of time each occurs in isolation with no other secondary structure, and the frequency with which pairs of secondary structure elements form simultaneously. It is seen that Turn7–8 and Turn34–35 are the turns most commonly formed, and each is present in almost half of the simulation ensemble; they often form simultaneously with each other (34% of the ensemble). For the minority of consensus secondary structure found in the conformational ensemble of Aβ21–30, the dominant feature was ~14% of the ensemble forming a classic β-turn structure centered at Val 24 and Gly 25 bringing together Asp 23 and Ser 26.[10]

The Aβ42 ensemble is dominated by secondary structure pairings with Turn7–8 and Turn34–35 (Figure 8a,b), impossible for Aβ21–30: Turn14–15, Helix17–18, and Turn24–25 are often present in conjunction with either of these turns in ~25% of the ensemble, Helix14–15 and Turn18–19 are most often paired with these turns in ~17% of the ensemble, and likewise Helix23–24 and Helix26–27 are most frequently present in conjunction with either Turn7–8 or Turn34–35 in ~9% of the ensemble. Table 2 also shows that Helix17–18 occurs with a high frequency (32% of the ensemble) compared to the other helical segments but never occurs in isolation, while Helix32–33 is the least frequent of the secondary structure categories. Even though the helical conformational populations are diverse, it is clear that helical conformations are more populated in the N-terminal region of the Aβ42 peptide (Figure 9a,b). We emphasize that while we have quantified the percentages of the ensemble that involve secondary structure pairings, most often these pairs are occurring with a third, fourth, or sometimes a greater number of additional secondary structure elements, resulting in a combinatorial explosion in the structural diversity of metastable conformations.

The most unusual secondary structure feature for Aβ42 involves Turn26–27. While it simultaneously pairs with either Turn7–8 or Turn34–35 in ~14% of the ensemble, 12% of the time the formation of Turn26–27 occurs simultaneously with β-sheet structure involving Beta16–21 and Beta29–36, and this β-sheet structure precludes the formation of many of the other secondary structures such as Turn14–15, Turn18–19, Turn24–25, Turn34–35, Helix14–15, Helix17–18, Helix23–24, Helix32–33, and, of course, Helix26–27, which are so common in the rest of the ensemble. In fact, 8% of the time Turn26–27 occurs in isolation [compared to 0–3% for all other turns or helices (Table 3)], i.e., with all other secondary structure categories completely absent, and when this happens, the β-sheet is almost always present (Figure 10a). This is unlike the case in which the N-terminal antiparallel β-sheet (Beta3–6 and Beta10–13) nucleates around Turn7–8 (9% of the ensemble), because this localized structure always forms simultaneously with other secondary structure categories in the latter two-thirds of the sequence (Figure 10b). Thus, while Turn26–27 is not the most dominant feature of secondary structure, present in only 28% of the ensemble, it is one of the most important features because it can have a much longer-range effect on the peptide's structure than all other secondary structure categories that are more structurally localized. There is still a significant percentage (9%) of the population in which the Beta16–21 and

Beta29–36 pair forms without Turn26–27 present (Figure 10c), so that in total this *β*-sheet occurs in 21% of the ensemble. In 5% of the ensemble, Beta16–21 forms a *β*-sheet with some other part of the peptide, and in 4% of the ensemble, Beta29–36 forms a *β*-sheet with another region besides Beta16–21, although in both cases the pairing partner region is highly promiscuous.

Another interesting structural feature of the ensemble is the presence of particular hydrogen bonding partners, and Table 4 emphasizes that most of these are stabilizing the secondary structure categories required by the DSSP definitions. Stabilizing salt bridges are also found in the A*β*42 ensemble, in which Arg 5 often participates in various salt bridges: the largest occupancy (16% of the ensemble) is between the Arg 5 and Glu 3 side chains, and Arg 5 also forms frequent salt bridges with Asp 1, Glu 11, and Glu 22. The Asp 23–Lys 28 salt bridge, found in the A*β*42 fibril conformation as well as 7% of the A*β*21–30 ensemble, and its competitor, the Glu 22–Lys 28 salt bridge, are found in 3 and 4% of the ensemble, respectively, and not surprisingly, they coform with helix and turns that they encapsulate: Turn24–25, Turn26–27, Helix23–24, and Helix26–27. In only 1% of the population does a salt bridge stabilize the Beta16–19 and Beta29–36 *β*-sheet (Figure 10d), which is thought to be necessary for stabilizing the A*β*42 monomer in the greater amyloid fibril assembly, although we observe it only for the competitor Glu 22–Lys 28 pairing.[41,49]

Finally, Figure 11 provides a contact map of strong experimental NOE intensities for A*β*42 that are dominated by a single contact. We note that these dominant NOEs correlate well both with the various secondary structures (including *β*-sheets) and the presence of NOEs that “compact” structure (i.e., those that involve multiple secondary structure categories occurring simultaneously) that are described in Figures 6-10. Moreover, they are highly consistent with the hydrogen bonds that occur with high frequency in Table 4 and lend strong support to the observed structural subpopulations.

DISCUSSION AND CONCLUSION

Our NMR–de novo MD study of the A*β*21–30 and A*β*42 conformational ensembles reveals a broad range of intrinsically disordered peptide structure. This range of structural complexity can be usefully described with techniques that combine the collection and analysis of both experimental NMR and simulation data using more modern force fields. Using a combination of NOESY cross-peaks, which reveal tertiary interactions via definitive intramolecular contacts, and MD simulations that provide the structural context for such contacts in an equilibrium ensemble under conditions that match those of the experiment, we achieve a picture of two very different intrinsically disordered peptides. While the A*β*21–30 peptide conforms to a rather homogeneous ensemble consistent with an extended random coil, A*β*42 exhibits traits of an extremely heterogeneous ensemble of peptide conformations that contain a diversity of localized as well as long-range tertiary structure.

Of interest here is the fact that functional but intrinsically disordered peptides and proteins are typically classified as having a dominant population corresponding to a particular “degree of disorder”: collapsed semiordered ensembles (radius of gyration similar to that of a globular protein of the same sequence length, typically containing well-formed secondary structure but little organized tertiary structure), collapsed disordered ensembles (radius of gyration typical of a globular protein, containing little well-formed secondary or tertiary structure), or an extended disordered ensemble (much larger radius of gyration than a globular protein and dominated by random coil). The A*β*42 peptide does not fall into any of these classifications exclusively; rather, NMR and MD show that its structural populations span the full range of classifications. For example, unlike molten globules (or ordered proteins), the A*β*42 hydrogen chemical shifts are not highly dispersed; ~98% of the chemical shifts assume values within a standard deviation of the average shift for each residue type along the Aβ42 sequence. The *J* coupling values also provide no evidence of secondary structure "blocks" at different points in the peptide sequence of Aβ42. Because Aβ42 samples so many distinct conformations, experimental observables that are based on simple ensemble averages like chemical shifts and spin–spin couplings yield values for these observables that are close to random coil values. While this might be interpreted as a signature of an extended IDP, in fact the homogeneity of the chemical shifts and *J* coupling are a result of averaging over many heterogeneous conformations involving compact, structured conformations, in addition to a significant population (~10%) of more extended conformers that span the R_g range from 9 to 30 Å. The NOESY experiment also confirms that these compact conformations must persist on sufficiently long time scales for the cross relaxation between protons to be detected, and thus, these interactions must be stabilized by (likely shallow) free energy minima. While the MD simulations show that many medium- to long-range NOEs in the Aβ42 NOESY spectra are reporting on semiorder in the collapsed conformations, we also observe numerous instances in which a single peak consists of many different contacts arising from different conformations with completely distinct tertiary structures and degrees of compactness. By contrast, Aβ21–30 can be more easily classified as an extended random coil.

