Investigating the Mechanism of Peptide Aggregation: Insights from Mixed Monte Carlo-Molecular Dynamics Simulations

Massimiliano Meli, Giulia Morra, and Giorgio Colombo
Istituto di Chimica del Ricoscimento Molecolare, Consiglio Nazionale delle Ricerche, 20131 Milano, Italy

ABSTRACT The early stages of peptide aggregation are currently not accessible by experimental techniques at atomic resolution. In this article, we address this problem through the application of a mixed simulation scheme in which a preliminary coarse-grained Monte Carlo analysis of the free-energy landscape is used to identify representative conformations of the aggregates and subsequent all-atom molecular dynamics simulations are used to analyze in detail possible pathways for the stabilization of oligomers. This protocol was applied to systems consisting of multiple copies of the model peptide GNNQQNY, whose detailed structures in the aggregated state have been recently solved in another study. The analysis of the various trajectories provides dynamical and structural insight into the details of aggregation. In particular, the simulations suggest a hierarchical mechanism characterized by the initial formation of stable parallel β-sheet dimers and identify the formation of the polar zipper motif as a fundamental feature for the stabilization of initial oligomers. Simulation results are consistent with experimentally derived observations and provide an atomically detailed view of the putative initial stages of fibril formation.

INTRODUCTION

Molecular self-organization and self-assembly are processes by which Nature builds complex three-dimensional (3D) multicomponent structures with well-defined functions starting from simple building blocks such as oligonucleotides, lipids, or peptides. In the case of polypeptides, these phenomena have been the focus of intense research in recent years, after the realization that self-assembly is at the basis of the formation of oligomers and of insoluble amyloid fibrils (1–6). These involve the most diffused neurodegenerative disorders such as Alzheimer's and Parkinson's disease, Creutzfeld-Jakobs Syndrome, type II diabetes, and others. However, the ability to form ordered amyloid aggregates also was shown for a large number of nonpathogenic polypeptides under suitable solvent, temperature, and pH conditions (7–9). These observations provided a unique opportunity for the rational development of ordered supramolecular structures for biotechnology. The use of molecular self-assembly in biomedicine is exemplified by Ellis-Behnke et al. (10). The authors report that a designed self-assembling peptide nanofiber scaffold creates a permissive environment in an animal model for axons to regenerate through the site of an acute injury and to knit the brain tissues together. Smith et al. (11) showed that fibrils have strength comparable to that of steel and a mechanical stiffness comparable to that of silk, revealing that fibrils possess attractive properties for future technological applications. In general, fibrillar aggregates are characterized by a common cross-β-sheet structure with the β-strands perpendicular to the fiber axis, and their deposition is considered a hallmark of many types of human diseases (4).

Considering the key role that the peptide self-assembly process holds for vast scientific, biomedical, and technological communities, it is of interest to characterize not only the structural properties of the final structures but, most importantly, the early steps of oligomer formation at the atomic level. From the biomedical side, there is mounting evidence suggesting that the toxicity of the fatal neurodegenerative diseases may be caused by the intermediate oligomers in addition to the mature fibrils (12,13). On the biotechnological side, the study of aggregation mechanisms may be the source of inspiration for the development of ordered, rationally designed nanostructures with potentially interesting applications from material sciences to tissue engineering, from molecular electronics to drug delivery (14–16).

The high degree of complexity of aggregation, combined with the fact that initial oligomers are metastable and short-lived, makes experimental data difficult to obtain. In contrast, this regime appears to be ideally suited for computer simulations. For example, the aggregation of small oligomers of amyloid-forming peptides ranging from dimers to octamers has been examined using a variety of chain representation and energy models and methods including molecular dynamics (MD) (17–33), discontinuous MD (34,35), replica exchange MD simulations (36–38), and Monte Carlo (MC)-based methods (39,40). However, predicting the aggregation paths and the detailed atomic structures of the intermediates and of the initial stable structures also proved a difficult task for computer simulations. The typical timescales and the complexity of the free-energy landscape for aggregation are beyond what can be thoroughly characterized by the most detailed all-atom simulation methods.

In this study, our aim is to improve the reach of aggregation simulations by combining the extensive sampling properties of coarse-grained methods to the refinement abilities of

Submitted August 31, 2007, and accepted for publication December 11, 2007.

Address reprint requests to Giorgio Colombo, Istituto di Chimica del Ricoscimento Molecolare, Consiglio Nazionale delle Ricerche, Via Mario Bianco 9, 20131 Milano, Italy. E-mail: giorgio.colombo@icrm.cnr.it.

Editor: Ruth Nussinov.

