JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Impact of Thermostats on Folding and Aggregation Properties of Peptides Using the Optimized Potential for Efficient Structure Prediction Coarse-Grained Model

Yannick G. Spill,[†] Samuela Pasquali, and Philippe Derreumaux*

Laboratoire de Biochimie Théorique, UPR 9080 CNRS et Université Paris Diderot (Paris 7), Institut de Biologie Physico Chimique, 13 rue Pierre et Marie Curie, 75005 Paris, France

ABSTRACT: The simulation of amyloid fibril formation is impossible if one takes into account all chemical details of the amino acids and their detailed interactions with the solvent. We investigate the folding and aggregation of two model peptides using the optimized potential for efficient structure prediction (OPEP) coarse-grained model and replica exchange molecular dynamics (REMD) simulations coupled with either the Langevin or the Berendsen thermostat. For both the monomer of blocked penta-alanine and the trimer of the 25−35 fragment of the Alzheimer's amyloid β protein, we find little variations in the equilibrium structures and heat capacity curves using the two thermostats. Despite this high similarity, we detect significant differences in the populations of the dominant conformations at low temperatures, whereas the configurational distributions remain the same in proximity of the melting temperature. Aβ_{25-35} trimers at 300 K have an averaged β-sheet content of 12% and are primarily characterized by fully disordered peptides or a small curved two-stranded β-sheet stabilized by a disordered peptide. In addition, OPEP molecular dynamics simulations of Aβ_{25-35} hexamers at 300 K with a small curved six-stranded antiparallel β-sheet do not show any extension of the β-sheet content. These data support the idea that the mechanism of Aβ_{25-35} amyloid formation does not result from a high fraction of extended β-sheet-rich trimers and hexamers.

1. INTRODUCTION

When an interacting particle is subject to Brownian motion in a solvent, under the assumption that the radius of the particle is significantly larger than the radius of the solvent constituents, its motion can be modeled by the Langevin equation:[1]

$$m\frac{\mathrm{d}^2\vec{\mathrm{X}}(t)}{\mathrm{d}t^2} = -\vec{\nabla}V(t) - m\gamma\frac{\mathrm{d}\vec{\mathrm{X}}(t)}{\mathrm{d}t} + \vec{\mathrm{R}}(t) \qquad (1)$$

$$\langle R_i(t)R_j(t')\rangle = \delta(t-t')\delta_{ij}6\gamma k_{\mathrm{B}}T_0 \text{ and } \langle \vec{\mathrm{R}}(t)\rangle = 0$$

The action of the solvent at a temperature T_0 is represented implicitly through the action of a friction and a random force. The friction force corresponds to the macroscopic decay in momentum when a particle moves through a viscous fluid colliding with the small fluid particles; the random force represents the microscopic shocks with the solvent that have the effect of changing the particle trajectory. These shocks are disordered and tend to "thermalize" the particle, meaning that at equilibrium, the velocity distribution is Gaussian, with zero mean and variance $\sigma^2 = k_{\mathrm{B}}T_0/m$. The parameter controlling the impact of viscosity on the dynamics is the friction constant γ, which can also be interpreted as the collision frequency between the particle and the solvent. Two limiting cases occur. When γ equals zero, the particle follows a Newtonian dynamic and evolves in the microcanonical ensemble. When γ tends to infinity, Newtonian forces are negligible compared to the Langevin forces, and the particle follows a purely diffusive Brownian dynamic. For all cases in between, the system is in the canonical ensemble, at constant temperature T_0.

Although originally postulated to describe the Brownian motion of a particle, this equation can be easily implemented to act as a heat bath in molecular dynamics (MD) simulations.[2] It can be shown that this formulation allows one to generate a canonical distribution of states.[3] In the context of molecular dynamics, a Langevin thermal bath has the disadvantage that the coupling of the system to the bath is not only global but also local due to the random shocks. For large values of γ, the local shocks between the particle and the solvent can become an important disturbance to the system's dynamics. Ideally one would want to control the temperature through a global coupling only in order to minimize the local disturbance. To overcome this limitation, a new thermostat was introduced by Berendsen et al.[4] The global Langevin coupling is maintained, while an integration over fast degrees of freedom smooths out local collisions. The modified equation of motion is

$$m\frac{\mathrm{d}^2\vec{\mathrm{X}}(t)}{\mathrm{d}t^2} = -\vec{\nabla}V - m\frac{1}{2\tau}\left(\frac{T_0}{T}-1\right)\frac{\mathrm{d}\vec{\mathrm{X}}(t)}{\mathrm{d}t} \qquad (2)$$

where $\tau = 1/(2\gamma)$ and γ is the Langevin friction constant. In practice one sets the Berendsen time constant $\tau = 1/(2\gamma)$ rather than γ itself. Thermodynamics governed by a Berendsen thermostat depart from the canonical distribution for finite and nonzero values of τ. The statistical distribution followed by a system obeying eq 2 has been rigorously determined only recently.[5] It reduces to the canonical distribution for $\tau = 0$ and to the microcanonical distribution for $\tau = +\infty$.

Received: October 29, 2010
Published: April 11, 2011

Several computational studies have already discussed the effect of the Berendsen thermostat in MD and replica exchange molecular dynamics (REMD) simulations. The systems studied included a single butane molecule,[6] bulk water alone or in the presence of a penta-alanine,[7] simulated by use of a fully atomistic representation, and the 56-residue SH3 protein with tails of various lengths, simulated by use of a single C_α representation and the native structure-based Go potential.[8] These studies indicate that the Berendsen thermostat produces a noncanonical phase-space distribution, but the magnitude of the deviation is system-dependent: small for bulk water and larger for a peptide model using AMBER/TIP3P.[7] UNRES MD simulations of two model α-helical systems with the Berendsen and Langevin thermostats showed also differences in the folding dynamics, with the presence of explicit friction forces slowing down the folding.[9]

Unless a worldwide network of computers[10] or a specially built supercomputer[11] is used, atomic-level characterization of protein folding and aggregation in explicit solvent is limited to short time scales.[12] To go beyond this limitation, one solution is to reduce the number of degrees of freedon by resorting to simpler protein representations. The implicit solvent optimized potential for efficient structure prediction (OPEP) coarse-grained force field has been recently used to predict the 3D structures of peptides in their monomeric states via a greedy approach[13] and the early formed oligomers of various amyloid peptides by use of REMD and the Berendsen thermostat.[14,15]