The many strong NOESY cross-peaks observed experimentally for Aβ42, which we are able explain with specific attributes of the peptide populations from our simulated ensemble, allow us to corroborate or question specific structural features proposed in previous work. The first NMR studies of the monomeric structural ensemble of Aβ42 (and Aβ40) by Hou et al. were aimed at changes that occurred upon oxidation of the Met 35 side chain, a chemical change associated with impeded fibrillization of the peptide.[8] On the basis of ^{1}H, ^{15}N, and ^{13}C NMR chemical shifts and backbone NOEs, that study found an absence of any well-defined secondary or tertiary structures, interpreting from the NMR data that the ensemble was well-classified as random, extended chains. Although our results show that Aβ42 is not really an extended random coil, that was a reasonable conclusion based on the more limited analysis of NMR data alone, because simple averaging over an ensemble of a highly diverse set of secondary and tertiary structures, a majority of which are collapsed, can generate average NMR observables consistent with extended random coil signatures. However, on the basis of NOE and chemical shift data, Hou et al. identified backbone C_α and C_β chemical shifts consistent with β-strand structure in two hydrophobic regions (Leu 17–Ala 21 and Ile 31–Val 36), as well as turn structures at two largely hydrophilic regions (Asp 7–Glu 11 and Phe 20–Ser 26). Those results are consistent with our study in which we found significant β-sheets (Beta16–21 and Beta29–36), as well as turns (Turn7–8, Turn24–25, and Turn26–27) largely in the same region of the sequence. What our study offers is a far richer picture of how these secondary structure elements organize together to create a diverse set of collapsed to extended conformations that contribute to the broad ensemble average.

A study by Sgouarkis and co-workers consisted of an MD simulation using the OPLS force field and the TIP3P water model, which they validated against measured scalar $^3J_{\mathrm{HNH}\alpha}$ data.[26] Although the agreement with the experimental *J* coupling data was not particularly good, they nonetheless characterized the conformational ensemble of Aβ42 (and Aβ40) that pointed to a unique feature of the longer peptide forming a β-hairpin at Met 35 and Val 36, bringing together short antiparallel β-strands at residues Ile 31–Leu 34 and Gly 38–Ile 41. This prediction of a large population of β-sheet structure in this region clearly contradicts our study and the earlier NMR results of Hou et al. because NOEs arising from these structures are not observed. More recently, Sgourakis and co-workers performed a second MD study on Aβ42,[28] this time using a combination of the ff99SB force field and the TIP4P-Ew water model that we successfully used on Aβ21–30[10] and used in this study. They analyzed 11570 Aβ conformations using a clustering algorithm that yielded thousands of small clusters, with the six largest clusters having populations in the range of 2–6% of the ensemble.[28] This analysis does not provide much structural insight, and we found classifications based on energetically stabilizing features such as hydrogen bonds and secondary structure to be more informative. Again they find central importance in the short *β*-strand at residues 38–40, which they argue may act as a conformational switch by forming alternative interactions with other *β*-strands along the sequence, in disagreement with this study and that of Hou and co-workers.

In summary, the full A*β*42 peptide and the shorter A*β*21–30 fragment we studied previously[10] represent two fundamentally different types of IDPs. Experimentally, while the average chemical shift and scalar coupling values suggest that they both are random coil polymers, it is only true for A*β*21–30, an outcome supported by a de novo molecular simulation ensemble that predicts its small number of 28 nonsequential NOESY cross-peaks.[10] By contrast, the NMR observables for A*β*42 represent an average over a rich and diverse set of tertiary structures, supported by a larger number of 179 nonsequential NOESY cross-peaks that are reasonably described by our MD ensemble. While the A*β*21–30 structural ensemble at 283 K is primarily extended, with ~65% of the structures containing no secondary structure or hydrogen bond pairing for any residue, the A*β*42 structural ensemble at 287 K has <1% of the population in which there is no secondary structure. In fact, key structures such as the *β*-sheet involving Beta16–21 and Beta29–36 could even be crucial to oligomerization of fibril nucleation, although this is currently speculation. A*β*21–30 has the same primary sequence as a key middle segment of the A*β*42 peptide; however, the additional interactions possible in the longer peptide shift the equilibrium to other stabilizing conformations dominated by a different set of residues, with the exception of ~5% of the population that contains the Asp 23–Lys 28 salt bridge that is observed in both peptides. The full A*β*42 sequence may be necessary to interpret any physiological behavior of the peptide, because it is only the complete peptide that adopts a diverse but structured conformational ensemble.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the National Energy Research Scientific Computing Center (NERSC) for computational resources. We especially thank NERSC Director Kathy Yelick for 1.5 million central processing unit hours from her Director's fund, which was vital for the completion of this study. All of us thank Dr. Robert Tycko for the A*β*42 peptide and Youcef Ouadah and Richard Zhu for help with the experiments.

Funding

T.H.-G. thanks the National Institutes of Health (NIH) for support of our research. K.A.B. is grateful for NIH Molecular Biophysics Training Grant T32 GM08295.

References

1. Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, et al. Intrinsically disordered protein. J Mol Graphics Modell. 2001; 19:26–59.
2. Dunker AK, Silman I, Uversky VN, Sussman JL. Function and structure of inherently disordered proteins. Curr Opin Struct Biol. 2008; 18:756–764. [PubMed: 18952168]
3. Uversky V, Gillespie J, Fink A. Why are "natively unfolded" proteins unstructured under physiologic conditions? Proteins. 2000; 41:415–427. [PubMed: 11025552]
4. Uversky VN, Dunker AK. Understanding protein non-folding. Biochim Biophys Acta. 2010; 1804:1231–1264. [PubMed: 20117254]