0006-3495/08/06/4414/13 $2.00

doi: 10.1529/biophysj.107.121061

all-atom MD simulations. To achieve this goal, we first ran a preliminary MC exploration of the free-energy landscape of a system consisting of several copies of a peptide using a simplified, coarse-grained energy model (41–43). At this stage, the peptides are modeled as rigid chains in extended conformations to speed up the initial search for aggregated starting structures. The free-energy landscape associated to the MC is a clear simplification of the real one, due to the neglect of fine atomic interactions and of possible intrapeptide conformational rearrangements. However, the coarse-grained search has the potential to select aggregated structures that provide models of relevant intermolecular contacts and relative chain arrangements. These structures may represent a (simplified) picture of conformations present at the initial stages of the molecular recognition process. We could thus expect that multiple all-atom MD simulations started from the sampled configurations would have significant advantages over, for example, those starting from fully random placements of peptide positions or from subjectively chosen preformed oligomers. This approach is an extension of the methodology we previously developed for folding studies (41–43). Several simulations, starting from random placement of monomers, were run on systems consisting of two to eight copies of the GNNQQNY sequence. Control simulations on a scrambled peptide sequence (QNYGNQN) and on a mutant sequence (GNNAANY), deleting the contribution of Q side chains, were run to check whether the stable structures obtained were really a property of the sequence or of the simulation protocol.

This peptide was already studied via MD simulations by other groups (36,44–46). It was chosen because it represented the first example of a fibrillogenic sequence for which an atomic-level, high resolution structure was obtained (47,48). The structure revealed unique features along with expected properties. The final fibril shows a hierarchical organization based on different types of stabilizing interactions. The basic unit of the assembly is a pair of β-sheets separated by a dry interface. Each β-sheet is formed by parallel strands, which are orthogonal to the fiber axis. In addition to the classical hydrogen bonds between backbone atoms, the β-sheets are stabilized by hydrogen bonds between polar side chains (the polar zipper) (47). At the dry interface, the side chains of residues N2, Q4, and N6 are tightly interdigitated with the corresponding residues of the opposing sheet (steric zipper) (47). All of these properties represent an attractive target for benchmarking theoretical aggregation studies and molecular simulations in particular.

MATERIALS AND METHODS

Simulations were run on several systems consisting of different numbers of copies of the GNNQQNY peptide, starting from random initial configurations. Each set of simulations contained two to eight copies of the peptide. For each system, three different MD simulations were run, each starting from the representative structure of the three most populated structural clusters identified at the transition temperature of the preliminary coarse-grained MC analysis (vide infra). MD simulations are labeled according to the number of chains in the simulation box (2, 3, 4, . . . 8) followed by a ''_c'' (Chains), the cluster number (1, 2, or 3), and the temperature at which the simulation is run. So, for instance, the simulation of four peptide chains starting from the representative of the second most populated cluster at 300 K will be labeled 2c_2_300K. The two simulations starting from two preformed tetrameric seeds will be labeled ''Seed_1'' and ''Seed_2'' (vide infra). The same protocol was applied to two chains of a scrambled peptide sequence (QNYGNQN) and of the GNNAANY mutant.

Preliminary coarse-grained MC

The preliminary MC exploration of the configurational space is designed to be as efficient as possible, and it follows the same lines as the procedure described by others (41–43). In this section, we provide a short version of the procedure. Effective C_α and C_β centroids were used for each amino acid (with the exception of Gly, which lacks the C_β centroid). The C_α centroids are the centers of the effective interactions among contacting residues. The C_β centroids are used to capture the steric hindrance of the side chains. In the extension of this approach to the case of self-organization of multiple peptide chains, the single peptide chains are represented in their extended structure, with no internal flexibility.

The dynamics of multiple chains in configuration space is carried out using an MC technique and the potential described previously (41–43). The collapse temperature is defined by monitoring the temperature-dependent evolution of the radius of gyration and of the energy of the system.

The 3D space in which multiple peptide copies move is discretized by means of a grid representation. The grid is cubic, and its points are separated by 3 Å. This makes the exploration of the otherwise continuous space that separates different copies easier and more efficient. In particular, at the beginning of each simulation, peptide chains are randomly positioned in a cubic box, such that peptide-peptide separation is bigger than the length of a single extended chain. This should avoid biases toward particular configurations of the system. Randomly selected peptides are then randomly translated such that the center of mass (CoM) occupies a new grid point. At the same time, single chains are rotated in 3D space using the algebra of quaternions. The configurational space sampling is based on a simulated annealing Metropolis MC algorithm.

As mentioned above, all the protein chain configurations generated through the MC procedure have a simplified structural representation essentially described by the C_α degrees of freedom. These coarse-grained structures need to be reconstructed with all the atomic detail before they can be processed in ordinary all-atom MD schemes. In this study, we have used a novel knowledge-based strategy developed by De Mori and colleagues (42).

MD simulations from the peptide configurations selected at Tc

For each system containing between two to eight peptide chains, three different simulations were run, each starting from the representative structure of the three most populated structural clusters identified at the transition temperature of the coarse-grained MC analysis. The structural clustering algorithm used was developed by Daura and co-workers (49).

The method is based on the pairwise root mean-square deviation calculated from the $C\alpha$ superposition of aggregates. The procedure can be summarized as follows. After counting the number of neighbors using a cutoff of 0.20 nm root mean-square deviation between the optimal $C\alpha$ superposition of different structures, take the structure with the largest number of neighbors with all its neighbors as cluster and eliminate it from the pool of clusters. This procedure is repeated for the remaining structures in the pool.