Our aim is to investigate the influence of the Langevin and Berendsen thermostats on both folding and aggregation properties of peptides using the OPEP coarse-grained force field. To this end, we performed long REMD simulations on the penta-alanine peptide and the trimer of Alzheimer's $A\beta_{25-35}$ peptide. REMD is one generalized ensemble method that goes beyond conventional MD to accelerate convergence to equilibrium[16,17] and allows one to extract thermodynamic information. The penta-alanine model enables a comparison with the atomistic simulations in explicit solvent carried out by Rosta et al.[7] $A\beta_{25-35}$, due to its small size, is a convenient model system for studying the formation of $A\beta$ amyloids. It is known experimentally that this molecule can readily form β-sheet aggregates that are highly toxic to neurons.[18] Simulations addressing its dimerization in explicit water have been recently performed.[19,20] In contrast to other $A\beta$ fragments such as $A\beta_{16-22}$[21,22] or $A\beta_{29-42}$[23,24], considered as toy models for computational studies of amyloid formation, $A\beta_{25-35}$ has been identified as a physiological proxy with mutiple effects in neuronal intracellular components including plasma membranes, mitochondria, and cytosol, and its use has shown many of the same biochemical changes in animal models as those detected in Alzheimer's disease patients.[25] To enlarge the part of $A\beta_{25-35}$ amyloid dynamics and stability, we also carried out OPEP-MD simulations of hexamers built from structured trimer conformations.

2. MATERIALS AND METHODS

2.1. General Setup. REMD simulations[26] were performed with the coarse-grained OPEP force field.[27,28] In this model, the backbone N, H, C_α, C ,and O atoms are represented explicitly, while the side chains of all amino acids (except for the proline's heavy atoms) are represented by a unique bead. The interaction parameters include bond lengths, bond angles, improper torsions, dihedral angles, van der Waals interactions, and two- and four-body interactions for the backbone hydrogen bonds.[28] By removing numerous degrees of freedom, the OPEP force field allows significant acceleration with realistic sampling of conformational space.[29]

Table 1. Summary of Simulations Performed

starting config[a]	method	temp (K)	thermostat (τ or γ)	time (ns)
		Penta-alanine		
E	REMD	200−350	B (0.5 ps)	300 × 12
R	REMD	200−350	B (1 ps)	300 × 12
R	REMD	200−350	L (1 ps^{-1})	300 × 12
R	REMD	200−350	L (0.5 ps^{-1})	300 × 12
		$A\beta_{25-35}$ Trimers		
R	REMD	250−500	B (0.5 ps)	1000 × 24
R	REMD	250−500	L (1 ps^{-1})	1000 × 24
P	MD	300	B (0.5 ps)	100
P	MD	300	L (1 ps^{-1})	100
		$A\beta_{25-35}$ Hexamers		
AP_1	MD	300	B (0.5 ps)	100
AP_1	MD	300	L (1 ps^{-1})	100
AP_2	MD	300	B (0.5 ps)	100
AP_2	MD	300	L (1 ps^{-1})	100

[a] The starting conformation is either extended (E), random (R), parallel (P), or antiparallel (AP).

The Langevin thermostat was implemented in the OPEP-MD program by scaling each velocity:

$$v_{ij} \rightarrow e^{-\gamma\delta t} v_{ij} + \sqrt{(1 - e^{-\gamma\delta t})(1 + e^{-\gamma\delta t})} \sqrt{\frac{k_B T_0}{m_i}}\, G(t) \quad (3)$$

where v_{ij} is the jth component of the velocity of particle i and m_i is its mass; γ is the Langevin collision frequency constant; δt is the integration time step; k_B is Boltzmann's constant; T_0 is the target temperature; and $G(t)$ is a random Gaussian number with zero mean and unit variance. A typical value for γ is, for example, 50 ps^{-1}, thought to reproduce best the dynamics in water,[30] but values as low as 0.15 ps^{-1} have been used for amyloids with an implicit solvent model.[31]

Velocity rescaling for the Berendsen thermostat is[29]

$$v_{ij} \rightarrow \sqrt{\left(1 - \frac{\delta t}{\tau}\right)\left(1 - \frac{T_0}{T}\right)}\, v_{ij} \quad (4)$$

where T is the instantaneous temperature and τ is the Berendsen coupling constant. A typical value of τ suited to study aggregation is τ = 0.5 ps (see ref 15), as it simultaneously allows for a quick thermalization while not overconstraining the system.

2.2. Simulations. A summary of the setup details of all simulations is given in Table 1. For each system, we describe the method used (REMD or MD), the initial configuration, the temperatures and thermostats used (Berendsen or Langevin and the corresponding values of τ or γ), and the total simulation time.

The first system studied is the penta-alanine blocked by acetyl and N-methyl groups. An exponential temperature distribution of the replicas assured fast convergence. The temperature range varying between 200 and 350 K spans across the folded and the unfolded states: 200.0, 210.4, 221.4, 233, 245, 257.9, 271.4, 285.6, 300.5, 316.1, 332.6, and 350.0 K. For each thermostat, two 300 ns REMD simulations were launched with 12 replicas,

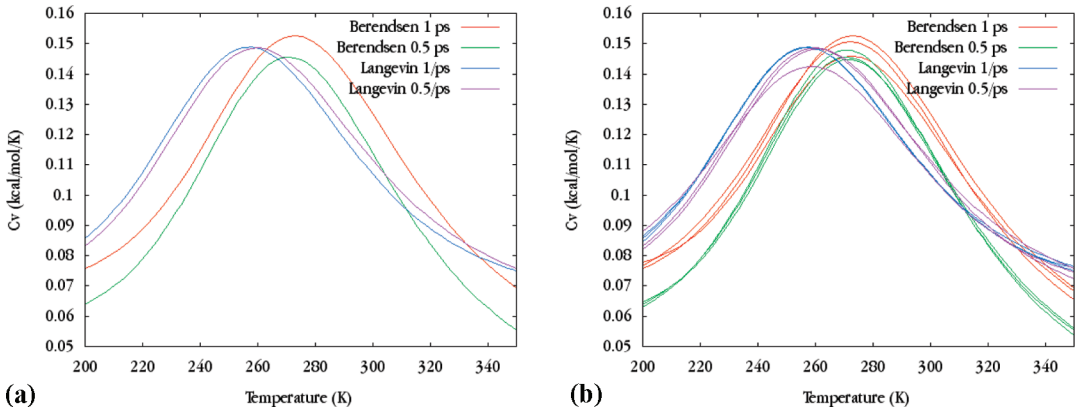


Figure 1. Plot of the heat capacity of the four penta-alanine simulations: (a) full 10−300 ns time interval; (b) 10−300, 10−150, and 150−300 ns intervals for each thermostat and coupling constant.