5. Romero P, Obradovic Z, Li X, Garner EC, Brown CJ, Dunker AK. Sequence complexity of disordered protein. Proteins. 2001; 42:38–48. [PubMed: 11093259]
6. Dunker AK, Obradovic Z. The protein trinity: Linking function and disorder. Nat Biotechnol. 2001; 19:805–806. [PubMed: 11533628]
7. Lu JX, Yau WM, Tycko R. Evidence from solid-state NMR for nonhelical conformations in the transmembrane domain of the amyloid precursor protein. Biophys J. 2011; 100:711–719. [PubMed: 21281586]
8. Hou L, Shao H, Zhang Y, Li H, Menon NK, Neuhaus EB, et al. Solution NMR studies of the A*β*(1–40) and A*β*(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]
9. Uversky VN. Intrinsic disorder in proteins associated with neurodegenerative diseases. Front Biosci. 2009; 14:5188–5238. [PubMed: 19482612]
10. Fawzi NL, Phillips AH, Ruscio JZ, Doucleff M, Wemmer DE, Head-Gordon T. Structure and dynamics of the A*β*(21–30) peptide from the interplay of NMR experiments and molecular simulations. J Am Chem Soc. 2008; 130:6145–6158. [PubMed: 18412346]
11. Teplow D. Preparation of Amyloid *β*-Protein for Structural and Functional Studies. Methods Enzymol. 2006; 413:20–33. [PubMed: 17046389]
12. Keller, R. Institute of Molecular Biology. Swiss Federal Institute of Technology; Zurich: 2004. The Computer Aided Resonance Assignment Tutorial; p. 159
13. Bax A, Davis DG. Practical Aspects of Two-Dimensional Transverse NOE Spectroscopy. J Magn Reson. 1985; 63:207–213.
14. Liu M, Mao X, Ye C, Huang H, Nicholson JK, Lindon JC. Improved WATERGATE pulse sequences for solvent suppression in NMR spectroscopy. J Magn Reson. 1998; 132:125–129.
15. Lippens G, Dhalluin C, Wieruszeski JM. Use of a Water Flip-Back Pulse in the Homonuclear NOESY Experiment. J Biomol NMR. 1995; 5:327–331. [PubMed: 22911506]
16. Piotto M, Saudek V, Sklenár V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. J Biomol NMR. 1992; 2:661–665. [PubMed: 1490109]
17. Sklenar V, Piotto M, Leppik R, Suadek V. Gradient-tailored water suppression for ^{1}H-^{15}N HSQC experiments optimized to retain full sensitivity. J Magn Reson. 1993; 102:241–245.
18. 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. 2006; 65:712–725. [PubMed: 16981200]
19. Horn HW, Swope WC, Pitera JW, Madura JD, Dick TJ, Hura GL, et al. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. J Chem Phys. 2004; 120:9665–9678. [PubMed: 15267980]
20. Best RB, Buchete NV, Hummer G. Are current molecular dynamics force fields too helical? Biophys J. 2008; 95:L07–L09. [PubMed: 18456823]
21. Best RB, Hummer G. Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. J Phys Chem B. 2009; 113:9004–9015. [PubMed: 19514729]
22. Nerenberg PS, Head-Gordon T. Optimizing Protein-Solvent Force Fields to Reproduce Intrinsic Conformational Preferences of Model Peptides. J Chem Theory Comput. 2011; 7:1220–1230.
23. Wickstrom L, Okur A, Simmerling C. Evaluating the Performance of the ff99SB Force Field Based on NMR Scalar Coupling Data. Biophys J. 2009; 97:853–856. [PubMed: 19651043]
24. Ruscio J, Fawzi N, Head-Gordon T. How hot? Systematic convergence of the replica exchange method using multiple reservoirs. J Comput Chem. 2010; 31:620–627. [PubMed: 19554556]
25. Yan Y, Wang C. A*β*42 is more rigid than A*β*40 at the C terminus: Implications for A*β* aggregation and toxicity. J Mol Biol. 2006; 364:853–862. [PubMed: 17046788]
26. Sgourakis NG, Yan Y, McCallum SA, Wang C, Garcia AE. The Alzheimer's peptides A*β*40 and 42 adopt distinct conformations in water: A combined MD/NMR study. J Mol Biol. 2007; 368:1448–1457. [PubMed: 17397862]
27. Yan Y, McCallum SA, Wang C. M35 oxidation induces A*β*40-like structural and dynamical changes in A*β*42. J Am Chem Soc. 2008; 130:5394–5395. [PubMed: 18376837]

28. Sgourakis NG, Merced-Serrano M, Boutsidis C, Drineas P, Du Z, Wang C, et al. Atomic-Level Characterization of the Ensemble of the Aβ(1–42) Monomer in Water Using Unbiased Molecular Dynamics Simulations and Spectral Algorithms. J Mol Biol. 2010; 405:570–583. [PubMed: 21056574]
29. Xu XP, Case DA. Automated prediction of ^{15}N, ^{13}C$_{\alpha}$, ^{13}C$_{\beta}$ and ^{13}C' chemical shifts in proteins using a density functional database. J Biomol NMR. 2001; 21:321–333. [PubMed: 11824752]
30. Karplus M, Grant DM. A Criterion for Orbital Hybridization and Charge Distribution in Chemical Bonds. Proc Natl Acad Sci U S A. 1959; 45:1269–1273. [PubMed: 16590503]
31. Vuister GW, Bax A. Quantitative J Correlation: A New Approach For Measuring Homonuclear 3-Bond $J_{H^NH^{\alpha}}$ Coupling Constants in N-15-Enriched Proteins. J Am Chem Soc. 1993; 115:7772–7777.
32. Peter C, Daura X, van Gunsteren WF. Calculation of NMR-relaxation parameters for flexible molecules from molecular dynamics simulations. J Biomol NMR. 2001; 20:297–310. [PubMed: 11563554]
33. Mallows CL. Some Comments on Cp. Technometrics. 1973; 15:661–675.
34. Zweckstetter M, Bax A. Prediction of sterically induced alignment in a dilute liquid crystalline phase: Aid to protein structure determination by NMR. J Am Chem Soc. 2000; 122:3791–3792.
35. Kabsch W, Sander C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. Biopolymers. 1983; 22:2577–2637. [PubMed: 6667333]
36. Merutka G, Dyson HJ, Wright PE. ‘Random coil’ ^{1}H chemical shifts obtained as a function of temperature and trifluoroethanol concentration for the peptide series GGXGG. J Biomol NMR. 1995; 5:14–24. [PubMed: 7881270]
37. Neal S, Nip AM, Zhang H, Wishart DS. Rapid and accurate calculation of protein ^{1}H, ^{13}C and ^{15}N chemical shifts. J Biomol NMR. 2003; 26:215–240. [PubMed: 12766419]
38. Ulrich EL, Akutsu H, Doreleijers JF, Harano Y, Ioannidis YE, Lin J, et al. BioMagResBank. Nucleic Acids Res. 2008; 36:D402–D408. [PubMed: 17984079]
39. Cavanagh, J.; Fairbrother, WJ.; Palmer, AG., III; Skelton, NJ.; Rance, M. Protein NMR Spectroscopy: Principles and Practice. 2. Elsevier Academic Press; London: 2007.
40. Wang C. private communication. 2011
41. Tycko R. Solid-state NMR studies of amyloid fibril structure. Annu Rev Phys Chem. 2011; 62:279–299. [PubMed: 21219138]
42. Barrow CJ, Yasuda A, Kenny PT, Zagorski MG. Solution conformations and aggregational properties of synthetic amyloid β-peptides of Alzheimer’s disease. Analysis of circular dichroism spectra. J Mol Biol. 1992; 225:1075–1093. [PubMed: 1613791]
43. 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]
44. Tarus B, Straub JE, Thirumalai D. Dynamics of Asp23-Lys28 salt-bridge formation in Aβ(10–35) monomers. J Am Chem Soc. 2006; 128:16159–16168. [PubMed: 17165769]
45. Kirschner DA, Inouye H, Duffy LK, Sinclair A, Lind M, Selkoe DJ. Synthetic peptide homologous to β protein from Alzheimer’s disease forms amyloid-like fibrils in vitro. Proc Natl Acad Sci U S A. 1987; 84:6953–6957. [PubMed: 3477820]
46. Flory, PJ. Statistical Mechanics of Chain Molecules. Wiley; New York: 1969.
47. Kohn JE, Millett IS, Jacob J, Zagrovic B, Dillon TM, Cingel N, et al. Random-coil behavior and the dimensions of chemically unfolded proteins. Proc Natl Acad Sci U S A. 2004; 101:12491–12496. [PubMed: 15314214]
48. Fitzkee NC, Rose GD. Reassessing random-coil statistics in unfolded proteins. Proc Natl Acad Sci U S A. 2004; 101:12497–12502. [PubMed: 15314216]
49. Luhrs T, Ritter C, Adrian M, Riek-Loher D, Bohrmann B, Dobeli H, et al. 3D structure of Alzheimer’s amyloid-β(1–42) fibrils. Proc Natl Acad Sci U S A. 2005; 102:17342–17347. [PubMed: 16293696]