For each complex, the MD simulation was carried out according to the following procedure:

The terminals were modeled as protonated to mimic the low pH (2.0) experimental conditions of crystal growth (47). The systems were solvated in a parallelepiped-shaped box large enough to contain 1.0 nm of solvent

around each initial aggregate. The simple point charge water model was used (50) to solvate each protein in the simulation box. Each system was subsequently energy-minimized with a steepest descent method for 5000 steps. The calculation of electrostatic forces used the particle mesh Ewald implementation of the Ewald summation method. The LINCS (51) algorithm was used to constrain all bond lengths. For the water molecules, the SETTLE algorithm (52) was used. A dielectric permittivity, $\epsilon = 1$, and a time step of 2 fs were used. All atoms were given an initial velocity obtained from a Maxwellian distribution at the desired initial temperature of 300 K. The density of the system was adjusted performing the first equilibration runs at constant number of particles, pressure, and temperature (NPT) condition by weak coupling to a bath of constant pressure ($P_0 = 1$ bar, coupling time $\tau_P = 0.5$ ps) (53). In all simulations, the temperature was maintained close to the intended values by weak coupling to an external temperature bath (53) with a coupling constant of 0.1 ps. The proteins and the rest of the system were coupled separately to the temperature bath. The Table 1 summarizes the simulation conditions and number of peptides for each simulation.

In all cases, the initial aggregates were simulated at 300 K for variable lengths; high temperature control simulations were also run at 360 K. All simulations were run at NPT conditions. All simulations and analysis were carried out using the GROMACS package (version 3.3; http://www.gromacs.org) (54), using the GROMOS96 43A1 force field (55,56). All calculations were performed on clusters of PCs, with the Linux operating system. Graphical display of structures was done using the PyMOL software (http://pymol.sourceforge.net/; 540 University Ave., Suite 325, Palo Alto, CA 94301).

RESULTS

The preliminary simplification of the free-energy landscape of systems of aggregating peptides is carried out following the same scheme as described previously (41–43). In summary, each peptide is described in terms of its C_α trace and of the associated C_β centroids. The residue-residue interaction energy is described by an energy functional that incorporates knowledge-based effective pairwise interactions and energetic penalties to avoid steric clashes. This simplified interaction potential derives from previous folding studies, where it proved to be effective in the description of the folding dynamics of small proteins with nontrivial tertiary structure features.

The rationale for extending the use of this potential to peptide aggregation stems from the consideration that the folding process is an intramolecular self-organization process governed by nonlocal interactions of the same physical nature as those governing the intermolecular aggregation processes. As a consequence, it is natural to apply the effective potential and simplified functional to the case of peptide self-assembly.

In this framework, the thermodynamics of the system consisting of multiple copies of the aggregating peptide (from two to eight copies) can be characterized by a simulated annealing MC evolution involving rotational and translational moves of the peptides on a discrete grid with a 3 Å distance between grid points. At this simulation stage, the peptides are considered as rigid chains in extended conformation.

The specific heat and radius of gyration (Rg) were reported as a function of the decreasing temperature of the coarse-grained system (Figs. 1 and 2). Analysis of these parameters

TABLE 1 Summary of MD simulations and simulation conditions

No. of chains	MC cluster*	Temperature (K)	Simulation label	Length (ns)	No. of atoms
2	I	300	2c_1_300	50	5698
2	II	300	2c_2_300	50	6790
2	III	300	2c_3_300	50	10615
2	I	330	2c_1_330	50	5698
2	II	330	2c_2_330	50	6790
2	III	330	2c_3_330	50	10615
2	Start from end of 2c_1_300	360	2c_1_360	100	5698
2	Start from end of 2c_1_300	360		100	6790
3	I	300	3c_1_300	50	4872
3	II	300	3c_2_300	50	24021
3	III	300	3c_3_300	50	30912
4	I	300	4c_1_300	50	8060
4	II	300	4c_2_300	50	9281
4	III	300	4c_3_300	50	13217
4	Start from end of 3c_1_300	330	4c_1_330	50	8060
4	Start from end of 3c_2_300	330	4c_2_330	50	9281
4	Start from end of 4c_3_300	330	4c_3_330	50	13217
4	Start from end of 4c_1_300	360	4c_1_360	50	8060
4	Start from end of 4c_2_300	360	4c_2_360	50	9281
4	Start from end of 4c_3_300	360	4c_3_360	50	13217
Preformed tetramer + 3 free chains	I	300	4_plus_3	50	17464
Preformed tetramer + 4 free chains	I	300	4_plus_4	50	7529
2 preformed tetramers	I	300	Seed_1	100	7477
2 preformed tetramers	II	300	Seed_2	100	13132

*Structural cluster obtained from the analysis of the coarse-grained MC simulation at the Cv peak from which the starting structure is extracted.

