Solution Structures of Rat Amylin Peptide: Simulation, Theory, and Experiment

Allam S. Reddy,[†] Lu Wang,[‡] Yu-Shan Lin,[‡] Yun Ling,[‡] Manan Chopra,[†] Martin T. Zanni,[‡] James L. Skinner,[‡] and Juan J. De Pablo[‡*]

[†]Department of Chemical and Biological Engineering and [‡]Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin

ABSTRACT Amyloid deposits of amylin in the pancreas are an important characteristic feature found in patients with Type-2 diabetes. The aggregate has been considered important in the disease pathology and has been studied extensively. However, the secondary structures of the individual peptide have not been clearly identified. In this work, we present detailed solution structures of rat amylin using a combination of Monte Carlo and molecular dynamics simulations. A new Monte Carlo method is presented to determine the free energy of distinct biomolecular conformations. Both folded and random-coil conformations of rat amylin are observed in water and their relative stability is examined in detail. The former contains an α-helical segment comprised of residues 7–17. We find that at room temperature the folded structure is more stable, whereas at higher temperatures the random-coil structure predominates. From the configurations and weights we calculate the α-carbon NMR chemical shifts, with results that are in reasonable agreement with experiments of others. We also calculate the infrared spectrum in the amide I stretch regime, and the results are in fair agreement with the experimental line shape presented herein.

INTRODUCTION

Amylin is a 37-residue peptide hormone that is produced by the islet β-cells in the pancreas. Amyloid deposits of human amylin have been identified as a hallmark of type 2 diabetes. It is generally accepted that the aggregated amylin peptide is a cause for the loss of insulin-producing pancreatic β-cells in type 2 diabetic patients (1,2).

The fibrillar structures formed by amylin have been studied extensively using experiments and simulations. Using molecular dynamics (MD) simulations, several groups have identified short amyloidogenic fragments within the amylin peptide (3,4). Simulations suggest the formation of a β-sheet structure with interdigitated side chains. Using coordinated x-ray diffraction and electron diffraction techniques, Makin and Serpell (5) determined that the fibrils of human amylin are made up of extended β-strands that run perpendicular to the fibril axis. Jayasinghe and Langen (6), using electron paramagnetic resonance spectroscopy, determined that the β-strands adopt a parallel orientation. More recently, Luca et al. (7) proposed an atomistic model of amylin aggregate based on solid-state nuclear magnetic resonance (NMR), scanning transmission electron microscopy, and atomic force microscopy. The proposed structure consists of striated ribbons containing four layers of parallel β-sheets that are formed from two symmetric layers of amylin molecules.

Although the structure of the aggregate is well understood, little is known about the aggregation pathway of the amylin peptide. Recently Shim et al. (8) studied that using isotope labeling and infrared (IR) spectroscopy. Their results indicate that during the aggregation of the peptide, the residues in the N-terminal half of the peptide form a β-sheet before a β-sheet on the C-terminal side is formed (8). This result is particularly intriguing as the residues near the N-terminus are exactly the same for the rat and human amylin peptides, but only human amylin peptide is known to form aggregates. Thus, it is important for the study of the solution structure of individual amylin molecules, and for understanding how the residues on the C-terminal side affect its conformational stability and perhaps influence its unfolding pathway.

In solution, human amylin is mostly regarded as being unstructured (9–12). However, the interpretation of experimental results is hampered by the extremely fast aggregation kinetics of the peptide. It is known that in 1,1,1,3,3,3-hexafluoro-2-propanol solution, in which amylin is predominantly monomeric, it forms a stable α-helical domain spanning residues 5–20 (13). Similarly, for the calcitonin gene-related peptide, which has a similar amino-acid sequence to that of human amylin, and belongs to the same hormone family, it is known that in trifluoric acid it forms a stable α-helical segment spanning residues 6–18 (14). Theoretical secondary-structure-prediction algorithms predict an α-helical secondary structure for residues 8–14 of human amylin (15). It has been recently shown that in the presence of phospholipid bilayers, human amylin adopts an α-helical conformation (16).

Rat amylin exhibits >80% sequence identity with human amylin; it differs from the human peptide at six residues, as shown in Fig. 1. Three proline substitutions at positions 25, 28, and 29 are of particular importance, and are believed to be the reason that rat amylin does not form aggregates (17). Given the relatively high sequence homology between the rat and human peptides, it is hypothesized that they may exhibit common structural features. Indeed, studies by

Submitted May 6, 2009, and accepted for publication October 19, 2009.

*Correspondence: depablo@engr.wisc.edu

Editor: Feng Gai.

0006-3495/10/02/0443/9 $2.00

doi: 10.1016/j.bpj.2009.10.029

Rat Amylin: KCNTATCATQRLANFLVRSSNNLGPVLPPTNVGSNTY-NH2

Human Amylin: KCNTATCATQRLANFLVHSSNNFGAILSSTNVGSNTY-NH2

FIGURE 1 Human and rat amylin sequences. Differences between the two are shown in red (for rat amylin).

Padrick and Miranker (12) suggest that rat and human amylin adopt similar prefibrillar conformations. Recently, using NMR chemical shifts, it was found that rat amylin adopts an α-helical conformation for residues 5–19 (18). To date there has been no computational effort to study the structure of full-length amylin peptides.

In this work we investigate the structure of rat amylin using a combination of atomistic Monte Carlo and molecular dynamics simulations in explicit water. We find, in fact, that there are two conformations in equilibrium—a folded structure in which residues 7–17 adopt an α-helical configuration, and a random-coil structure. Using a new variant of replica exchange umbrella sampling (REUS) Monte Carlo simulations, we evaluate the relative thermodynamic stability of the two structures, finding roughly equal free energies at room temperature. From representative configurations of each structure, and the relative weights, we calculate the NMR secondary chemical shifts for the α-carbons, and our results are in reasonable agreement with experiment (18). IR spectroscopy is another widely-used method to probe structure and dynamics of proteins (19–24). In particular, the amide I mode, primarily involving peptide-bond carbonyls, is highly sensitive to protein secondary structures (25–27). Using the MD trajectories, the relative weights of the two configurations, and local frequency and coupling maps (28–30), we calculate the IR spectrum of rat amylin in solution. To compare with these calculations, herein we measure the same spectrum experimentally. The theoretical IR spectrum is in fair agreement with experiment.

MATERIALS AND METHODS

Solution structure simulations

As mentioned above, the amino-acid sequence of rat amylin is given in Fig. 1. Note that in both NMR and IR experiments, and in the simulations, the C-terminal is capped with an NH_2 group, and there is a disulfide bond between Cys-2 and Cys-7. The starting structure used for the peptide is an α-helix, which is constructed by restricting the φ and ψ dihedral angles to –57° and –47°, respectively. The peptide is immersed in simple point charge H_2O (31); it is modeled using the GROMOS96 53a6 force field (32–34), which has been used extensively in the literature to study the folding of protein molecules (35–37). Molecular dynamics (MD) simulations were performed with the GROMACS molecular simulation package (38,39), suitably modified to incorporate the new sampling techniques described in this article (see below). The system included a single rat amylin peptide molecule, 11,218 water molecules, and four chloride counterions. Long-range electrostatic interactions were treated with a particle-mesh Ewald sum (40,41). All simulations were performed with rigid bonds (using the linear constraint solver method) and with an integration time step of 2 fs. The equilibrium simulations were performed at a temperature of 298 K and pressure of 1 bar using Berendsen coupling (42).