ABBREVIATIONS

NMR	nuclear magnetic resonance
A*β*42	amyloid-*β* 1–42
IDPs	intrinsically disordered peptides

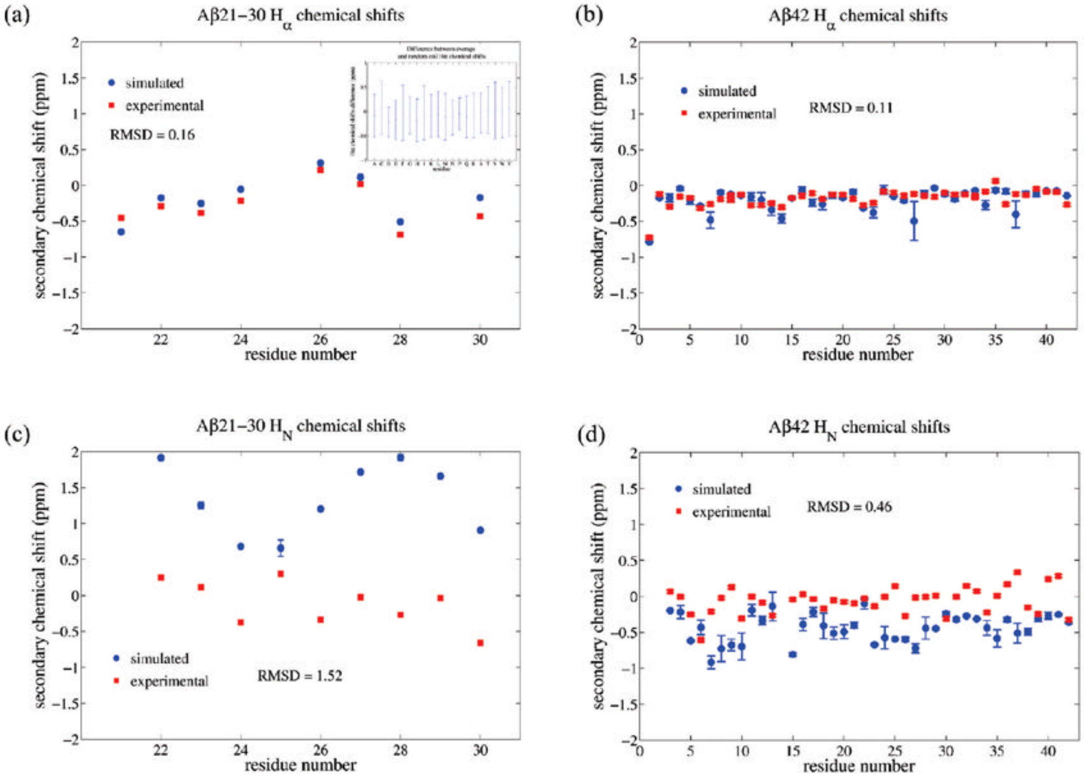


Figure 1.

Experimental and calculated proton secondary chemical shifts by residue: (a) A*β*21–30 H_α chemical shifts, (b) A*β*42 H_α chemical shifts, (c) A*β*21–30 H_N chemical shifts, and (d) A*β*42 H_N chemical shifts. Red squares represent experimental data, while blue circles represent the data calculated from simulation. Random coil residue specific values are subtracted from both experimental and simulation values. The experimental data for A*β*21–30 were taken from ref 10. The inset in panel a shows the difference between the average chemical shift for each residue and the random coil H_α chemical shift from ref 36.

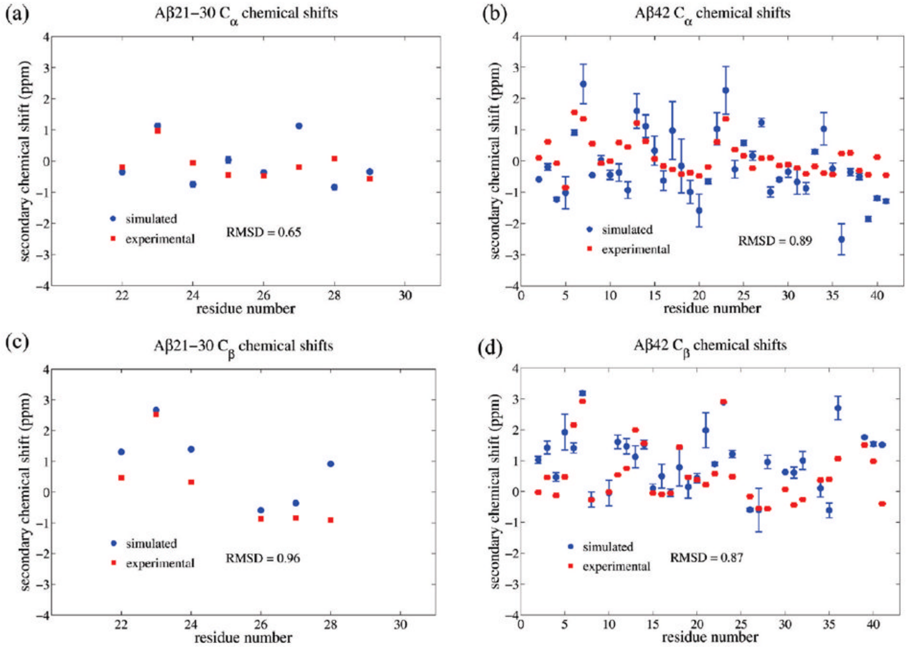
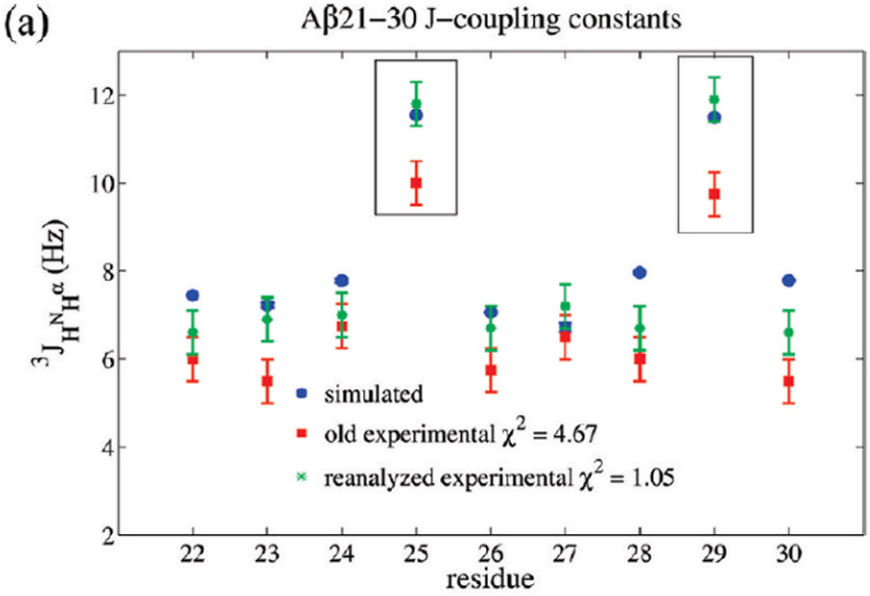


Figure 2.