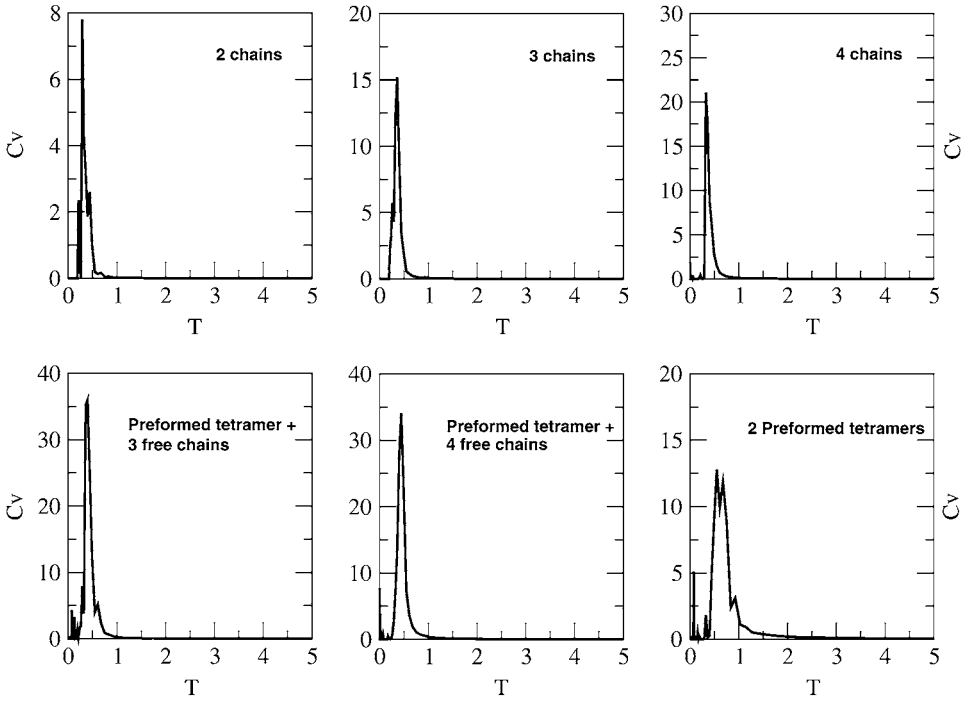

FIGURE 1 Specific heat as a function of the temperature of the coarse-grained MC models for the different peptide ensembles simulated.

enables the identification of a collapse temperature, *T*c, at which each system undergoes a compaction transition. The collapse is represented by a sudden (cooperative-like) decrease in the radius of gyration and a peak in the specific heat. At *T*c, the specific heat exhibits a peak that reflects the large energy fluctuations associated to a phase transition characterized by the simultaneous presence of multiple configurations for the systems under examination. The multiple chain system is in fact undergoing a transition from a totally disordered condition to an ensemble of aggregate structures with a certain degree of compactness. The latter ensemble is characterized by the coexistence of polymeric structures with one to two residues per chain involved in interchain contacts with other structures in which the number of intermolecular interactions is higher, and the chains lie in the same plane with either antiparallel or parallel orientations. Similar to what was

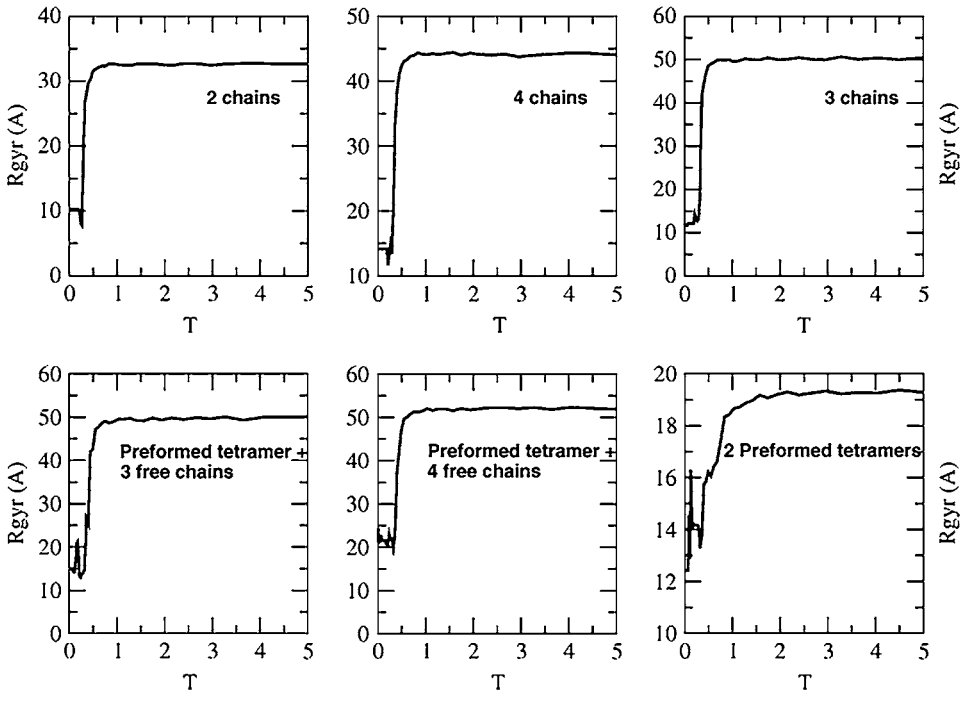

FIGURE 2 Radius of gyration as a function of the temperature of the coarse-grained MC models for the different peptide ensembles simulated.

observed for previous folding studies, the polymeric ensembles at *T*c become even more compact at lower temperatures in the coarse-grained MC, suggesting that, at this temperature, the aggregates can be prone to collapse into more compact conformations with interesting structural properties.