The relative thermodynamic stability of the folded and random-coil conformations, sampled during the equilibrium MD simulations, was determined using a new variant of the REUS method that we describe in what follows (43–47). Similar to traditional parallel tempering or replica exchange molecular dynamics (REMD) (43,48,49), in this method M independent replicas of the same system are simulated under different temperature conditions. For our problem we used 38 replicas, at temperatures chosen between 273 K and 600 K according to the procedure of Rathore et al. (50). In a typical REMD calculation for protein folding, the random-coil states are easily sampled, but the α-helical conformations are not efficiently sampled because of their larger folding timescale. To enable better sampling of the folded α-helical states of the peptide, a weak umbrella potential, represented by a set of harmonic springs, was imposed between the C_α atoms of the native contacts. Labeling the residues with an index j, for an α-helical conformation a native contact is defined by two residues with j and $j + 4$, and the distance between these residues is r_j. The corresponding equilibrium distance (determined from an ideal α-helical conformation), r_0, was taken to be 5.9 Å. The umbrella potential is therefore given by

$$U = k \sum_{j=p}^{p+q-1} (r_j - r_0)^2, \tag{1}$$

where the spring constant k was chosen to be equal to $k = 100$ kJ/mol/nm^2. For the particular case of amylin in water, we found that the α-helical domain extends over residues 7–17 (see the Results and Discussion), so that in the sum of Eq. 1 the starting residue is $p = 7$, and the number of springs is $q = 7$. The umbrella potential helps sample folded states, whereas replicas at high temperature permit sampling of unfolded states. The transitions between the folded and unfolded states can be quantified using an order parameter ξ defined as

$$\xi^2 = \frac{1}{q} \sum_{j=p}^{p+q-1} (r_j - r_0)^2. \tag{2}$$

With this choice of ξ, the umbrella potential is reduced to a simple explicit function of the order parameter:

$$U = qk\xi^2. \tag{3}$$

In a conventional REMD simulation, replicas at different temperatures are exchanged at regular intervals. Similarly, in REUS, configurations from two different replicas are exchanged at regular intervals. The acceptance criterion for the exchange is obtained by satisfying detailed balance criteria. We label each replica with an index m, and this replica's inverse temperature by β_m. One can then label the configuration residing at an instant t in this replica by the pair mt. We define the state X to have configurations mt and nt in the two replicas m and n (at time t), whereas in state X' these two configurations are exchanged. By invoking detailed balance we can write

$$P(X)w_p(X \to X')w_{acc}(X \to X') = P(X')w_p(X' \to X)w_{acc}(X' \to X), \tag{4}$$

where $P(X)$ denotes the probability of finding the system in state X, $w_p(X \to X')$ represents the probability of proposing an exchange between state X and state X', and w_{acc} $(X \to X')$ represents the probability of accepting the proposed exchange between state X and state X'. If the forward and reverse moves are proposed with equal probability, then $w_p(X \to X') = w_p(X' \to X)$. Hence, the acceptance criteria for the exchange between states X and X' is given by

$$w_{acc}(X \to X') = \min\left(1, \frac{P(X')}{P(X)}\right). \tag{5}$$

The ratio of the equilibrium probabilities is given by

$$\frac{P(X')}{P(X)} = \frac{\exp(-\beta_m E_{nt} - \beta_m U_{nt})\exp(-\beta_n E_{mt} - \beta_n U_{mt})}{\exp(-\beta_m E_{mt} - \beta_m U_{mt})\exp(-\beta_n E_{nt} - \beta_n U_{nt})}, \quad (6)$$

where E_{mt} is the potential energy of configuration *mt* (and U_{mt} is the umbrella potential in configuration *mt*). Defining $\Delta E = E_{nt} - E_{mt}$, $\Delta U = U_{nt} - U_{mt}$, and $\Delta\beta = \beta_n - \beta_m$, Eq. 5 becomes

$$w_{acc}(X \to X') = \min(1, \exp(\Delta\beta\Delta E)\exp(\Delta\beta\Delta U)). \quad (7)$$

For each replica, the configurations from the simulation were saved every 4 ps for a total of 50 ns.

The potential of mean force or free energy for folding the peptide from a random coil to an α-helical state was calculated as a function of the order parameter ξ using the weighted histogram analysis method (51). In this method the order parameter ξ is discretized, and then one considers the (unnormalized) probability histogram $P_{k,\beta}(\xi)$ as a function of order parameter ξ, which is obtained from

$$P_{k,\beta}(\xi) = \sum_{m=1}^{M}\sum_{t=1}^{T} \frac{\exp(-\beta E_{mt} - \beta U_{mt})\delta_{\xi,\xi_{mt}}}{\sum_{n=1}^{M}\exp(f_n - \beta_n E_{nt} - \beta_n U_{nt})}, \quad (8)$$

and

$$f_m = -\ln\sum_{\xi} P_{k,\beta_m}(\xi). \quad (9)$$

In the above equation, T is the total number of data points (snapshots) in each replica. These two equations are solved iteratively, until f_m and $P_{k,\beta}(\xi)$ are consistent with each other. The potential of mean force is determined from the probability distribution $P_{k,\beta}(\xi)$ according to

$$\phi(\xi) = -kT \ln P_{0,\beta}(\xi). \quad (10)$$

NMR secondary chemical shifts

NMR chemical shifts are known to be a sensitive probe of secondary structure (52). Secondary chemical shifts are the measured shifts minus their corrected random-coil values (18,53). A variety of secondary chemical shifts were determined for each residue, using ^{13}C- and ^{15}N-labeled wild-type rat amylin, and triple-resonance techniques (18). In all (H_α, H_N, carbonyl C, C_α, and C_β) cases, the secondary chemical shifts were indicative of α-helical structure from residues 5–19. For example, for these residues the C_α chemical shifts are all positive.

From our ensemble of simulation configurations for both the α-helix and random-coil states, together with the calculated weights of these two configurations, we can make a quantitative comparison with these experiments. For each configuration, NMR secondary chemical shifts were calculated using the SPARTA (53) and SHIFTX (54) programs, available from http://spin.niddk.nih.gov/bax and http://redpoll.pharmacy.ualberta.ca/shiftx/. For each conformer, the calculated shifts were obtained as an average over a 100-ns simulation trajectory. We focused on the α-carbon ^{13}C chemical shifts, and for each residue the random-coil value was subtracted (53).

IR line shape

The absorption line shape can be calculated from the Fourier transform of the quantum dipole time-correlation function (55). If the electric field of the excitation light is polarized in the $\hat{\epsilon}$ direction, the linear absorption line shape is

$$I(\omega) \sim Re\int_0^\infty dt\, e^{-i\omega t}\langle\hat{\epsilon}\cdot\vec{\mu}(0)\vec{\mu}(t)\cdot\hat{\epsilon}\rangle, \quad (11)$$

where $\vec{\mu}$ is the dipole operator for the system. The brackets indicate a quantum equilibrium statistical mechanical average, which is impossible to evaluate for a nontrivial condensed phase system.