Experimental and calculated carbon secondary chemical shifts by residue: (a) Aβ21–30 C_α chemical shifts, (b) Aβ42 C_α chemical shifts, (c) Aβ21–30 C_β chemical shifts, and (d) Aβ42 C_β chemical shifts. Red squares represent experimental data, while blue circles represent the data calculated from simulation. Random coil residue specific values are subtracted from both experimental and simulation values. The experimental data for Aβ42 were taken from ref 8, while the Aβ21–30 data were taken from ref 10.



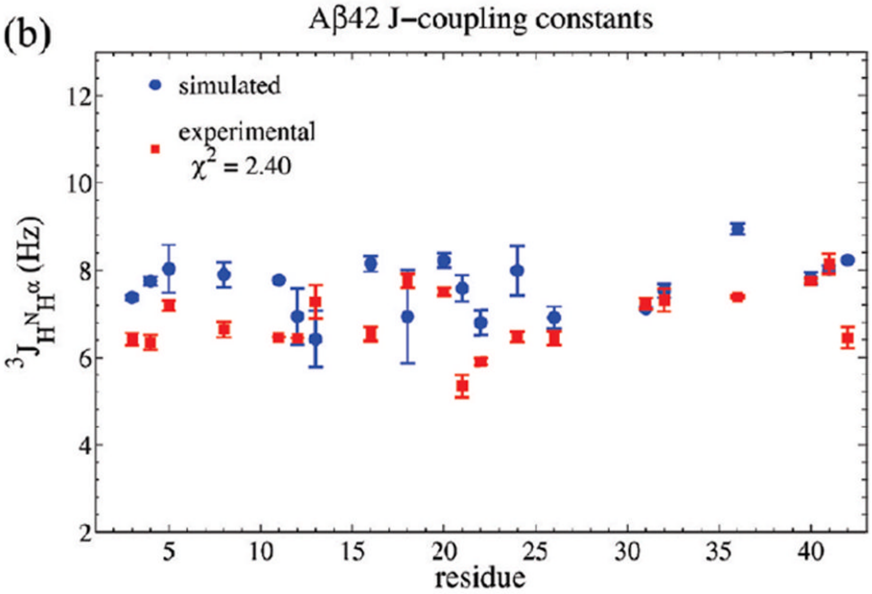


Figure 3.

J coupling constants for backbone amides of (a) A*β*21–30 and (b) A*β*42. The red symbols depict the A*β*42 experimental data from ref 27 and the A*β*21–30 data from ref 10. The green stars depict the reprocessed experimental data for A*β*21–30. The blue circles depict the data calculated from the simulation ensemble using the Karplus parameters of Vuister and Bax.[31] Glycine residues are boxed to emphasize that they can only be measured as the sum of two coupling constants and therefore have higher values than the other residues. Simulation uncertainty bars represent rms differences between two independent simulations and the average. The data of Yan et al.[27] have not accounted for a relaxation correction that makes *J* couplings determined from a H_NH_α three-dimensional experiment consistently lower than those from COSY splittings[8] by a small amount (maybe as much as ~10%[31]); if this relaxation effect is included, then the qualitative agreement between experiment and simulation is excellent.

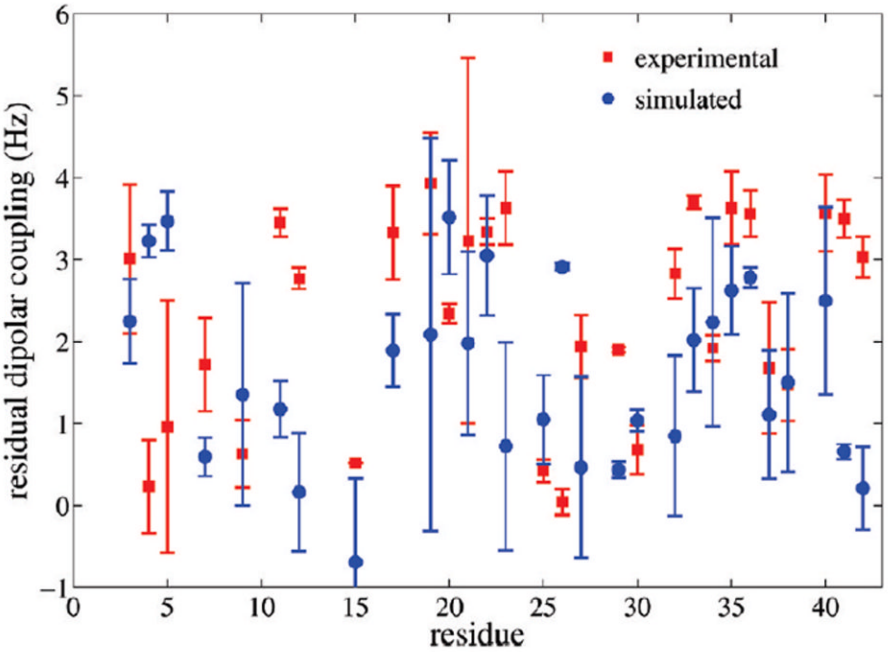


Figure 4.

Experimental vs calculated residual dipolar couplings for Aβ42. The red symbols depict experimental data for Aβ42 from ref 27. The blue circles depict the data calculated from the simulation ensemble.

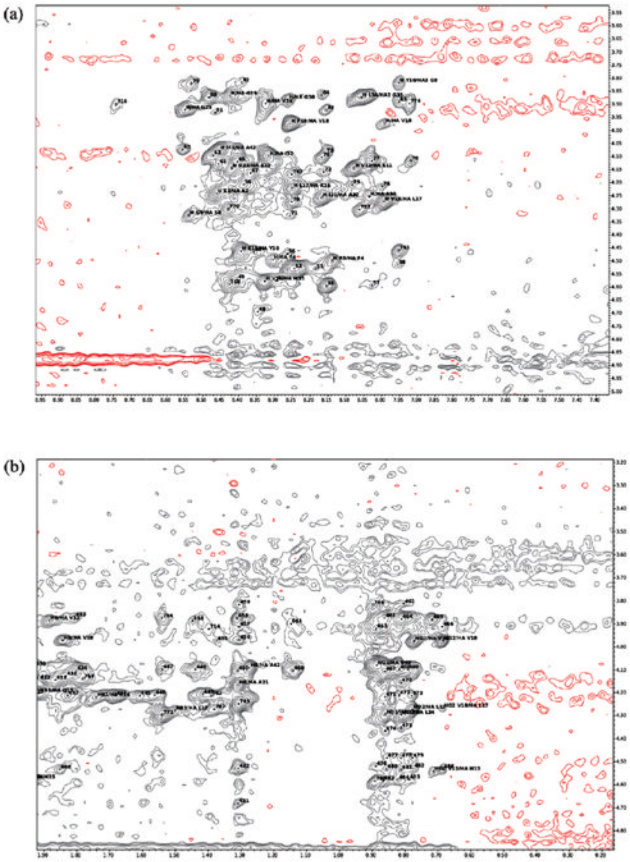


Figure 5.
Experimental NOESY spectra. (a) Region of the spectrum with large chemical shift degeneracy (47 H chemical shifts along the *x* axis and 52 H chemical shifts along the *y* axis). (b) Region of the spectra with high t1 noise and a large background signal arising from the water peak.