Representative aggregated structures are thus selected at the *T*c of the MC dynamics through the application of a structural clustering algorithm. The conformational variability of the aggregates at *T*c is such that significantly different structures can be picked and so make these structures viable candidates to investigate the dynamics of formation of oligomeric species and their structural properties at atomic resolution

These properties justify the choice of starting conformations for explicit solvent all-atom MD runs from structures at *T*c and ensure that MD runs will be ''time-advanced'' with respect to ones started from peptides distributed in a completely random fashion. In this latter case, in fact, most of the all-atom simulation effort would be used to bring the peptides close enough to interact with one another.

The structures at *T*c are grouped in families of similar conformations (41–43) using the conformational clustering algorithm developed by Daura et al. (49). The representative structures of the three most populated clusters from each of the MC runs are then selected as starting structures for the all-atom MD runs in explicit water at 300 K. These structures had to be first reconstructed in atomic detail according to the procedure described in Methods and by De Mori and colleagues (42). In the following sections, we will describe the results of simulations involving different numbers of peptides.

Two chains

As outlined in the previous paragraph, the representatives of the structural ensemble at the collapse temperature (*T*c) in the MC run are identified through the application of the clustering algorithm devised by Daura (41–43,49). The structures corresponding to the representatives of the three most populated clusters, after reconstruction of the whole atomic detail and solvation in explicit water, are then used as starting structures for long timescale MD simulations. The latter are then analyzed in terms of different structural and conformational parameters. Moreover, representative structures are selected through the application of a second structural cluster analysis to the all-atom MD trajectories. The analysis of the trajectories immediately shows interesting features. In the 50-ns, all-atom simulation starting from the representative of the first MC cluster at 300 K (simulation labeled 2c_1_300K), the two peptides form a stable antiparallel structure, with the amidic side chains interdigitated at the interface between the two peptides with remarkable shape complementarity. In this particular arrangement, the interface between the two peptides is not occupied by water molecules. This structure closely resembles the one of the dry interface (47,48,57), with two antiparallel β-sheets in the same plane held together by Van der Waals contacts between sterically complementary side chains that exclude water molecules from the environment (Fig. 3, *top panel*).

In the trajectory started from the representative structure of the second MC cluster determined at the *Tc*, 2c_2_300K, the formation of a stable parallel β-sheet is evident immediately after the first few nanoseconds of simulation. This structural motif is stabilized by in-register backbone-backbone H-bonding interactions. Additional stabilization and ordering is provided by side chain-side chain interactions between residues facing each other in the parallel arrangement. These involve not only the formation of stable hydrogen bonds between N and Q amidic groups but also packing interactions between the aromatic rings of the terminal tyrosines (Fig. 3).

The initial structure of the trajectory starting from the representative structure of the third MC cluster, 2c_3_300K, is characterized by the two terminal Y residues forming a packing interaction, with the remaining part of each peptide chain pointing into the solvent without additional interchain contacts. Interestingly, this type of Y-Y packing is also observed in the crystal structure of the fibril, with the role of connecting different interdigitated dimeric complexes. In the all-atom MD simulations, however, this starting structure is not stable and evolves toward the antiparallel arrangement, already observed for simulation 2c_1_300K, with a reptation movement.

Control simulations at the temperature of 360 K were run to check the stability of the structures obtained in the above-reported all-atom simulations at 300 K. The antiparallel arrangement observed in simulations 2c_3_300K and 2c_1_300K is disrupted at around 5 ns. The two chains explore an ensemble of bent conformations and, at around 21 ns, the two chains relax into a parallel arrangement that is disrupted and recovered during the rest of the simulation.

In contrast, running a simulation at 360 K starting from the final stable parallel conformation of simulation 2C_2_300K does not cause any significant changes in terms of structure or relative organization of the strands, suggesting that the parallel arrangement is particularly stable for this sequence (Fig. 3).

Three chains

Increasing the number of chains in the simulation box to three determines a different and more varied conformational scenario. In the simulation started from the first cluster obtained at the *T*c of the MC run, labeled 3c_1_300K, one chain is juxtaposed to a preformed antiparallel dimer similar to that observed in simulation 2c_1_300K. The preformed dimer is held together by interactions between the side chains forming the dry interface, and the third chain binds parallel to one of the previous two (Fig. 3).

Simulation 3c_2_300K starts from a dimer where the dry interface is formed, and the remaining chain is far in space. The latter, however, is eventually able to dock on the preformed complex forming a parallel structure with one of the two prestructured strands, similar to the one described above.