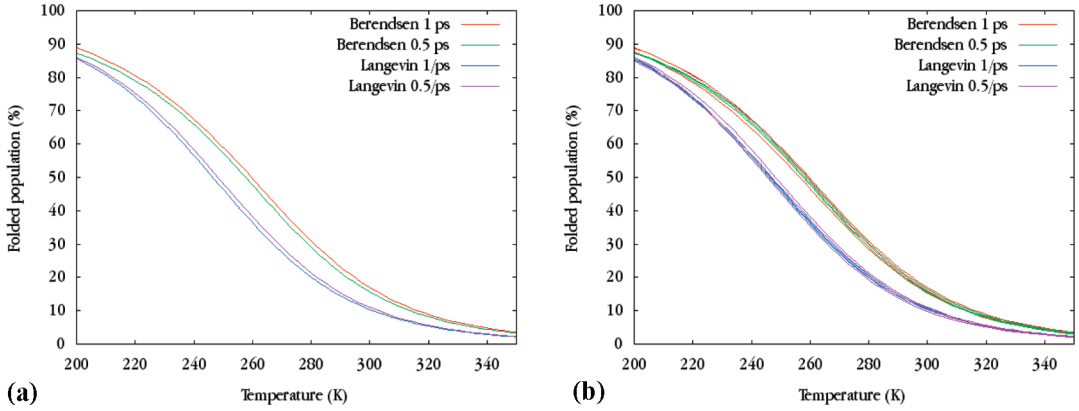


Figure 2. Populations of the folded state for four different penta-alanine simulations. The capped penta-alanine was considered to be folded if it had at least two $i, i+3$ or $i, i+4$ hydrogen bonds. (a) Full 10−300 ns time interval; (b) 10−300, 10−150, and 150−300 ns intervals for each thermostat and coupling constant.

with the following coupling constants: Berendsen τ, 0.5 ps and 1 ps; Langevin γ, 1 ps^{-1} and 0.5 ps^{-1}. This choice corresponds to two series of simulations with reciprocal time constants with respect to the equality $\tau = 1/(2\gamma)$. The initial configuration was an extended configuration.

The second system is a trimer of the hydrophilic−hydrophobic Aβ_{25-35} peptide, of sequence acetyl-Gly-Ser-Asn-Lys-Gly-Ala-Ile-Ile-Gly-Leu-Met-NH_2. Since it is a homotrimer, five different topologies are expected in principle:

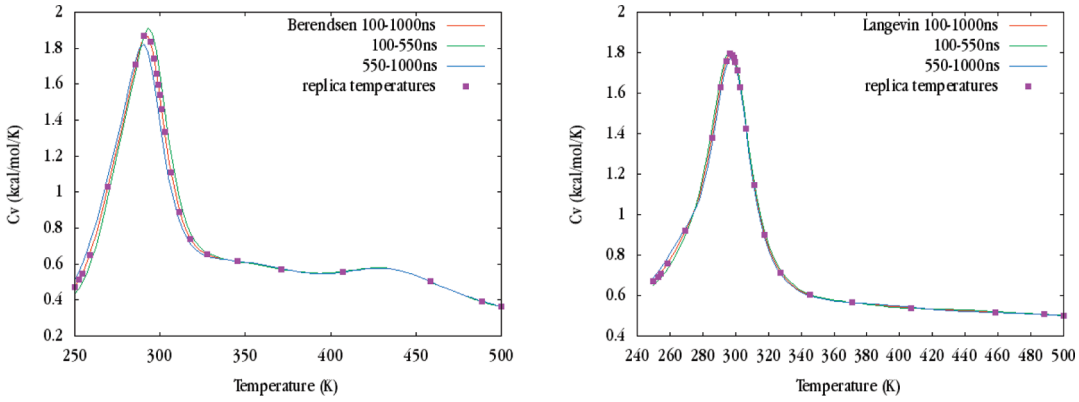
- antiparallel (AP): the three chains form a β-sheet where neighboring peptides are oriented opposite to each other
- mixed (Mix): the three chains form a β-sheet where two neighboring peptides are oriented parallel, while the third one is oriented in the opposite direction
- parallel (P): the three chains form a β-sheet with all peptides oriented in the same direction
- partially folded (2 + 1): two peptides form a β-sheet while the third one is random coil
- coil: all three peptides are random coil

Two REMD runs were performed, one with Berendsen $\tau = 0.5$ ps and the other with Langevin $\gamma = 1$ ps^{-1}. For each run, 24 replicas were used with the same set of temperatures: 250, 252.69, 254.48, 258.66, 269.33, 285.64, 290.74, 294.25, 296.55, 297.99, 298.96, 299.83, 300.96, 302.81 306.45, 311.48, 317.74, 327.46, 345.32, 370.95, 407.15, 458.41, 488.43, and 500.00 K. These temperatures were determined from the configurations obtained by a preliminary REMD simulation of 500 ns with 18 replicas starting from fully disordered peptides separated from each others by 15 Å via the procedure presented in ref 32. This choice corresponds to a replica flux-optimized temperature distribution for our system, as described in refs 17, 33, and 34. It is of interest to note that equilibration was not achieved within 500 ns in the preliminary REMD.

Each REMD simulation was performed for 1 μs per replica, starting with conformations belonging to the antiparallel, mixed, 2 + 1, and coil basins and obtained through the preliminary simulation. The integration time step was 1.5 fs, a sphere of 80 Å with reflecting boundary conditions was used, and energy and structure snapshots were taken every 1.5 ps. Replica exchanges were attempted every 6 ps, a time interval highly used and tested.[26,35] Two MD simulations of 100 ns at 300 K were also carried out on the trimer starting from a fully extended parallel β-sheet.

The third system is a hexamer of Aβ_{25-35} consisting of a single-layer antiparallel small β-sheet configuration. Four MD runs of 100 ns at 300 K were performed, two with Berendsen $\tau =$ 0.5 ps (B_1 and B_2) and two with Langevin $\gamma = 1$ ps^{-1} (L_1 and L_2). The runs Bi and Li use the same structure, and the rmsd deviation between the two starting conformations is 10 Å.

2.3. 2D and 3D Structure Analysis. For each system, the secondary structure of all conformations was determined. For the



(a) Berendsen simulation (b) Langevin simulation

Figure 3. Block analysis of the Aβ_{25-35} trimer heat capacity for each thermostat.

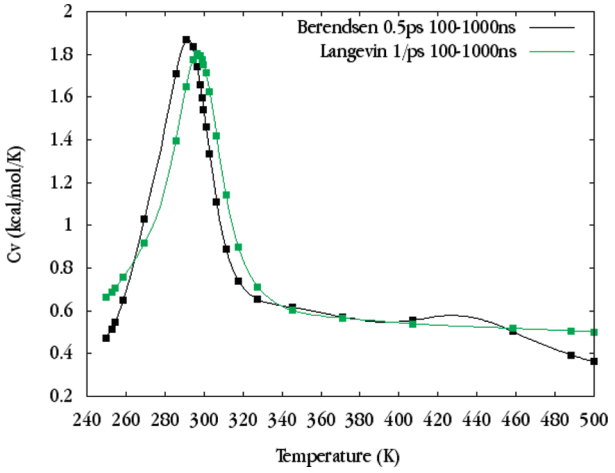


Figure 4. Heat capacities of the Aβ_{25-35} trimer with Langevin and Berendsen thermostats.