If one is focused on a particular IR band, in this case the amide I band of the peptide, one useful approximation is to treat the amide I vibrational modes quantum mechanically, ignore other high-frequency modes, and treat the low-frequency degrees of freedom (translations, rotations, and torsions) classically. Within such a mixed quantum/classical approach, the IR line shape for a multi-chromophore system can be written as (56)

$$I(\omega) \sim Re\int_0^\infty dt\, e^{-i\omega t}\sum_{ij}\langle m_i(0)F_{ij}(t)m_j(t)\rangle e^{-t/2T_1}. \quad (12)$$

In the above equation, i and j index the amide I vibrational chromophore; $m_i(t) = \vec{\mu}_i(t)\cdot\hat{\epsilon}$, where $\vec{\mu}_i(t)$ is the transition dipole of the i^{th} chromophore at time t; and $F_{ij}(t)$ values are the elements of the matrix $F(t)$, which satisfies the equation

$$\dot{F}(t) = iF(t)\kappa(t), \quad (13)$$

subject to the initial condition $F_{ij}(0) = \delta_{ij}$, and with

$$\kappa_{ij}(t) = \omega_i(t)\delta_{ij} + \omega_{ij}(t)(1 - \delta_{ij}). \quad (14)$$

$\omega_i(t)$ values represent the local-mode transition frequencies and $\omega_{ij}(t)$ values are the vibrational couplings. The angular brackets now indicate a classical equilibrium statistical mechanical average. T_1 is the lifetime of the first excited state of an isolated amide I vibration, and the term $e^{-t/2T_1}$ is added phenomenologically to include lifetime broadening. T_1 is set to be 600 fs (57).

The amide I local-mode frequencies are calculated with the empirical frequency map developed by Lin et al. (28):

$$\omega_i = \{1717 + 4213E_{i,C} + 2108E_{i,N}\}\ cm^{-1}. \quad (15)$$

In the map, local-mode frequencies are related to the electric fields due to nearby water molecules, counterions, and peptide atoms that are more than three covalent bonds away. $E_{i,C}$ is the electric-field component for the i^{th} chromophore in the C=O direction on the C atom and $E_{i,N}$ is that on the N atom (both in atomic units). A 20 Å cutoff is used for electric field calculations in the MD simulations. It has been shown that couplings between adjacent peptide units are mainly due to overlapping charge densities, which cause through-bond interactions, whereas those between nonadjacent peptide units are primarily long-distance electrostatic interactions (30). The former depend on nearby dihedral angles, and for which we have adopted the approach of Jansen et al. (29) and used their nearest-neighbor coupling map. The latter are calculated with the transition-dipole coupling approximation (30,58)

$$\hbar\omega_{ij} = \frac{\vec{\mu}_i\cdot\vec{\mu}_j}{r_{ij}^3} - 3\frac{(\vec{\mu}_i\cdot\vec{r}_{ij})(\vec{\mu}_j\cdot\vec{r}_{ij})}{r_{ij}^5}. \quad (16)$$

In the above equation, the vector $\vec{r}_{ij}$ connects the two transition dipoles $\vec{\mu}_i$ and $\vec{\mu}_j$. The transition dipole has the strength of 0.37 D. It points toward the nitrogen atom with an angle of 20° with respect to the C=O bond, and its location is on the C=O bond with a distance of 0.868 Å from the carbon atom (30,58).

If the rotational dynamics of the system are sufficiently fast, then during the course of the simulation the peptide will sample all possible orientations with respect to the electric field unit vector $\hat{\epsilon}$. However, it is not the case here, and so one needs instead to average over all possible orientations of $\hat{\epsilon}$ with respect to the lab-fixed axes. This is equivalent to averaging over the three cases where $\hat{\epsilon}$ is $\hat{x}$, $\hat{y}$, and $\hat{z}$, which is what we do here.

To calculate the IR spectrum, MD simulations with the same parameters described in the solution structure simulations section were performed. The MD simulation code was modified to report the local-mode frequencies

using Eq. 15. The local-mode frequencies and coordinate trajectories were saved every 10 fs for a total of 2 ns for each conformer. The latter were used to calculate vibrational couplings. The IR line shape evaluation included 37 carbonyls in the main-chain peptide bonds and seven carbonyls from Asn and Gln side chains. The line shape was calculated by direct numerical integration of Eq. 13 (59,60).

Experimental peptide preparation and IR spectrum

We first note that these experiments were performed in D_2O (rather than H_2O) solution, to avoid the strong spectral overlap between the H_2O bend and amide I stretch. However, the simulations were performed with H_2O. Because all potentials involving D_2O are the same as for H_2O, the ensembles of configurations in the two cases are identical. This means, for example, that calculations of relative free energies and secondary chemical shifts are not affected by this change. IR spectra are often affected by molecular dynamics, and because the masses of H_2O and D_2O are different, there will be differences in time-dependent solvent-induced frequency and coupling fluctuations. These differences, however, are expected to be small.

Rat amylin were purchased from Bachem (Bubendorf, Switzerland) and dissolved in deuterated hexafluoroisopropanol (d-HFIP) at a concentration of 0.5 mM. An aliquot of the d-HFIP stock solution was made into D_2O solution after evaporating d-HFIP. The rat amylin D_2O solution was lyophilized in 0.1 mM DCl multiple times and then one time in D_2O to remove residual trifluoric acid (which was present in the purchased sample, and which has a strong absorption in the amide I band region). The final peptide sample was in D_2O at pH 6 with a concentration of 0.5 mM. The spectrum was taken in an IR cell, which consisted of two CaF_2 windows separated by a 100-μm Teflon spacer.

RESULTS AND DISCUSSION

Solution structure determination

First, we investigate the equilibrium solution structures of rat amylin. The starting structure for the MD simulations was prepared by folding all of the peptide's residues into an α-helix conformation. The energy of the resulting structure was then minimized to remove bad contacts. Within a short period of time ($<$1 ns), the molecule unfolded into a conformation where residues 7–17 assumed an α-helix; the rest of the peptide adopted an unstructured, random-coil-like state. The system was equilibrated at length before starting actual production runs. From this point on, four independent simulations were performed to analyze the stability of the structure thus obtained. Simulation runs 1, 2, and 3 were performed using the GROMOS96 53a6 force field. A representative snapshot from this run is presented in Fig. 2. Run 4 was performed using the AMBER99 (61) force field, to also analyze the effect of force field on the structure. The structure showed no further important conformational changes even after relatively long simulation times in all the four independent runs (225 ns for Runs 1 and 4, and 100 ns for Runs 2 and 3). The root mean-square deviation (RMSD) of the Cα atoms belonging to domain 7–17, with respect to the starting structure, is presented in Fig. 3. In the RMSD calculation, the peptide is rotated and translated so as to minimize the RMSD (with respect to the starting structure). The low value for RMSD throughout the entire length of the