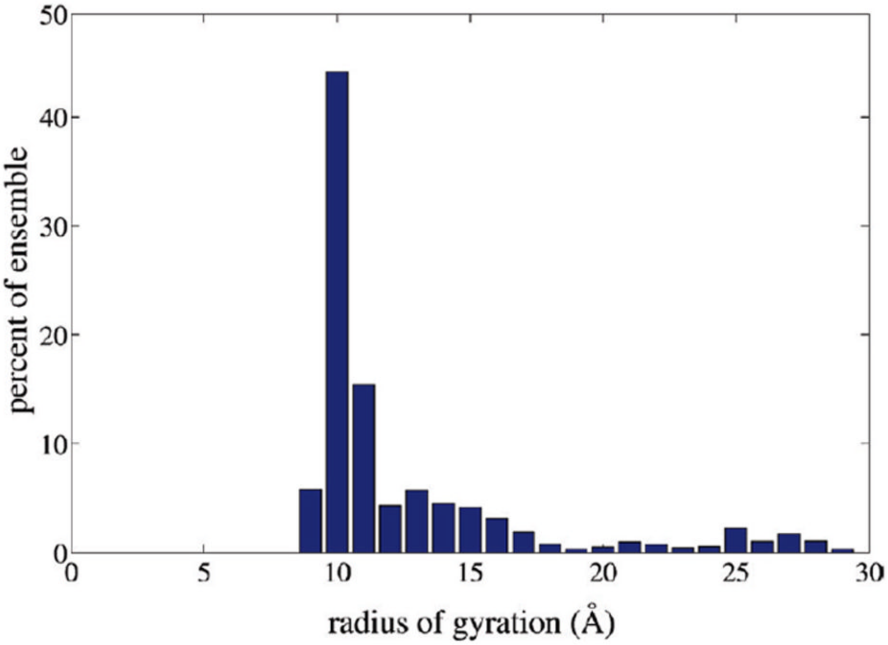


Figure 6.

Distribution of the radii of gyration of Aβ42. The peptide's radius of gyration ranges from ~9 Å, which is similar to that of globular proteins with the same number of residues, to ~30 Å, which is fully extended, with an average radius of gyration of 13 ± 4.5 Å.

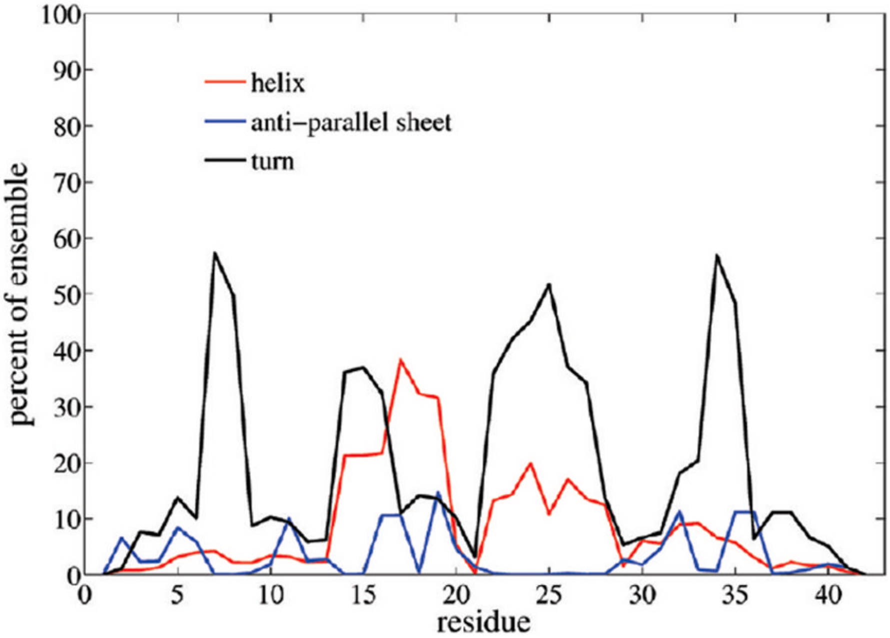
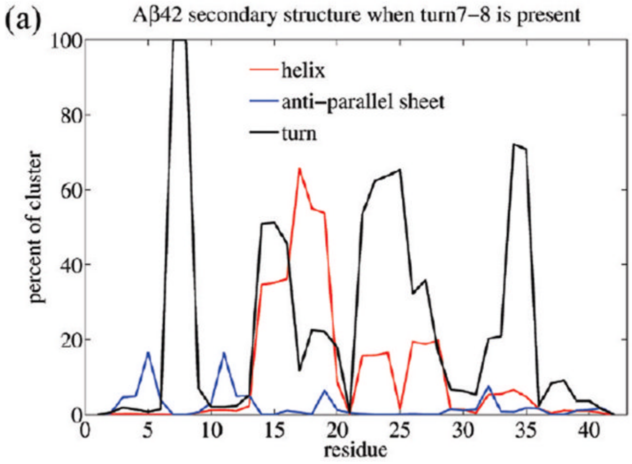
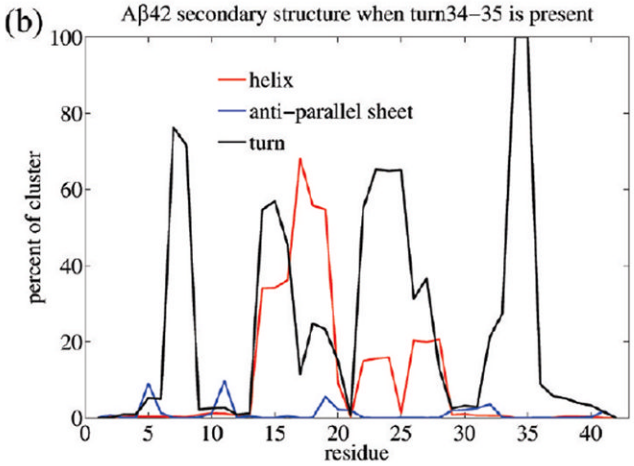


Figure 7.

Percentage of the A*β*42 ensemble in different types of secondary structure by residue. The red line represents helix, the blue line antiparallel sheet, and the black line *β*-turns.



(a) Aβ42 secondary structure when turn7–8 is present



(b) Aβ42 secondary structure when turn34–35 is present

Figure 8.

Percentages of the A*β*42 ensemble in secondary structure by residue involving the most ubiquitous turns. Fractions of ensemble in which (a) Turn7–8 is always present and (b) Turn34–35 is always present.