Aβ_{25-35} trimer, prior to cluster analysis based on C_α rmsd, we determined the topology of each configuration by using a combination of three metrics that includes vector cosines (where we compute the angles between the vectors representing the direction of each peptide), number of hydrogen bonds formed, and distance between the peptides. We also analyzed the transition times between the Mix and AP configurations and defined that a conformational change from Mix to AP (or vice versa) has occurred if, before and after the transition, the trimer explores both topologies for at least 100 ps. For the Aβ_{25-35} hexamer, we followed the C_α rmsd as a function of time using the core β-sheet region.

2.4. Thermodynamic Quantities. At the end of an REMD simulation, thermodynamic quantities can be extracted. Even though in principle these could be obtained from their definitions with respect to the independent variables of the system, in practice this direct evaluation leads, sometimes, to very large error bars in the estimates and alternative derivations turn out to be more convenient. This is the case for the specific heat C_V. With both thermostats, we could obtain C_V by computing the potential energy derivative with respect to the temperature, but different strategies give better results. For the Langevin thermostat, where the system obeys the canonical distribution, the heat capacity can be calculated from the potential energy as

$$C_V(T) = \frac{\langle (E - \langle E \rangle)^2 \rangle}{k_B T^2} \quad (5)$$

For the weak-coupling Berendsen thermostat, where the distribution is noncanonical, the appropriate formula has been shown to be[5]

$$C_V(T) = \frac{k_B \langle (\delta\Phi)^2 \rangle}{(k_B T)^2 - \sqrt{\langle (\delta\Phi)^2 \rangle \langle (\delta K)^2 \rangle} 2/(3N)} \quad (6)$$

where $\langle (\delta\Phi)^2 \rangle$ and $\langle (\delta K)^2 \rangle$ represent the variance of potential and kinetic energy, respectively, and broken brackets denote a time average.

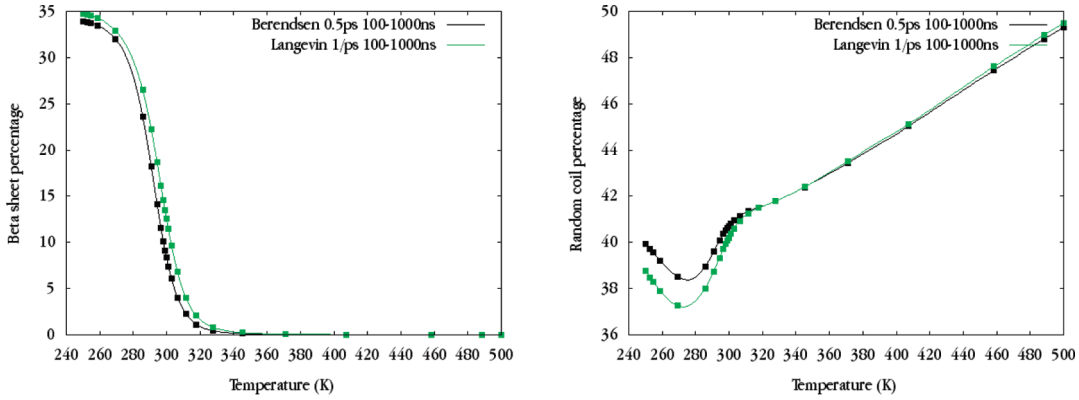
Thermodynamic quantities are meaningful only if computed after the full convergence of the system has been established with respect to the particular quantity in question. Different quantities have different convergence times. We used a block analysis scheme to assess convergence over a given time frame. The quantity is evaluated once over the entire time interval, a second time over the first half of the interval, and a third time over the second half. The three curves are then superposed and the convergence is reached when the curves coincide within a preset error.

In practice, to extract thermodynamical quantities, histogram-based methods like WHAM are typically used.[36] We preferred implementing our own script based on the recently introduced MBAR method,[37] which has the conceptual advantage of producing curves accompanied by error bars and the technical advantage of reducing the computing time by avoiding some of the slow histogram calculations.

3. RESULTS

3.1. Penta-alanine. The first 10 ns of each REMD simulation was discarded, and the remaining 290 ns was analyzed. The acceptance rate was always higher than 30% in all four simulations, reaching its minimal values for the lowest and highest replicas.

When the heat capacity curves obtained from the two thermostats with reciprocal values of τ and γ are compared (Figure 1), the position of the peak (T_m) changes from 260 K for Langevin to 273 K for Berendsen. Compared to the Berendsen thermostat,



(a) β-sheet % (b) Random coil %

Figure 5. (a) β-Sheet and (b) random coil percentages in the two REMD simulations of Aβ_{25-35} trimer.

Table 2. Structural and Energetic Contents of the Two Aβ_{25-35} Trimer Simulations at the Lowest Temperature[a]

	Berendsen		Langevin	
class	%	energy (kcal/mol)	%	energy (kcal/mol)
AP	71.6	-134.1 ± 6.3	56.3	-133.8 ± 8.4
Mix	24.2	-130.6 ± 6.9	37.7	-131.4 ± 8.9
P	0.0		0.0	
2 + 1	4.3	-119.9 ± 6.7	5.9	-118.7 ± 8.9
coil	0.0		0.0	

[a] Obtained by classifying every tenth frame of the converged period of each simulation, totalling 60 000 structures. The population of coil conformations is too small to get accurate statistics at this temperature.

the Langevin thermostat shifts the melting temperature 13 K toward lower temperature.

If the coupling constant is modified, the position of T_m and the full width at half-maximum of the peak (fwhm) remain unchanged in the Langevin simulations, whereas in Berendsen simulations T_m remains constant while fwhm increases by 20% $\pm$ 5%. This variation in fwhm is larger than the error bars obtained from block analysis.

A recent all-atom REMD simulation on the same penta-alanine system in explicit solvent reported that, for a weak-coupling Berendsen thermostat, the folded state is overpopulated by about 10% at low temperatures and underpopulated at high temperatures.[7] To verify this hypothesis, we compared the population distribution as a function of temperature, obtained with both Langevin and Berendsen thermostats. In our simulations we define the folded state by a criterion in terms of the hydrogen-bond (H-bond) network. We considered a H-bond to be formed if the distance between donor and acceptor was smaller than 3 Å and the donor−H−acceptor angle was less than 60°. In all simulations, we found that H-bonds were established mainly between residues i and $i + 4$ (α-helix) and to a lesser extent between i and $i + 3$ (3_{10}-helix). We also observed a very small amount of $i + 5$ contacts (π-helix). We thus considered the peptide to be folded if it had two or more hydrogen bonds between residues three or four amino acids apart. The resulting folded populations are shown in Figure 2.