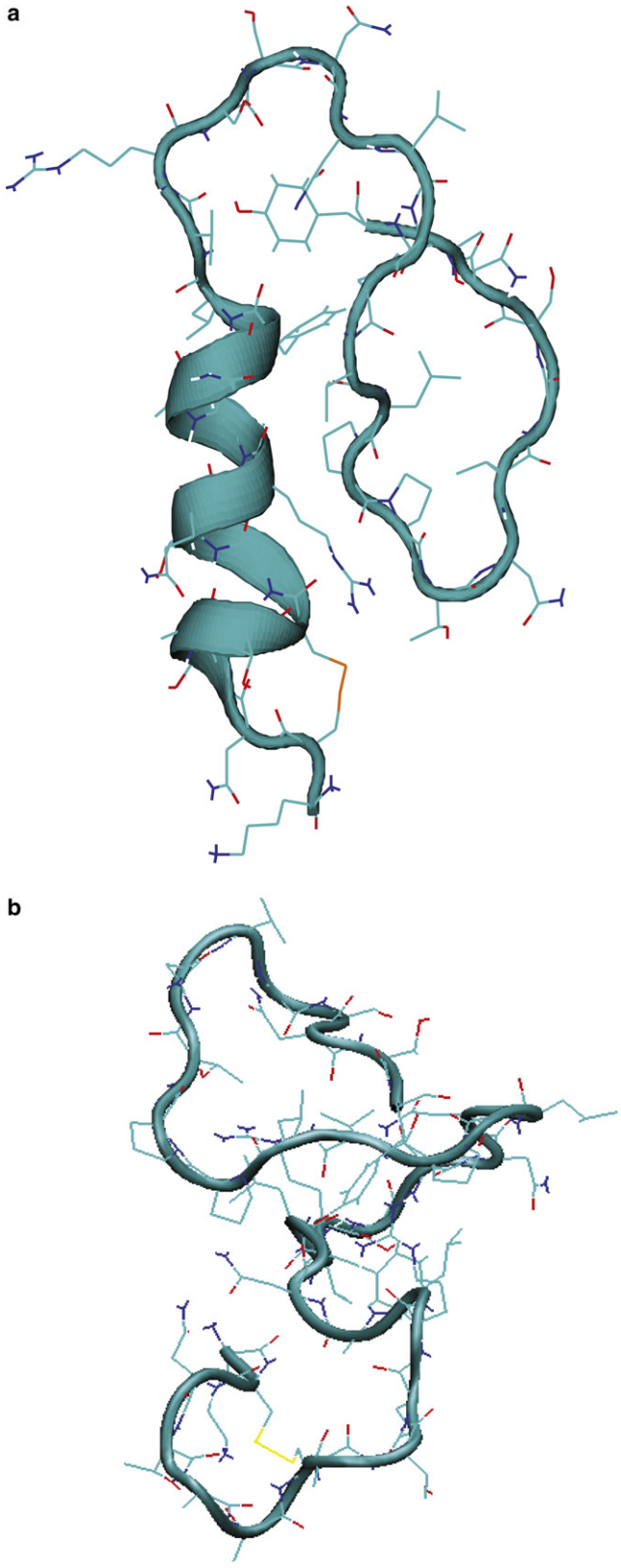

FIGURE 2 A snapshot of rat amylin in (*a*) α-helical conformation and (*b*) random coil conformation.

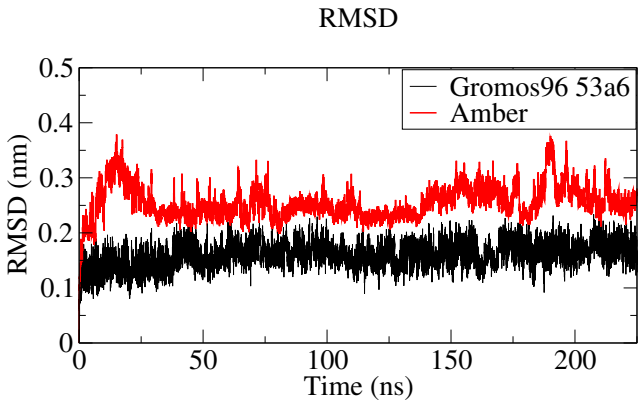

FIGURE 3 For the α-helical conformation: root mean-square deviation (RMSD), from the configuration at time 0, of the C_α atoms belonging to domain 7–17.

simulation in all four runs indicates a high stability for the predicted structure. The evolution of secondary structure of rat amylin calculated using the Dictionary of Protein Secondary Structure (62) definitions is presented in Fig. 4. It is important to note that in all four runs considered here, residues 7–17 remained folded in the α-helix conformation throughout the entire length of the trajectory.

Direct MD runs at constant temperature (such as those shown in Fig. 2) serve to provide an indication of the stability of a particular conformational state, but they cannot provide an unambiguous measure of the thermodynamic stability of a particular structure, particularly if large free energy barriers separate one conformational state of the system from another. To determine the true relative stability of the α-helical conformation vis-a-vis that of the random coil, the free energy of the system was evaluated by resorting to REUS simulations. Such simulations require that the probability distribution curves corresponding to neighboring replicas exhibit a sufficient degree of overlap. Fig. S1 in the Supporting Material shows the combined energy (potential energy + umbrella energy) distribution of the protein at different temperatures used in our REUS simulations. As can be seen from the figure, the overlap between adjacent replicas is appropriate and does lead to a near optimal acceptance rate of ~15%, as prescribed by Rathore et al. (50). Fig. S2 follows the history of two configurations throughout an REUS simulation; by history, we refer to the replica and order parameter that a configuration will occupy and adopt at any given time during the simulation. As can be seen in Fig. S2, the exchange of configurations at regular intervals between replicas running at different temperatures allows the peptides to fold and unfold repeatedly over the course of a simulation, thereby ensuring efficient sampling of conformational space and guaranteeing that the α-helical structure of the molecule is erased and reformed multiple times. Fig. 5 *a* shows the corresponding free-energy change of the peptide during the helix-coil transformation at three temperatures: 298, 346, and 416 K. In all cases the free-energy curves show (at least) two minima: at ~$\xi = 0.1$ nm (the α-helical state) and ~$\xi = 0.4$ nm (the random-coil state). At room temperature, the peptide exhibits a slightly greater stability in the α-helical state (with probability of roughly 0.55) compared to the random coil (whose probability is ~0.45). At the highest temperature, the random coil is considerably more stable. The convergence of the simulation was assessed by confirming that the free energy profile of

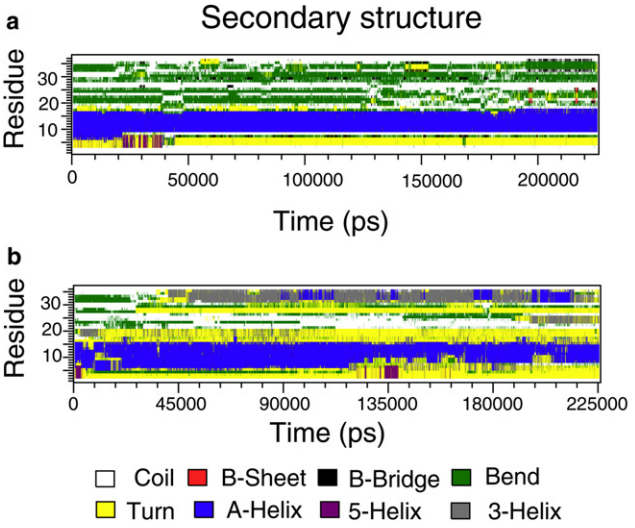

FIGURE 4 Evolution of secondary structure as a function of residue calculated using the Dictionary of Protein Secondary Structure definitions (62). (*a*) GROMOS96 53a6 force field. (*b*) AMBER99 force field.