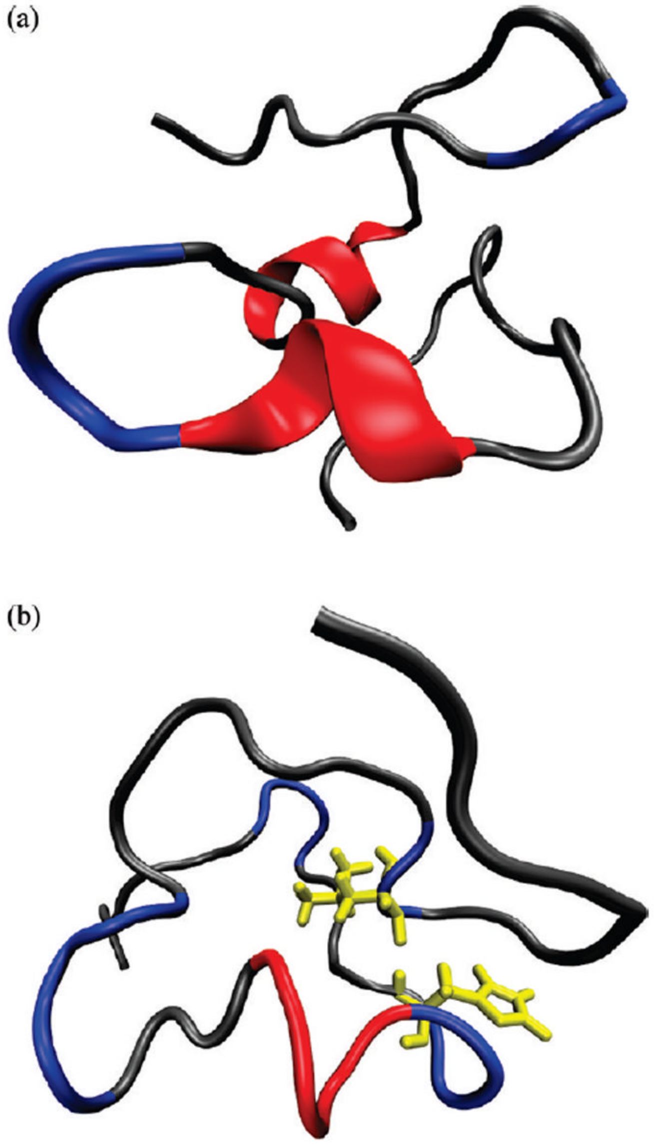


Figure 9.

Representative ribbon conformations of the Aβ42 populations: turns (blue) and helices (red). (a) Turn7–8 and Turn24–25 and Helix14–15 and Helix26–27. Each of these pairs occurs with a high frequency with each other (Table 2). (b) The hydrogen bond between backbone residues 13 and 34, which occurs in ~40% of the ensemble, occurs with a high frequency when many secondary structure elements occur simultaneously, such as Turn7–8, Turn14–15, Turn24–25, Turn26–27, and Helix17–18 shown here.

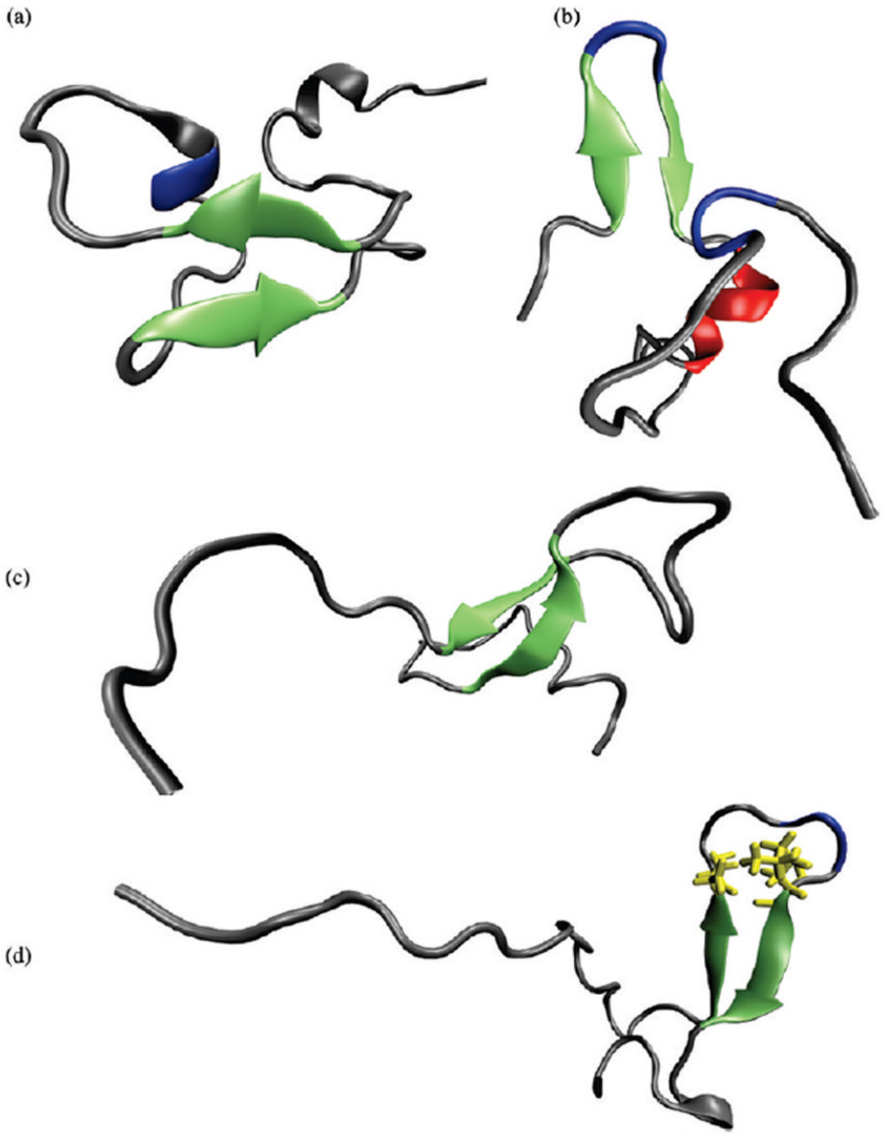


Figure 10.
Representative ribbon conformations of the A*β*42 populations: turns (blue), helices (red), and *β*-sheets (green). (a) Turn26–27 forms with *β*-sheet (Beta16–21 and Beta29–36), and no other major secondary structure is present. (b) Turn7–8 nucleating *β*-sheet Beta3-6 and Beta10–13, along with Helix14-15 and Helix17–18, and Turn34–35. (c) *β*-Sheet involving Beta16–21 and Beta29–36 in the absence of Turn26–27, or any other of the major secondary structure features. (d) Turn26–27 with *β*-sheet Beta16–21 and Beta29–36, stabilized by a salt bridge between Glu 22 and Lys 28, in the absence of other secondary structure.

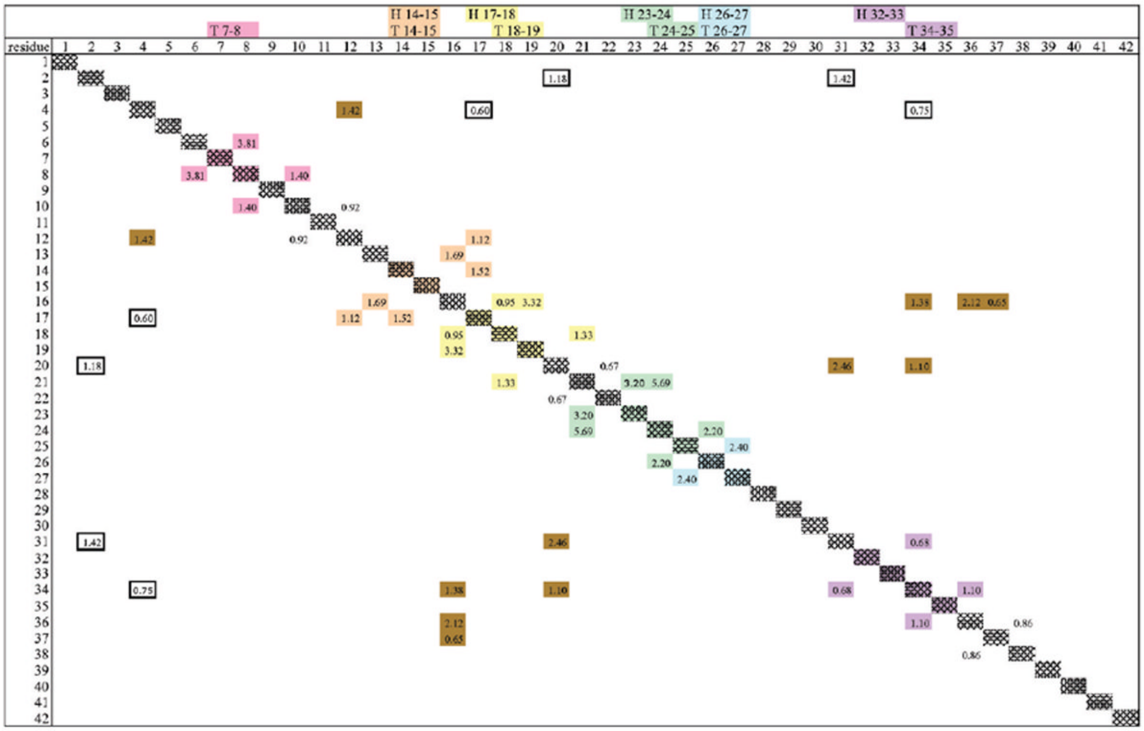


Figure 11.