When we focus on the temperature range 290−340 K, our simulations lead to a small shift in the folded populations varying between 5% and 1% between the two thermostats, with an increased population for the Berendsen thermostat. At lower temperatures, the difference in the folded population slightly increases but never exceeds 10%. This overpopulation of the folded (helical) state at low temperatures with the weak Berendsen coupling constant is fully consistent with what was observed by Rosta et al.[7] using all-atom REMD simulations. In contrast to Rosta's results at 350 K, where the fraction of folded state is lowered by about 10% for Berendsen compared to Langevin but remains around 0.6 at 350 K, both thermostats with OPEP give a negligible population of folded state at 350 K.

3.2. Aβ_{25-35} Trimer. Two REMD simulations were performed on the Aβ_{25-35} trimer, with reciprocal Langevin (γ = 1 ps^{-1}) and Berendsen (τ = 0.5 ps) thermostats starting with 24 conformations from different energy basins. From a block analysis of the potential energies and C_V values, the first 100 ns were discarded for each replica. Figure 3 shows the convergence of each simulation by comparing their heat capacity profiles over two time intervals, 100−550 and 550−1000 ns. The conformational properties are therefore based on 21.6 μs for each thermostat.

In Figures 4 and 5, we show the heat capacity and the percentages of β-sheet and random coil as determined by the STRIDE program[38] as a function of temperature. Overall, the characteristic transitions are shifted by approximately 5 K between the two thermostats, as assessed by the peak of heat capacity (297 K for Langevin vs 291 K for Berendsen) and the inflection point of β-sheet percentage (296 K for Langevin and 292 K for Berendsen). Random coil profiles also superpose well with a maximum difference of 1% at low temperatures. We note, however, that the excess heat capacity reaches slightly lower values in Berendsen simulation than in Langevin simulation at low (below 260 K) and high (>460 K) temperatures.

Of the five possible topologies of the Aβ_{25-35} trimer, all except the parallel (P) three-stranded β-sheet have been detected in each one of the REMD simulations. It should be noted that two MD simulations starting from the fully parallel geometry evolve toward the 2 + 1 topology within 100 ns. Table 2 reports the structural and energetic content of the REMD-generated topologies at the lowest temperature (250 K). As can be seen, the antiparallel conformation (AP) is the most populated, followed by the mixed (Mix) and 2 + 1 topologies in both simulations. Looking at the three populated AP, Mix, and 2 + 1 topologies representing 99.99% of the conformations, we find that the averaged

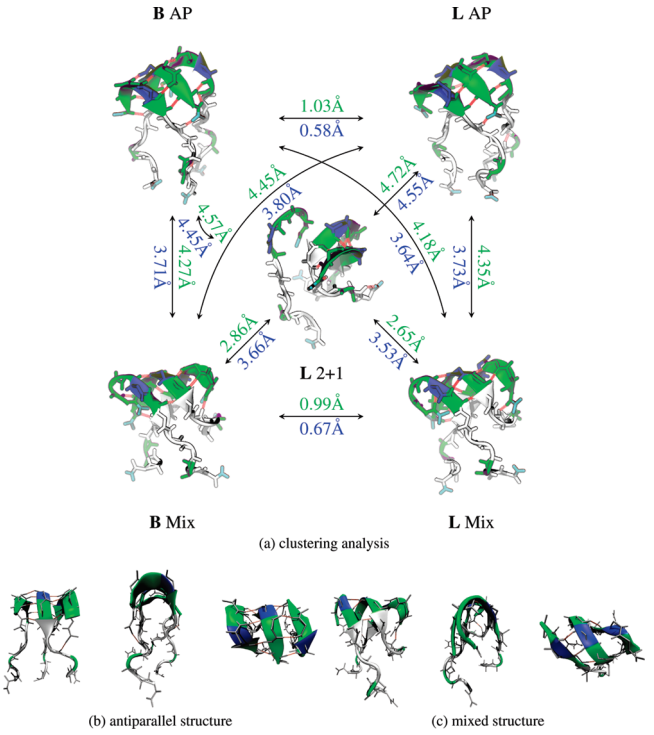


Figure 6. Main structures encountered in the REMD simulations of Aβ_{25-35} trimer at the lowest temperature. (a) Cluster analysis (images were generated by PyMOL).[39] **B** AP, first cluster of the Berendsen simulation (65% population); **L** AP, first cluster of the Langevin simulation (57%); **B** Mix, second cluster of the Berendsen simulation (16%); **L** Mix, second cluster of the Langevin simulation (16%); **L** 2 + 1, an example of a partially unfolded 2 + 1 topology (fifth cluster of Langevin simulation, 1% population). White, hydrophobic residues; green, hydrophilic residues; blue, polar residues; yellow, H-bonds. Numbers indicated are root-mean-square deviations (rmsds) in angstroms: numbers in green, rmsd on the whole backbone; numbers in blue, rmsds on the structured part only (amino acids 26−30). (b, c) Three different view angles of the (b) antiparallel and (c) mixed structures.

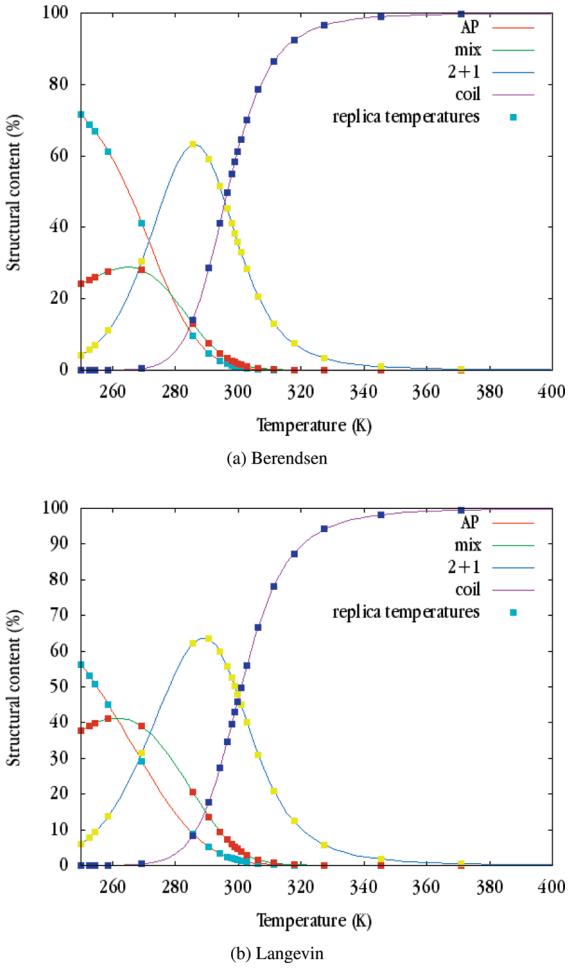


Figure 7. Topological content percentage analyzed over the converged part of each Aβ trimer simulation (60 000 structures): (a) Berendsen and (b) Langevin simulations.

potential energies for both thermostats are very similar. For AP topology, we find −133.8 kcal/mol for Langevin versus −134.1 kcal/mol for Berendsen, and for 2 + 1 topology, we find −118.7 kcal/mol for Langevin versus −119.9 kcal/mol for Berendsen. However, the Berendsen simulation displays a narrower energy distribution than the Langevin simulation for each topology, with a difference in the root-mean-square fluctuation of the potential energy amounting to 2.0 kcal/mol.