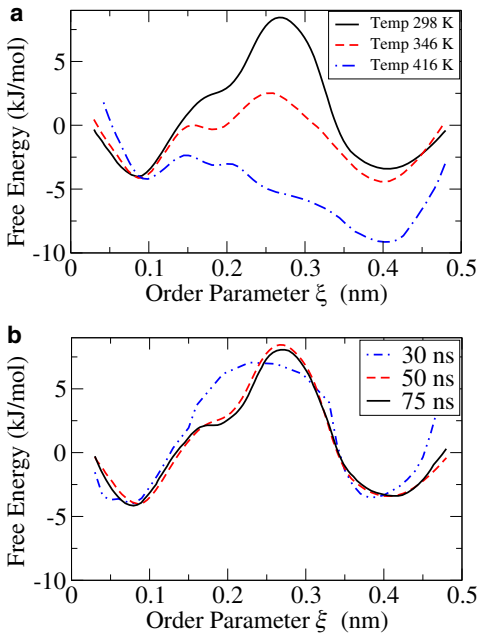

FIGURE 5 (*a*) Free energy as a function of α-helix order parameter, as described in the text. (*b*) Free energy curve at 298 K for three different lengths (30, 50, and 75 ns) of simulation time.

Fig. 5 *a* did not change appreciably after the simulations were extended over longer periods of time (results shown in Fig. 5 *b*). As can be seen in the figure, at early times (e.g., 30 ns) the free energy profiles change appreciably, but after ~50 ns they reach a converged steady-state value.

NMR secondary chemical shifts

The NMR $^{13}C_{\alpha}$ secondary chemical shifts were calculated as described above for each of the two conformers predicted by our REUS simulations (the α-helical and the random coil). The theoretical secondary NMR shifts of the α-helical and random coil conformations are shown in Fig. 6. One sees that, for the α-helical conformer (Fig. 6, *a* and *b*), the secondary chemical shifts are uniformly positive, with some magnitudes reaching values as large as 3 ppm in strongly α-helical regions (residues 7–17). In contrast, for the random coil conformation (Fig. 6, *c* and *d*), the shifts are smaller and appear to be randomly distributed. The relative probability analysis described above indicates that the equilibrium solution structure is composed of 55% of the α-helical state and 45% of the random-coil state. The theoretical secondary NMR shifts averaged over the two conformations, weighted by their relative stabilities, are shown in Fig. 7, together with the experimental results (18). As can be seen in Fig. 7, *a* and *b*, the predicted NMR shifts can change appreciably depending on the algorithm or method (SPARTA or SHIFTX) used to determine NMR spectra from results of simulations. Given such limitations, we first

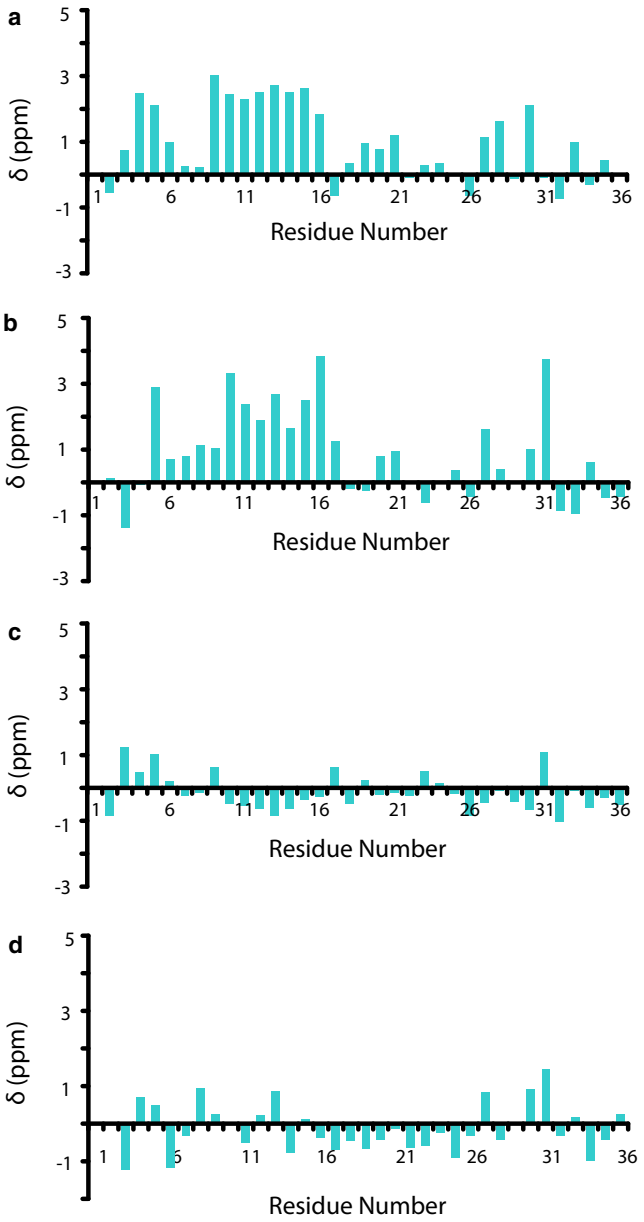

FIGURE 6 Calculated NMR $^{13}C_{\alpha}$ secondary chemical shifts for: α-helical (*a* and *b*) and random-coil conformations (*c* and *d*) using the SPARTA (*a* and *c*) and SHIFTX (*b* and *d*) programs (53,54).