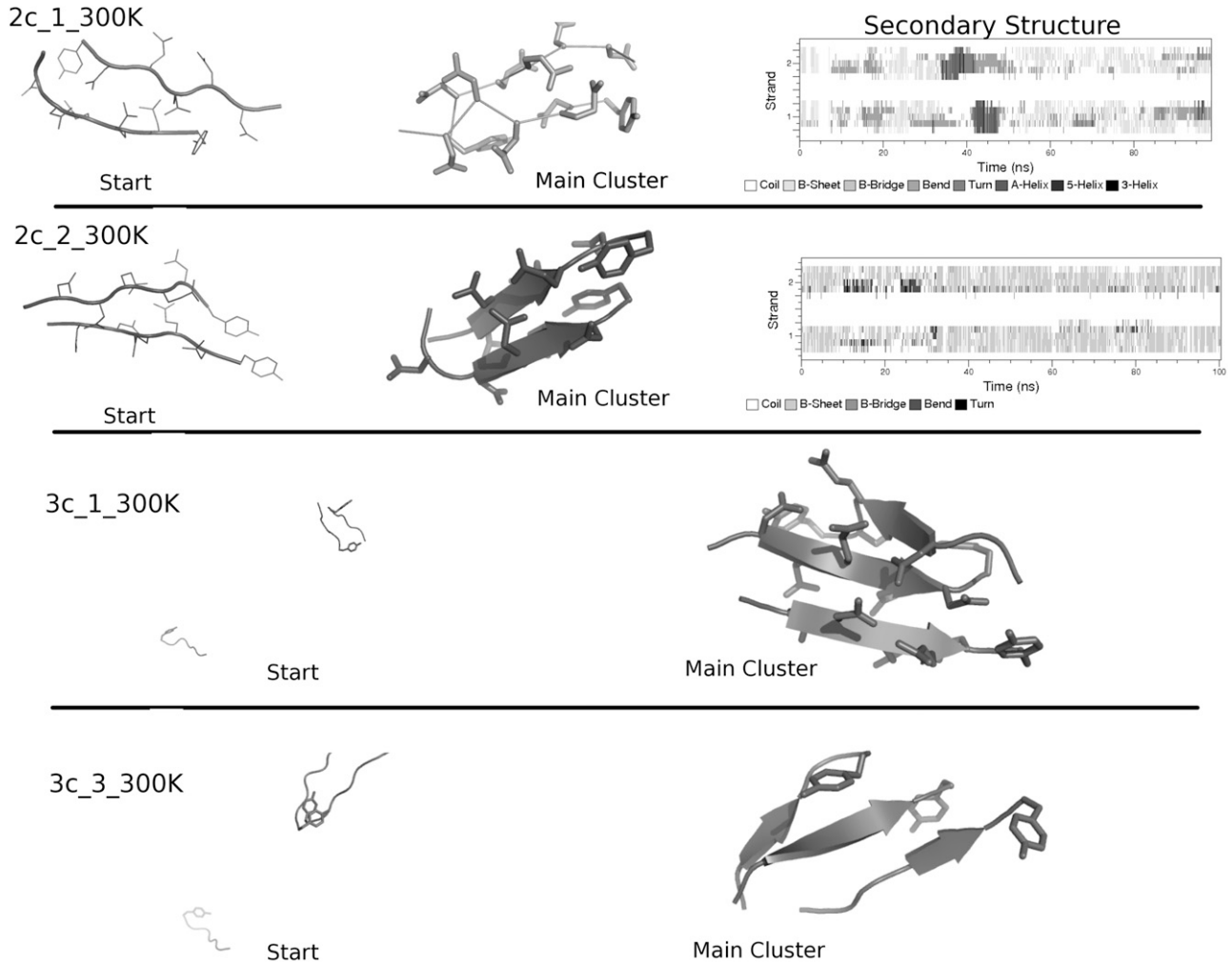

FIGURE 3 The all-atom 3D structures of the starting and most representative conformations for MD simulations 2c_1_300K, 2c_2_300K, 3c_1_300K, 3c_3_300K (*from top to bottom*). In the top two panels, the time-dependent evolution at 360 K of the secondary structure of the dimeric aggregates obtained from MD at 300 K is also reported. The secondary structure of the parallel arrangement is stable at 360 K (*second panel from top*) contrary to the antiparallel one. Legend for the secondary structure is in the figure in grayscale. Light gray represents ordered β-sheet structures.

Finally, simulation 3c_3_300K starts with a preformed dimer with the two tyrosines packing with their aromatic rings, the remaining part of the chains pointing toward the same direction in space, and the third chain separated by ~50 Å from the other two. Interestingly, the evolution in the all-atom MD simulation at 300 K leads to the formation of a parallel trimer with a significant degree of packing of the three Y residues and backbone-backbone interchain hydrogen bonds stabilizing the parallel conformation (Fig. 3). Heating this structure to high temperatures (360 K) does not change the situation, once again suggesting an enhanced stability of the parallel conformations for oligomeric complexes.

Four chains

The simulations involving four different peptide chains were run to investigate the dynamics of formation of the putative fibril nucleus, as identified by Nelson and co-workers (47). Similar to the previous cases, three possible starting configurations are identified from the cluster analysis of the *T*c structural ensemble in the coarse-grained MC trajectory. The representative structure of the first cluster shows a ''micelle-like'' arrangement for the four aggregated chains: the four Y residues point toward the center of the micelle forming a hydrophic core with their aromatic rings, whereas the rest of the peptide points outward into the solvent. During the 50-ns, all-atom, MD simulation at 300 K, 4c_1_300K, this construct reorganizes to form a stable parallel β-sheet involving two central chains. The remaining two chains remain basically disordered (Fig. 4).