The most populated clusters of each simulation at the lowest temperature are represented in Figure 6. Structural differences between the configurations sampled by the thermostats are very subtle, as reported by the cross rms deviations between the centers of the clusters. For instance, the mixed topologies or the antiparallel topologies in the two simulations deviate by less than 1 Å. Similar rms deviations are found at higher temperatures (data not shown).

We analyzed our trajectories in search of conformational events from the antiparallel to mixed topologies or vice versa. These events always involved a transient coil or partially folded (2 + 1) topology. Looking at each replica below the melting temperatures, we found the following statistics. For the Langevin simulation, we counted four events from Mix to AP, with an averaged $\tau_{Mix\rightarrow AP}$ of 268 ns, and five events from AP to Mix, with an averaged $\tau_{AP\rightarrow Mix}$ of 111 ns. For the Berendsen simulation we observed the same number of transitions as in Langevin, $\tau_{Mix\rightarrow AP}$ = 199 ns and $\tau_{AP\rightarrow Mix}$ = 89 ns. The two simulations thus behave very similarly. It should be noted that the interconversion between Mix and AP three-stranded β-sheets is slow despite the use of REMD and having optimized the replica temperatures distribution to ensure adequate exchange rates. This suggests the existence of significant potential and free energy barriers, as found also in other systems such as dimers of the Aβ_{16-22}[21] and GNNQQNY peptides.[40]

A final comparison between the thermostats is reported in Figure 7, which compares the percentage of topological content generated by each simulation as a function of temperature. Although the distribution of each topology looks very similar for both thermostats, there is one striking difference. At low temperatures (<269 K), while the difference in the populations of the 2 + 1 topology is small (<1%), there are more antiparallel structures in the Berendsen simulation than in the Langevin simulation (see also Table 2). For instance, at 259 K (replica 4), 63% of the structures are antiparallel for Langevin and 46% for Berendsen. The reverse is true for the mixed state, where we observe a population of 26% for Berendsen against 41% for Langevin.

Around the melting temperatures, that is, between 280 and 340 K, the populations of the most ordered AP and mixed topologies

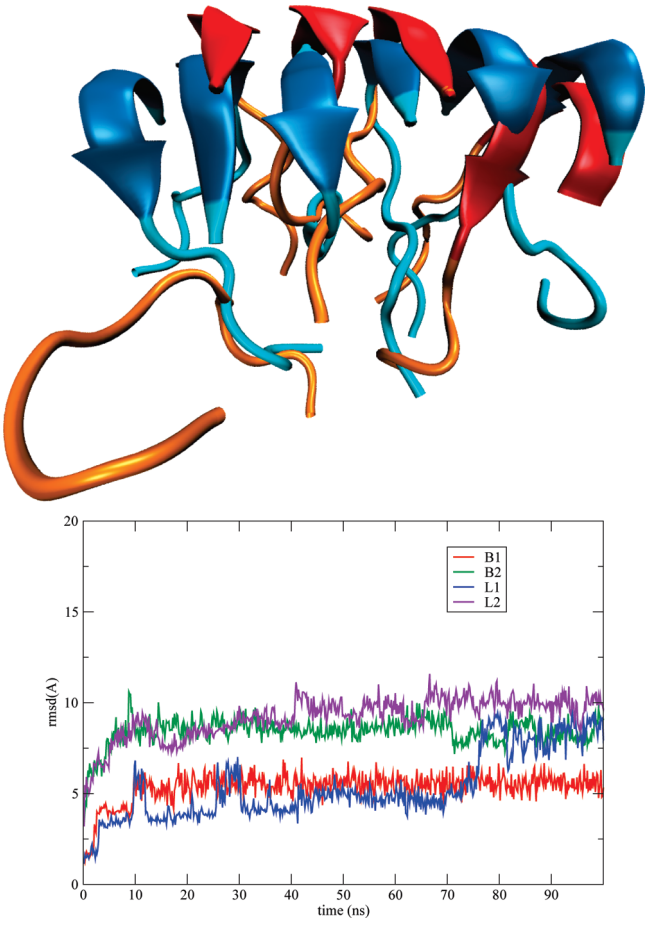


Figure 8. Aβ_{25-35} hexamer. (Top) Superposition of the initial structure on that obtained by Berendsen after 100 ns (image generated by VMD).[50]. The rmsd is 6.35 Å for the core. (Bottom) Time evolution of the Cα rmsd of all chains for residues 26−30 at 300 K, that is, a few degrees above the melting temperature predicted by Langevin and Berendsen REMD simulations.

vary little between the two thermostats. For instance, at 291 K (replica 9), the AP populations are 5.9% in one case and 6.2% in the other. The peak of the 2 + 1 topology is shifted 3 K toward higher temperatures for Langevin than Berendsen, and its height is the same in both simulations. However, its distibution is larger, with a full width at half-maximum of 36.4 K for Langevin versus 31.3 K for Berendsen. Similarly, the coil curve is shifted 3 K to the right for the Langevin simulation compared to Berendsen.

3.3. Aβ_{25-35} Hexamer. It is instructive to enlarge the part of the amyloid dynamics and stability by examining oligomers higher than the trimer. While changes in electrical fields associated with membranes and the presence of metals can play a role in Alzheimer's disease and are studied by standard or Car−Parrinello-type MD simulations,[41,42] we limit ourselves to mainly Aβ oligomers in aqueous solution.

Ma and Nussinov[43] studied the stability of Aβ_{25-35} octamers consisting of two β-sheets using short (5 ns) all-atom MD simulations at 330 K. Recent all-atom MD simulations by Shea and co-workers[19] showed that a V-shaped protofibril structure consisting of six Aβ_{25-35} peptides was stable at 310 K for 55 ns. Röhrig et al.[44] studied the stability of oligomers of Aβ_{16-22} from the dimer to the 32-mer. In all-atom MD simulations of ∼30 ns at 300 and 348 K, a single-layer β-sheet of eight peptides was not stable in contrast to a two-layered octamer β-sheet, suggesting that the minimum nucleus size is on the order of 8−16 Aβ_{16-22} peptides.