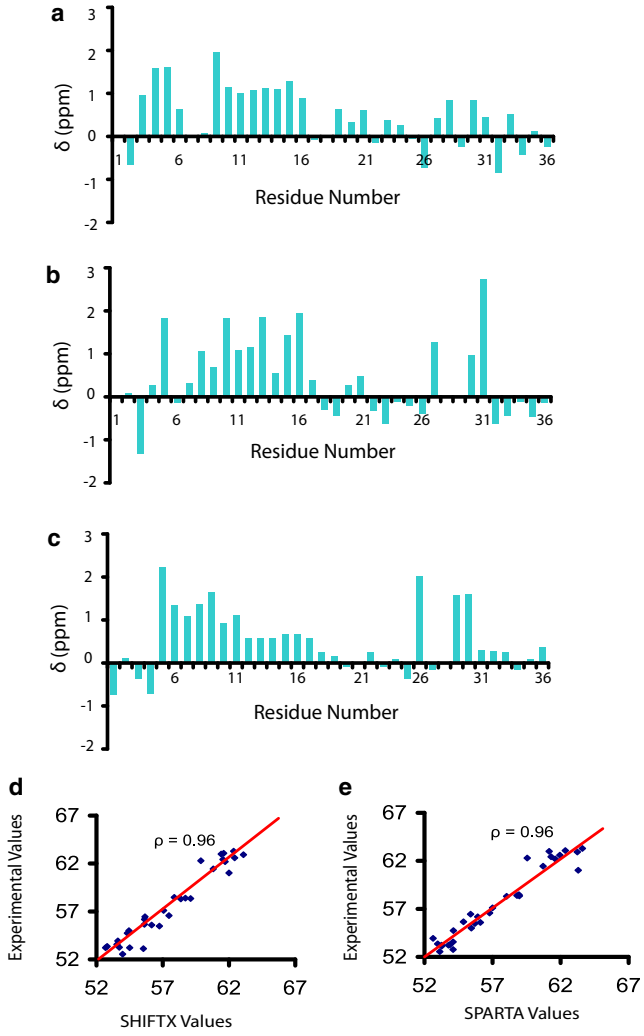

FIGURE 7 (*a* and *b*) Theoretical NMR secondary chemical shifts from the weighted average of the results in the previous figure using the SPARTA (*a*) and SHIFTX (*b*) programs. (*c*) Experimental NMR secondary chemical shifts (18). (*d* and *e*) Correlation between experimentally observed chemical shifts and those obtained from SPARTA and SHIFTX programs.

focus our discussion on the most salient features of the experimental and calculated spectra: in all cases the shifts are positive. They have approximately the same magnitude, particularly for residues between 7 and 17, which is where the α-helix is located. Our results can thus be viewed as generally consistent with experiment, and serve to confirm the overall α-helical character of the region comprised between such residues.

If we now make a more-quantitative comparison between simulated and experimental spectra, one finds that the primary differences arise for residues outside the helical (7–17) region. To gain some perspective into such differences, it is important to specify several details pertaining to the experimental spectra reported by Williamson and Miranker (18). First, in experiments Pro is observed to make a *cis-trans* conformational change. However, this being a very slow transition, it is usually not observed with any of the force fields generally used for protein simulations (63). This could explain the differences for residues 25, 28, and 29. Second, in calculating deviations from random coil, in their article, Williamson and Miranker subtract a random coil value that is sequence-dependent. For example, for Asparagine at residue numbers 3, 14, 22, and 31, a value of 52.9 is subtracted, whereas for Asparagine at residue number 21, a value of 52.3 is subtracted. Similarly for Leu-12, 16, and 23, they subtract a value of 55.1, whereas for Leu-27, a value of 52.7 is subtracted. In contrast, the algorithms subtracts an equal value for all residues. When these details are taken into consideration, we see that the agreement between simulations and experiment is in fact reasonable. Again, we emphasize that the key piece of information from the experiments and simulations is the α-helical nature of residues 7–17.

IR line shape calculation

Fig. 8 shows the theoretical IR spectra of the α-helical and random-coil conformations, each weighted by their relative populations. The total theoretical spectrum is obtained by adding the results for the α-helix and random-coil, and the experimental spectrum measured herein. Note that the theoretical and experimental spectra are normalized to have the same peak height. The theoretical line shape for the α-helical conformer exhibits a shoulder at ~1657 cm^{-1}, which is characteristic of α-helices (64). However, its main peak is ~1631 cm^{-1}, which is very similar to the peak in the theoretical line shape for the random-coil state. Thus, the total theoretical line shape is unstructured (as in experiment), but it is too red-shifted by ~17 cm^{-1}, and too broad by ~21 cm^{-1}.

CONCLUSIONS

Atomistic-scale theoretical simulations of rat amylin in solution have been presented. The folded state of rat amylin is found to contain a helical segment spanning residues 7–17. A newly proposed variant of replica exchange simulations permitted careful quantification of the relative stability of the helical and random-coil conformations as a function of temperature. It was found that at room temperature the α-helical conformation was marginally more stable than the random-coil state.

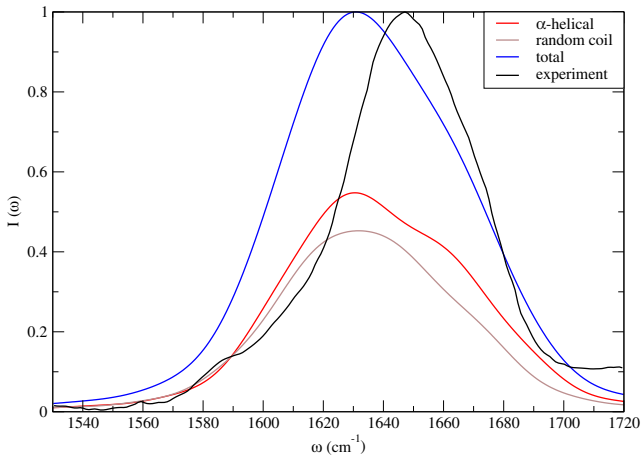

FIGURE 8 Theoretical (total) and experimental IR line shapes in the amide I stretch region. Also shown are line shapes for the α-helical and random-coil states, weighted by their relative probabilities.

Trajectories for the two conformers, together with their relative probabilities, were used to calculate both NMR secondary chemical shifts, and amide I IR spectra. In both cases agreement with experiment was only qualitative. Discrepancies between theory and experiment for the NMR observables could be ascribed to deficiencies in the force field used, in the SPARTA algorithm to calculate the shifts, and/or to insufficient sampling. Discrepancies between theory and experiment for the IR observable could be ascribed to force-field and sampling issues, as well as to deficiencies in our theoretical approach for calculating IR spectra—in particular, those involving frequency and coupling maps. Thus, these findings indicate the need for further work on developing and comparing force fields, and on theoretical spectroscopic methodology. Nonetheless, given the complexity of the amylin molecule and the level of detail at which structure is being resolved, even the qualitative agreement between theory and experiment is encouraging. It shows the promise of a combined simulation, theoretical, and experimental approach for elucidating protein structure and dynamics, particularly in circumstances (such as membrane proteins, or systems with fast aggregation kinetics) in which a single approach may well be inadequate.

The similarity of rat and human amylin sequences near the N-terminus suggests that human amylin might also exhibit a similar α-helical secondary structure. In fact, our preliminary simulations of human amylin protein suggest that the human amylin protein exhibits conformations with a similar α-helical structure between residues 7 and 17; however,

in the human case one also finds an additional β-hairpin appearing between residues 23–27 and 30–34 (Allam S. Reddy, Sadanand Singh, Lu Wang, Yun Ling, Martin T. Zanni, James L. Skinner, and Juan J. De Pablo, unpublished results). Our preliminary results also indicate that the unfolding pathway of the human amylin protein involves formation of contacts between the β-hairpin segment of the protein with the α-helical structure, leading to an overall loss of α-helicity and unfolding of the peptide. The proline mutations in rat amylin (at residues 25, 28, and 29) prevent the formation of such a β-hairpin, thereby precluding the occurrence of an unfolding process similar to that observed in the human version of the peptide.