The simulation involving the representative of the second cluster, 4c_2_300K, shows the formation of a stable parallel β-sheet involving four strands (Fig. 4). It is worth emphasizing that this structure is totally uncorrelated with the starting one. The latter is characterized by the presence of two dimeric complexes, contacting through the aromatic side

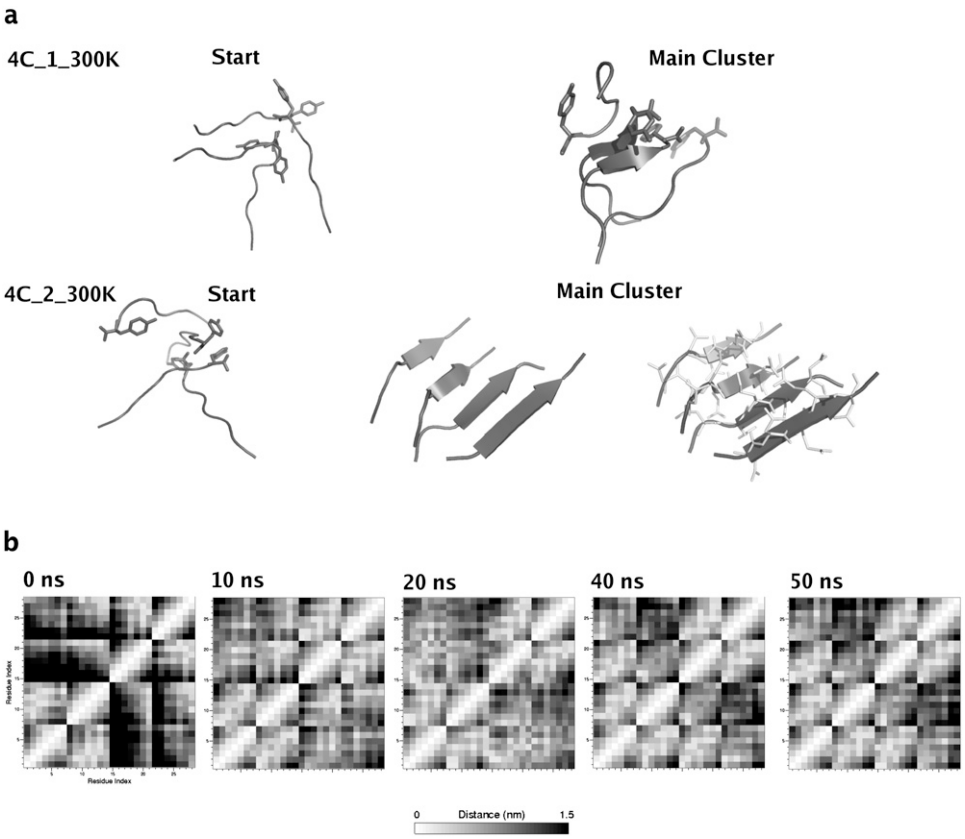

FIGURE 4 (*a*) The all-atom 3D structures of the starting and most representative conformations for MD simulations 4c_1_300, 4c_2_300. In the case of the most representative structure for simulation 4c_2_300, besides the ordered four-stranded β-sheet structure, the interplay among interdigitating amidic side chains is reported. (*b*) Time-dependent evolution of the contact matrix in simulation 4c_2_300 based on atom-atom distances—from a disordered situation, evolution toward a typical β-sheet contact matrix is apparent.

chains of the tyrosines. Each of the dimeric complexes is initially devoid of ordered secondary structure at the beginning of the simulation. Very interestingly, however, we could observe the rapid alignment of the two chains in each dimer to form a parallel in-register β-sheet. During the simulation, the two β-sheets evolve spontaneously to form a flat parallel β-sheet involving all the four chains present in the simulation box. This structure is remarkably stable at 300 K and at higher temperatures (330 K and 360 K). The rapid formation and the stability of this supramolecular complex are in substantial agreement with the suggestion by Nelson and coworkers that the four chains may represent a stable nucleus for fibril formation (47).

The results obtained in the simulation starting from the third structure is similar to that described for 4c_2_300K, and will not be discussed further.

Seeded aggregation

The stability of the four-stranded parallel β-sheet, combined with the fact that it forms rapidly even on the MC-MD time scales, suggested additional simulations in which this complex could be used as a preformed seed. Two MC simulations were thus run on two systems consisting of the following:

1. the preformed, four-stranded, parallel β-sheet plus three free chains.
2. the preformed, four-stranded, parallel β-sheet plus four free chains.

In the preliminary coarse-grained MC run the additional (three or four) peptides are completely free to translate and rotate in space, the only constraint being on the structure of the seed. In the subsequent all-atom MD, none of the peptides is subjected to any constraint. As a consequence, the whole system is free to move and reorganize.

In both cases, we only reconstructed and simulated the representative structure of the most populated cluster from the MC exploration at the all-atom level. This was done to save on computational time, considering the dimensions of the boxes in the presence of such a high number of free and unconstrained chains.

In the all-atom MD simulation starting from the first ensemble of structures, labeled 4_plus_3, the preformed four-stranded β-sheet remains stable during the entire simulation time. What is most interesting to observe is that all the three remaining peptides dock on the preformed plane in an antiparallel orientation with respect to the seed and parallel to one another. This arrangement is similar to the one observed in the crystal structure in the interface between the two β-sheet layers running parallel to the fibril axis (Fig. 5).