It is well-established, however, that all-atom 100-ns MD simulations, while allowing one to study the stability of preformed oligomers, do not sample equilibrium structures, and one must resort to enhanced conformational technique and/or simplified protein−water representations. All-atom REMD simulations in explicit solvent showed that seven β2 m(83−89) peptides are in equilibrium between numerous topologies.[45] From different coarse-grained models, simulations pointed to the complexity of the free energy landscape of Aβ_{16-22} 6-mers and 7-mers.[46,47] Similarly, Masman et al.[48] found that pentamers of Aβ_{1-42} with fibril geometries remain stable by all-atom MD for 100 ns at 310 K, while Urbanc et al.[49] found disordered pentamers for the same system using very long discrete MD simulations.

Here, we constructed a hexamer of Aβ_{25-35} consisting of a single-layer antiparallel β-sheet configuration based on the antiparallel structure we found for the trimer, that is, with a β-sheet core spanning amino acids 26−30, and performed four MD runs of 100 ns at 300 K. Our goal is not to explore the full configurational space but rather to determine the differences between Langevin and Berendsen OPEP-MD simulations on a reasonable time scale. As can observed in Figure 8, the L_1 and B_1 runs lead to 5.2 ± 1.8 and 5.3 ± 0.8 Å rmsd, respectively, on the entire trajectory, while the L_2 and B_2 runs lead to 7.5 ± 1.2 and 6.7 ± 0.9 Å rmsd. This suggests (i) the same plasticity of the oligomers for both thermostats, characterized by the detachment of one or two external peptides from the β-sheet core, and (ii) the heterogeneity of the energy landscape, characterized by multiple isoenergetic conformations in dynamic equilibrium, consistent with many computational studies on other sequences.[31,46,47,49] As reported for the REMD simulations of the Aβ trimer, we observe similar averaged potential energies for the hexamer with both thermostats (−301 ± 15 kcal/mol for B_1 + B_2 vs −301 ± 20 kcal/mol for L_1 + L_2) and a narrower energy distribution for the Berendsen simulation.

4. CONCLUSIONS

In this study, we have determined the impact of Langevin and Berendsen thermostats on folding properties of penta-alanine and aggregation properties of the trimer of the 25−35 fragment of the Alzheimer's amyloid β protein with the OPEP coarse-grained model. Using long REMD simulations, we find small variations in the heat capacity curves for the two thermostats. There is, however, a small distortion in thermodynamic descriptions of the two systems with the Berendsen thermostat at low temperatures. While the structural contents of the folded state for the penta-alanine peptides and of the topologies for the Aβ_{25-35} trimer remain the same (rms deviations of less than 1 Å), their populations can vary by 15%. This finding is fully consistent with previous reports[7] and the physics behind this variation is the same: narrowed potential energy fluctuations modify the relative populations of the configurations. At higher temperatures, and precisely above and around the melting temperatures, REMD simulations with the Berendsen thermostat result in small effects. We can thus conclude that the aggregation properties of all the amyloids, with the OPEP force field and the Berendsen thermostat,[14,15] would be unaffected around their melting temperatures.

It is of interest to examine the results of the Aβ_{25-35} trimer simulations, since little is known about their structures experimentally, as they are transient. Contrary to other fragments of the

full-length Aβ_{1-42} protein, the Aβ_{25-35} peptide has been subject to a small number of MD and REMD simulations. Wei et al.[19] and Kittner and Knecht[20] studied its dimerization in explicit water. Ma and Nussinov[43] studied the impact of N27Q substitution on the stability of preformed β-sheets, and Shea and coworkers[19] studied the stability of protofibrils with various topologies. Fu et al.[51] investigated the initial adsorption features and dynamics of Aβ_{25-35} on a single-walled carbon nanotube surface using MD in explicit solvent. Recently, Yu et al.[52] reported a hybrid computational approach to construct, search, optimize, and rank soluble micellelike Aβ_{25-35} structures with different side-chain packings at the atomic level.

Equilibration of Aβ_{25-35} trimers, as measured by the convergence of the heat capacity, is a very difficult task with the OPEP force field and REMD simulations. A replica flux-optimized temperature distribution and a larger number of replicas contribute to sampling efficiency. Our converged simulations, which total 48 μs in length, show that trimers at 300 K have an averaged β-sheet content of 12%. This content is very similar to the values found at 310 K in the monomer (13%) and the dimer (20%) by atomistic simulations.[19] The trimer configurational ensemble is primarily characterized by fully random coils or a small curved two-stranded β-sheet stabilized by a disordered peptide. Approximately 10% of the conformations consist, however, of curved small three-stranded β-sheets spanning the Ser_{26}-Asn-Lys-Gly-Ala_{30} amino acids with mixed or fully antiparallel orientations of the chains. This impossibility to stabilize fully extended amyloid-like conformations has been also observed for dimers by Wei et al.[19] in explicit solvent. Whether the presence of straight-extended β-sheet-rich dimers and trimers in the assembly of Aβ_{25-35} peptides by use of an implicit water−atomistic protein model[53] results from finite-size effects[54] or is an artifact of the force field used remains to be determined, but the present OPEP-MD simulations of the hexamer do not reveal any extension of the small β-sheet core within 100 ns.

AUTHOR INFORMATION

Corresponding Author

*E-mail: philippe.derreumaux@ibpc.fr.

Present Addresses

†Unité de Bioinformatique Structurale, Institut Pasteur, 25 rue du Docteur Roux, 75015 Paris, France.

ACKNOWLEDGMENT

We thank Massimiliamo Bonomi and Alessandro Barducci (ETH Zurich) for useful discussions on the Langevin thermostat. Y.G.S. thanks Michael Nilges for his complementary computational support, John Chodera for his help in interfacing the MBAR library, and École Normale Supérieure for financial support. P.D. thanks CNRS and Institut Universitaire de France (IUF) for financial support.

REFERENCES

(1) Langevin, P. *C. R. Acad. Sci.* **1908**, *146*, 530–532.

(2) Schlick, T. *Molecular Modeling and Simulation: An Interdisciplinary Guide*; Interdisciplinary Applied Mathematics; Elsevier: Amsterdam, 2002; pp 435−440.

(3) McQuarrie, D. A. *Statistcal Mechanics*; Harper & Row: New York, 1976; pp 452−456.

(4) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(5) Morishita, T. *J. Chem. Phys.* **2000**, *113*, 2976–2982.

(6) D'Alessandro, M.; Tenenbaum, A.; Amadei, A. *J. Phys. Chem. B* **2002**, *106*, 5050–5057.