Contact map of strong experimental NOE intensities for A*β*42 that are dominated by a single contact. Strong experimental intensities that define Turn7–8 (pink), Turn14–15/Helix14–15 (orange), Turn18–19/Helix17–18 (yellow), Turn24–25/Helix23–24 (green), Turn26–27/Helix26–27 (light blue), Turn34–35/Helix32–33 (purple), and *β*-sheets (brown). Hydrogen bonds that "compact" structure and involve multiple secondary structure categories occurring simultaneously are outlined in black. Intensity values are reported in each reported single contact pair.

Table 1

Aβ42 RDC Values from Experiment[27] and the Simulated Ensemble

residue	experimental RDC	experimental error	RDC from simulation	simulation error
Glu 3	3.01	0.91	2.25	0.52
Phe 4	0.23	0.57	3.23	0.20
Arg 5	0.96	1.54	3.47	0.36
Asp 7	1.72	0.57	0.59	0.23
Gly 9	0.63	0.41	1.36	1.36
Glu 11	3.45	0.17	1.18	0.34
Val 12	2.77	0.13	0.16	0.72
Gln 15	0.52	0.00	−0.69	1.02
Lys 16	7.59	2.60	0.73	1.00
Leu 17	3.33	0.57	1.89	0.44
Phe 19	3.93	0.62	2.09	2.40
Phe 20	2.34	0.12	3.52	0.69
Ala 21	3.23	2.23	1.98	1.12
Glu 22	3.34	0.16	3.05	0.73
Asp 23	3.63	0.45	0.72	1.27
Gly 25	0.42	0.14	1.05	0.54
Ser 26	0.04	0.16	2.91	0.04
Asn 27	1.94	0.38	0.47	1.10
Gly 29	1.90	0.03	0.44	0.10
Ala 30	0.68	0.30	1.04	0.13
Ile 32	2.83	0.30	0.85	0.98
Gly 33	3.70	0.08	2.02	0.63
Leu 34	1.92	0.16	2.23	1.27
Met 35	3.63	0.45	2.62	0.54
Val 36	3.56	0.28	2.78	0.12
Gly 37	1.68	0.80	1.11	0.78
Gly 38	1.47	0.44	1.50	1.09
Val 40	3.57	0.47	2.50	1.14
Ile 41	3.50	0.23	0.66	0.09
Ala 42	3.03	0.25	0.21	0.51

Table 2

Experimental and Calculated NOEs and Analysis of the Number of False Positives and False Negatives

	total no. of NOE peaks	**no. of long-range peaks (*i*−*i* + 2 or greater)**
experimentally observed peaks	705	176
experimentally assigned peaks	196	11
experimentally unassigned peaks	509	165
simulated peaks	914	188
simulated peaks agree with experimentally assigned peaks	176	0
simulated peaks agree with experimentally unassigned peaks	355	54
false negatives	174	122
false negatives found in high-noise regions	−62	−47
false negatives explained by atomic contacts on the same residue pairs	−66	−29
false negatives inconsistent with experiment	**46**	**46**
false positives	383	134
false positives found in high-noise regions	−82	−19
false positives explained by atomic contacts on the same residue pairs	−223	−81
false positives consisting of numerous weak contacts	−16	−6
false positives inconsistent with experiment	**62**	**28**

Table 3

Categorization of Secondary Structure Features Most Prevalent in the Ensemble[a]

	secondary structure region percentages						
secondary structure	**Turn7–8**	**Turn14–15/Helix14–15**	**Turn18–19/Helix17–18**	**Turn24–25/Helix23–24**	**Turn26–27/Helix26–27**	**Turn34–35**	**Helix32–33**
observed in isolation	48.90	34.02/20.78	12.52/31.82	35.24/14.23	28.32/13.36	47.25	8.16
	3.11	1.24/0.59	0.19/0.00	3.22/1.04	8.41/2.68	2.36	0.89

	simultaneous secondary structure pairing percentages						
secondary structure	Turn7–8	Turn14–15/Helix14–15	Turn18–19/Helix17–18	Turn24–25/Helix23–24	Turn26–27/Helix26–27	Turn34–35	Helix32–33
Turn7–8		24.28/16.91	10.03/26.81	24.87/7.68	14.48/9.14	33.75	2.54
Turn14–15			5.06/19.55	16.81/5.00	9.68/5.56	25.82	0.66
Turn18–19				8.28/2.08	5.14/2.96	10.52	0.09
Turn24–25					10.20/6.83	24.30	2.62
Turn26–27						14.16	0.73
Turn34–35							0.27
Helix14–15			6.74/11.85	12.78/3.56	5.96/4.76	16.05	1.82
Helix17–18				18.56/6.18	8.73/7.01	26.20	2.36
Helix23–24					2.92/1.49	7.28	4.09
Helix26–27						9.38	0.16

[a] Each column is a different secondary structural feature (see the text). In the first part of the table, we report the observed secondary structure percentages, as well as the percentage of the time they are observed in isolation. The second half of the table reports percentages of time that two secondary structure elements occur simultaneously.

Table 4

Hydrogen Bonds between Residue Pairs Most Frequently Found (>5%) in the Aβ42 Ensemble[a]

Residue –Residue hydrogen-bonds	***% in Aβ42* Ensemble**	**Residue –Residue hydrogen-bonds**	***% in Aβ42* Ensemble**
2 5	6%	18 34	10%
4 7	11%	19 32	8%
4 8	7%	21 24	16%
6 9	50%	21 25	10%
6 10	13%	21 26	24%
7 9	6%	21 27	20%
8 10	6%	22 25	42%
8 11	5%	23 26	13%
8 12	5%	24 27	18%
10 12	6%	24 28	11%
13 16	50%	25 28	24%
13 17	39%	25 29	5%
13 34	41%	26 28	9%
14 17	7%	26 29	13%
14 18	5%	29 31	9%
16 19	40%	29 32	7%
16 20	12%	31 34	13%
16 36	10%	31 35	10%
17 20	30%	32 35	8%
17 21	9%	33 36	32%

[a] Hydrogen bonds that define Turn7–8 (pink), Turn14–15/Helix14–15 (orange), Turn18–19/Helix17–18 (yellow), Turn24–25/Helix23–24 (green), Turn26–27/Helix26–27 (light blue), Turn34–35/Helix32–33 (purple), and β-sheets (brown). See Figure 10 in which colors correspond to the contact map.