Recently, Engel et al. (65) studied the interaction of amylin with phospholipid monolayers. Their results suggest that amylin inserts into lipid monolayers as a monomer. Furthermore, they suggest that insertion occurs at the N-terminus. It is known that the insertion ability of proteins into monolayers is influenced by their secondary structure in solution. The α-helical segment near the N-terminus, as predicted in our work (and also by the NMR experiments (18)), would promote its insertion into the monolayers. In addition, amylin in its aggregated state would lose the α-helical secondary structure, thereby making the insertion of the peptide less favorable, as observed by Engel et al. (65). Ongoing work in our labs (66,67), particularly using isotope labels and 2DIR spectroscopy, is focused on human amylin and its aggregation pathways, in the absence and presence of membranes.

SUPPORTING MATERIAL

Two figures are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(09)01671-3.

This work was supported in part by a grant from the National Science Foundation (No. CHE-0832584) to J.L.S. and M.T.Z. J.L.S. also thanks the National Science Foundation for support of this work through grant No. CHE-0750307. M.T.Z. thanks the National Institutes of Health (grant No. DK79895) for support. J.J.dP. thanks the National Science Foundation for support of this work through the University of Wisconsin *Materials Research Science and Engineering Center* on Nanostructured Interfaces, through grant No. CBET-0755730, and through grant No. DMR-0503940.

REFERENCES

1. Clark, A., G. J. S. Cooper, …, R. C. Turner. 1987. Islet amyloid formed from diabetes-associated peptide may be pathogenic in type-2 diabetes. *Lancet*. 2:231–234.
2. Lorenzo, A., B. Razzaboni, …, B. A. Yankner. 1994. Pancreatic islet cell toxicity of amylin associated with type-2 diabetes mellitus. *Nature*. 368:756–760.
3. Zanuy, D., and R. Nussinov. 2003. The sequence dependence of fiber organization. A comparative molecular dynamics study of the islet amyloid polypeptide segments 22–27 and 22–29. *J. Mol. Biol.* 329: 565–584.
4. Wu, C., H. Lei, and Y. Duan. 2005. Elongation of ordered peptide aggregate of an amyloidogenic hexapeptide NFGAIL observed in molecular dynamics simulations with explicit solvent. *J. Am. Chem. Soc.* 127:13530–13537.
5. Makin, O. S., and L. C. Serpell. 2004. Structural characterization of islet amyloid polypeptide fibrils. *J. Mol. Biol.* 335:1279–1288.
6. Jayasinghe, S. A., and R. Langen. 2004. Identifying structural features of fibrillar islet amyloid polypeptide using site-directed spin labeling. *J. Biol. Chem.* 279:48420–48425.
7. Luca, S., W.-M. Yau, …, R. Tycko. 2007. Peptide conformation and supramolecular organization in amylin fibrils: constraints from solid-state NMR. *Biochemistry*. 46:13505–13522.
8. Shim, S. H., R. Gupta, …, M. T. Zanni. 2009. Two-dimensional IR spectroscopy and isotope labeling defines the pathway of amyloid formation with residue-specific resolution. *Proc. Natl. Acad. Sci. USA*. 106:6614–6619.
9. Kayed, R., J. Bernhagen, …, A. Kapurniotu. 1999. Conformational transitions of islet amyloid polypeptide (IAPP) in amyloid formation in vitro. *J. Mol. Biol.* 287:781–796.
10. Dunker, A. K., J. D. Lawson, …, Z. Obradovic. 2001. Intrinsically disordered protein. *J. Mol. Graph. Model.* 19:26–59.
11. Jaikaran, E. T., and A. Clark. 2001. Islet amyloid and type 2 diabetes: from molecular misfolding to islet pathophysiology. *Biochim. Biophys. Acta*. 1573:179–203.
12. Padrick, S. B., and A. D. Miranker. 2001. Islet amyloid polypeptide: identification of long-range contacts and local order on the fibrillogenesis pathway. *J. Mol. Biol.* 308:783–794.
13. Cort, J., Z. Liu, …, N. H. Andersen. 1994. Beta-structure in human amylin and two designer β-peptides: CD and NMR spectroscopic comparisons suggest soluble β-oligomers and the absence of significant populations of β-strand dimers. *Biochem. Biophys. Res. Commun.* 204:1088–1095.
14. Breeze, A. L., T. S. Harvey, …, I. D. Campbell. 1991. Solution structure of human calcitonin gene-related peptide by 1H NMR and distance geometry with restrained molecular dynamics. *Biochemistry*. 30: 575–582.
15. Jaikaran, E. T., C. E. Higham, …, P. E. Fraser. 2001. Identification of a novel human islet amyloid polypeptide β-sheet domain and factors influencing fibrillogenesis. *J. Mol. Biol.* 308:515–525.
16. Knight, J. D., J. A. Hebda, and A. D. Miranker. 2006. Conserved and cooperative assembly of membrane-bound α-helical states of islet amyloid polypeptide. *Biochemistry*. 45:9496–9508.
17. Green, J., C. Goldsbury, …, U. Aebi. 2003. Full-length rat amylin forms fibrils following substitution of single residues from human amylin. *J. Mol. Biol.* 326:1147–1156.
18. Williamson, J. A., and A. D. Miranker. 2007. Direct detection of transient α-helical states in islet amyloid polypeptide. *Protein Sci.* 16:110–117.
19. Fayer, M. D. 2001. Ultrafast Infrared and Raman Spectroscopy. Marcel-Dekker, New York.
20. Hochstrasser, R. M. 2007. Two-dimensional spectroscopy at infrared and optical frequencies. *Proc. Natl. Acad. Sci. USA*. 104:14190–14196.
21. Mukamel, S. 2000. Multidimensional femtosecond correlation spectroscopies of electronic and vibrational excitations. *Annu. Rev. Phys. Chem.* 51:691–729.
22. Park, S., K. Kwak, and M. D. Fayer. 2007. Ultrafast 2D-IR vibrational echo spectroscopy: a probe of molecular dynamics. *Laser Phys. Lett.* 4:704–718.
23. Cho, M. 2008. Coherent two-dimensional optical spectroscopy. *Chem. Rev.* 108:1331–1418.
24. Manor, J., P. Mukherjee, …, I. T. Arkin. 2009. Gating mechanism of the influenza A M2 channel revealed by 1D and 2D IR spectroscopies. *Structure*. 17:247–254.
25. Surewicz, W. K., H. H. Mantsch, and D. Chapman. 1993. Determination of protein secondary structure by Fourier transform infrared spectroscopy: a critical assessment. *Biochemistry*. 32:389–394.