(7) Rosta, E.; Buchete, N.-V.; Hummer, G. *J. Chem. Theory Comput.* **2009**, *5*, 1393–1399.

(8) Mor, A.; Ziv, G.; Levy, Y. *J. Comput. Chem.* **2008**, *29*, 1992–1998.

(9) Khalili, M.; Liwo, A.; Jagielska, A.; Scheraga, H. A. *J. Phys. Chem. B* **2005**, *109*, 13798–13810.

(10) Snow, C. D.; Nguyen, H.; Pande, V. S.; Gruebele, M. *Nature* **2002**, *420*, 102–106.

(11) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; Wriggers, W. *Science* **2010**, *330*, 341–346.

(12) Huet, A.; Derreumaux, P. *Biophys. J.* **2006**, *91*, 3829–3840.

(13) Maupetit, J.; Derreumaux, P.; Tuffery, P. *Nucleic Acids Res.* **2009**, *37*, W498–W503.

(14) Melquiond, A.; Dong, X.; Mousseau, N.; Derreumaux, P. *Curr. Alzheimer Res.* **2008**, *5*, 244–250(7).

(15) Chebaro, Y.; Mousseau, N.; Derreumaux, P. *J. Phys. Chem. B* **2009**, *113*, 7668–7675.

(16) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141–151.

(17) Trebst, S.; Troyer, M.; Hansmann, U. H. E. *J. Chem. Phys.* **2006**, *124*, 174903.

(18) Mattson, M. P.; Begley, J. G.; Mark, R. J.; Furukawa, K. *Brain Res.* **1997**, *771*, 147–153.

(19) Wei, G.; Jewett, A. I.; Shea, J.-E. *Phys. Chem. Chem. Phys.* **2010**, *12*, 3622–3629.

(20) Kittner, M.; Knecht, V. *J. Phys. Chem. B* **2010**, *114*, 15288–15295.

(21) Santini, S.; Wei, G.; Mousseau, N.; Derreumaux, P. *Structure* **2004**, *12*, 1245–1255.

(22) Nguyen, P. H.; Li, M. S.; Stock, G.; Straub, J. E.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 111–116.

(23) Itoh, S. G.; Okamoto, Y. *J. Phys. Chem. B* **2008**, *112*, 2767–2770.

(24) Lu, Y.; Wei, G.; Derreumaux, P. *J. Phys. Chem. B* **2011**, *115*, 1282–1288.

(25) Kaminsky, Y. G.; Marlatt, M. W.; Smith, M. A.; Kosenko, E. A. *Exp. Neurol.* **2010**, *221*, 26–37.

(26) Chebaro, Y.; Dong, X.; Laghaei, R.; Derreumaux, P.; Mousseau, N. *J. Phys. Chem. B* **2009**, *113*, 267–274.

(27) Forcellino, F.; Derreumaux, P. *Proteins: Struct., Funct., Bioinf.* **2001**, *45*, 159–166.

(28) Maupetit, J.; Tuffery, P.; Derreumaux, P. *Proteins: Struct., Funct., Bioinf.* **2007**, *69*, 394–408.

(29) Derreumaux, P.; Mousseau, N. *J. Chem. Phys.* **2007**, *126*, No. 025101.

(30) Pastor, R.; Brooks, B.; Szabo, A. *Mol. Phys.* **1988**, *65*, 1409–1419.

(31) Kim, S.; Takeda, T.; Klimov, D. K. *Biophys. J.* **2010**, *99*, 1949–1958.

(32) Nadler, W.; Meinke, J. H.; Hansmann, U. H. E. *Phys. Rev. E* **2008**, *78*, No. 061905.

(33) Katzgraber, H. G.; Trebst, S.; Huse, D. A.; Troyer, M. *J. Stat. Mech.* **2006**No. P03018.

(34) Nadler, W.; Hansmann, U. H. E. *Phys. Rev. E* **2007**, *75*, No. 026109.

(35) Sindhikara, D.; Meng, Y.; Roitberg, A. E. *J. Chem. Phys.* **2008**, *128*, No. 024103.

(36) Kumar, S.; Bouzida, D.; Swendsen, R.; Kollman, P.; Rosenberg, J. *J. Comput. Chem.* **1992**, *13*, 1011–1021.

(37) Shirts, M. R.; Chodera, J. D. *J. Chem. Phys.* **2008**, *129*, No. 124105.

(38) Frishman, D.; Argos, P. *Proteins: Struct., Funct., Bioinf.* **1995**, *23*, 566–579.

(39) PyMOL, version 1.3; Schrödinger, LLC.

(40) Strodel, B.; Whittleston, C. S.; Wales, D. J. *J. Am. Chem. Soc.* **2007**, *129*, 16005–16014.

(41) Lugli, F.; Toschi, F.; Biscarini, F.; Zerbetto, F. *J. Chem. Theory Comput.* **2010**, *6*, 3516–3526.
(42) Morante, S. *Curr. Alzheimer Res.* **2008**, *5*, 508–524.
(43) Ma, B.; Nussinov, R. *Biophys. J.* **2006**, *90*, 3365–3374.
(44) Röhrig, U. F.; Laio, A.; Tantalo, N.; Parrinello, M.; Petronzio, R. *Biophys. J.* **2006**, *91*, 3217–3229.
(45) Simone, A. D.; Derreumaux, P. *J. Chem. Phys.* **2010**, *132*, No. 165103.
(46) Lu, Y.; Derreumaux, P.; Guo, Z.; Mousseau, N.; Wei, G. *Proteins: Struct., Funct., Bioinf.* **2009**, *75*, 954–963.
(47) Irbäck, A.; Mitternacht, S. *Proteins: Struct., Funct., Bioinf.* **2008**, *71*, 207–214.
(48) Masman, M. F.; Eisel, U. L. M.; Csizmadia, I. G.; Penke, B.; Enriz, R. D.; Marrink, S. J.; Luiten, P. G. M. *J. Phys. Chem. B* **2009**, *113*, 11710–11719.
(49) Urbanc, B.; Betnel, M.; Cruz, L.; Bitan, G.; Teplow, D. B. *J. Am. Chem. Soc.* **2010**, *132*, 4266–4280.
(50) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.
(51) Fu, Z.; Luo, Y.; Derreumaux, P.; Wei, G. *Biophys. J.* **2009**, *97*, 1795–1803.
(52) Yu, X.; Wang, Q.; Zheng, J. *Biophys. J.* **2010**, *99*, 666–674.
(53) Cheon, M.; Chang, I.; Mohanty, S.; Luheshi, L. M.; Dobson, C. M.; Vendruscolo, M.; Favrin, G. *PLoS Comput. Biol.* **2007**, *3*, e173.
(54) Pawar, A.; Favrin, G. *PLoS ONE* **2008**, *3*, e2641.