26. Susi, H., and D. M. Byler. 1986. Resolution-enhanced Fourier transform infrared spectroscopy of enzymes. *Methods Enzymol.* 130:290–311.
27. Haris, P. I., and D. Chapman. 1992. Does Fourier-transform infrared spectroscopy provide useful information on protein structures? *Trends Biochem. Sci.* 17:328–333.
28. Lin, Y.-S., J. M. Shorb, …, J. L. Skinner. 2009. Empirical amide I vibrational frequency map: application to 2D-IR line shapes for isotope-edited membrane peptide bundles. *J. Phys. Chem. B.* 113: 592–602.
29. Jansen, T. L., A. G. Dijkstra, …, J. Knoester. 2006. Modeling the amide I bands of small peptides. *J. Chem. Phys.* 125:044312.
30. Gorbunov, R. D., and G. Stock. 2007. Ab initio based building block model of amide I vibrations in peptides. *Chem. Phys. Lett.* 437:272–276.
31. Berendsen, H. J. C., J. P. M. Postma, …, J. Hermans. 1981. *In* Intermolecular Forces., B. Pullman, editor. Reidel, Dordrecht.
32. van Gunsteren, W., S. Billeter, …, I. Tironi. 1996. Biomolecular Simulations: the GROMOS96 Manual and User Guide. BIOMOS bv. Laboratory of Physical Chemistry, ETH Zentrum, Groningen, The Netherlands.
33. Scott, W., P. Hunenberger, …, W. van Gunsteren. 1999. The GROMOS biomolecular simulation program package. *J. Phys. Chem. A.* 103:3596–3607.
34. Oostenbrink, C., A. Villa, …, W. Van Gunsteren. 2004. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J. Comp. Chem.* 25:1656–1676.
35. Shirts, M. R., and V. S. Pande. 2005. Solvation free energies of amino acid side chain analogs for common molecular mechanics water models. *J. Chem. Phys.* 122:134508.
36. Zagrovic, B., and W. Van Gunsteren. 2006. Comparing atomistic simulation data with the NMR experiment: how much can NOEs actually tell us? *Proteins Struct. Funct. Bioinform.* 63:210–218.
37. Zhou, Y., C. Oostenbrink, …, J. Jongejan. 2006. Computational study of ground-state chiral induction in small peptides: comparison of the relative stability of selected amino acid dimers and oligomers in homochiral and heterochiral combinations. *J. Comp. Chem.* 27:857–867.
38. Lindahl, E., B. Hess, and D. van der Spoel. 2001. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model.* 7:306–317.
39. Berendsen, H. J. C., D. van der Spoel, and R. van Drunen. 1995. GROMACS: a message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* 91:43–56.
40. Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald: an $N\cdot$ log (N) method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089.
41. Essmann, U., L. Perera, …, L. Pedersen. 1995. A smooth particle mesh Ewald method. *J. Chem. Phys.* 103:8577–8593.
42. Berendsen, H., J. Postma, …, J. Haak. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81:3684–3690.
43. Yan, Q., and J. de Pablo. 1999. Hyper-parallel tempering Monte Carlo: application to the Lennard-Jones fluid and the restricted primitive model. *J. Chem. Phys.* 111:9509.
44. Sugita, Y., A. Kitao, and Y. Okamoto. 2000. Multidimensional replica-exchange method for free-energy calculations. *J. Chem. Phys.* 113:6042–6051.
45. Yan, Q., and J. J. de Pablo. 2000. Hyperparallel tempering Monte Carlo simulations of polymeric systems. *J. Chem. Phys.* 113:1276–1282.
46. Faller, R., Q. Yan, and J. J. de Pablo. 2002. Multicanonical parallel tempering. *J. Chem. Phys.* 116:5419–5423.
47. Reddy, A. S., A. Izmitli, and J. J. de Pablo. 2009. Effect of trehalose on amyloid β (29–40)-membrane interaction. *J. Chem. Phys.* 131:085101.
48. Sugita, Y., and Y. Okamoto. 1999. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* 314:141–151.
49. Mitsutake, A., Y. Sugita, and Y. Okamoto. 2001. Generalized-ensemble algorithms for molecular simulations of biopolymers. *Biopolymers.* 60:96–123.
50. Rathore, N., M. Chopra, and J. J. de Pablo. 2005. Optimal allocation of replicas in parallel tempering simulations. *J. Chem. Phys.* 122:024111.
51. Kumar, S., J. Rosenberg, …, P. Kollman. 1992. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* 13:1011–1021.
52. Cornilescu, G., F. Delaglio, and A. Bax. 1999. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR.* 13:289–302.
53. Shen, Y., and A. Bax. 2007. Protein backbone chemical shifts predicted from searching a database for torsion angle and sequence homology. *J. Biomol. NMR.* 38:289–302.
54. Neal, S., A. M. Nip, …, D. S. Wishart. 2003. Rapid and accurate calculation of protein ^{1}H, ^{13}C and ^{15}N chemical shifts. *J. Biomol. NMR.* 26:215–240.
55. McQuarrie, D. A. 1976. Statistical Mechanics. Harper and Row, New York.
56. Auer, B. M., and J. L. Skinner. 2008. IR and Raman spectra of liquid water: theory and interpretation. *J. Chem. Phys.* 128:224511.
57. Mukherjee, P., A. T. Krummel, …, M. T. Zanni. 2004. Site-specific vibrational dynamics of the CD3ζ membrane peptide using heterodyned two-dimensional infrared photon echo spectroscopy. *J. Chem. Phys.* 120:10215–10224.
58. Torii, H., and M. Tasumi. 1992. Model calculations on the amide-I infrared bands of globular proteins. *J. Chem. Phys.* 96:3379–3387.
59. Gorbunov, R. D., P. H. Nguyen, …, G. Stock. 2007. Quantum-classical description of the amide I vibrational spectrum of trialanine. *J. Chem. Phys.* 126:054509.
60. Jansen, T. C., and J. Knoester. 2006. Nonadiabatic effects in the two-dimensional infrared spectra of peptides: application to alanine dipeptide. *J. Phys. Chem. B.* 110:22910–22916.
61. Sorin, E. J., and V. S. Pande. 2005. Exploring the helix-coil transition via all-atom equilibrium ensemble simulations. *Biophys. J.* 88:2472–2493.
62. Kabsch, W., and C. Sander. 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers.* 22:2577–2637.
63. Creighton, T. 1993. Proteins: Structures and Molecular Properties. WH Freeman, New York.
64. Jackson, M., P. I. Haris, and D. Chapman. 1989. Fourier transform infrared spectroscopic studies of lipids, polypeptides and proteins. *J. Mol. Struct.* 214:329–355.
65. Engel, M. F. M., H. Yigittop, …, J. Antoinette Killian. 2006. Islet amyloid polypeptide inserts into phospholipid monolayers as monomer. *J. Mol. Biol.* 356:783–789.
66. Strasfeld, D. B., Y. L. Ling, …, M. T. Zanni. 2008. Tracking fiber formation in human islet amyloid polypeptide with automated 2D-IR spectroscopy. *J. Am. Chem. Soc.* 130:6698–6699.
67. Ling, Y. L., D. B. Strasfeld, …, M. T. Zanni. 2009. Two-dimensional infrared spectroscopy provides evidence of an intermediate in the membrane-catalyzed assembly of diabetic amyloid. *J. Phys. Chem. B.* 113:2498–2505.