The second simulation, started from a ''seeded'' system, contains the preformed four-stranded β-sheet plus the four free peptides, labeled 4_plus_4. The results of the all-atom,

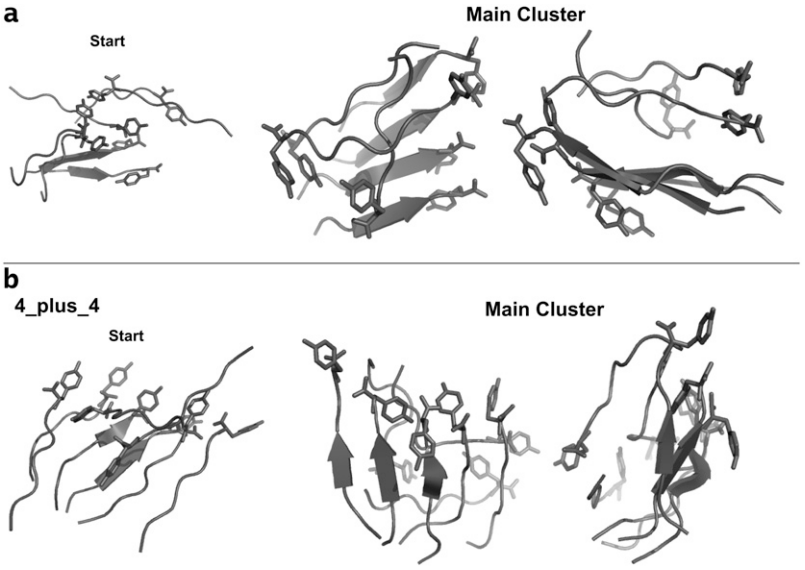

FIGURE 5 The all-atom 3D structures of the initial and most representative conformations for MD simulation started from the results of the coarse-grained MC on (*a*) preformed tetramer plus three free chains; (*b*) preformed tetramer plus four free chains. The structures on the extreme right-hand side of the figure are the same as those in the middle of the panel rotated by 90°.

unconstrained, MD run demonstrate that one of the free peptides adds up to the seed structure in a parallel fashion. One of the three remaining peptides packs on one side of the plane of this complex, with an antiparallel orientation with respect to the five-stranded β-sheet. The two remaining peptides do not attain any particularly ordered structure during the simulation time, although they show a tendency toward packing with an antiparallel orientation with respect to the preformed seed.

Aggregation of preformed tetramers

One final MC dynamics simulation was run using two preformed, four-stranded, parallel β-sheets as input structures. This choice was dictated by the previous observation of the stability of this particular conformation, and by the importance of this structure as a possible nucleus for the formation of the whole fibril suggested by experiments. Two main conformational families are identified by the cluster analysis of the coarse-grained MC search. The representatives of these two main clusters are thus reconstructed with the full atomic detail, and MD simulations were run for 100 ns at 300 K (simulations Seed_1 and Seed_2).

The starting structure of the first MD simulation is characterized by the two four-stranded β-sheets packed on top of each other in a parallel-like fashion. The rings of the Y residues form an initial stable aromatic core. During the simulation, the aromatic rings pack mainly in a face-to-face conformation, and the overall topology of the starting conformation is conserved with little perturbation. Interestingly, water is completely excluded from the intersheet packing. Conformational changes involve a partial breaking of one of the four-stranded β-sheets into parallel dimers, which tend to align with the remaining stable nucleus. Interestingly, the spontaneous formation of tetramer-dimer complexes has already been observed in previous simulations of the aggregation of the short fibrillogenic NFGAIL sequence with the coarse-grained OPEP force field (40).

The second MD simulation is started from a conformation where one of the two parallel four-stranded β-sheets is rotated by $\sim$90° with respect to the second one. This structure evolves toward a β-barrel–like organization of the peptides, which are held together by H-bonding and aromatic-packing interactions. Interestingly, the barrel-like structure represents a minimum energy conformation among all the structural cluster representatives isolated for this system, according to GB/SA calculations (Fig. 6). Barrel-like structures similar to the one observed here were previously reported also for different sequences with different simulation methods (59). Experimentally, this type of arrangement of the monomers is reminiscent of the synuclein protein rings described by Lashuel et al. (60) and of the initial structures formed during the aggregation of hydrophobins (61). Moreover, the β-barrel arrangement has been observed in the formation of peptide pores in membranes mimicking environments (62).

β-barrel-like structures form fast on the experimental scale, and they can be considered as intermediates on the pathway to the evolution of the final fibrillar architecture.

To check on the relative stability at high temperatures of the two starting structures, two more MD simulations were run at 360 K. In both cases, the starting aggregates appear to be quite labile at high temperatures, and the only ordered structural motif surviving at these conditions is represented by parallel, in-register, β-sheet dimers.

DISCUSSION

In this article, we have presented the results of multiple simulations on the first stages of the ordered self-organization of the peptide GNNQQNY. The availability of high resolution structures of the aggregate makes this system an optimal model to test aggregation simulations. This work builds on

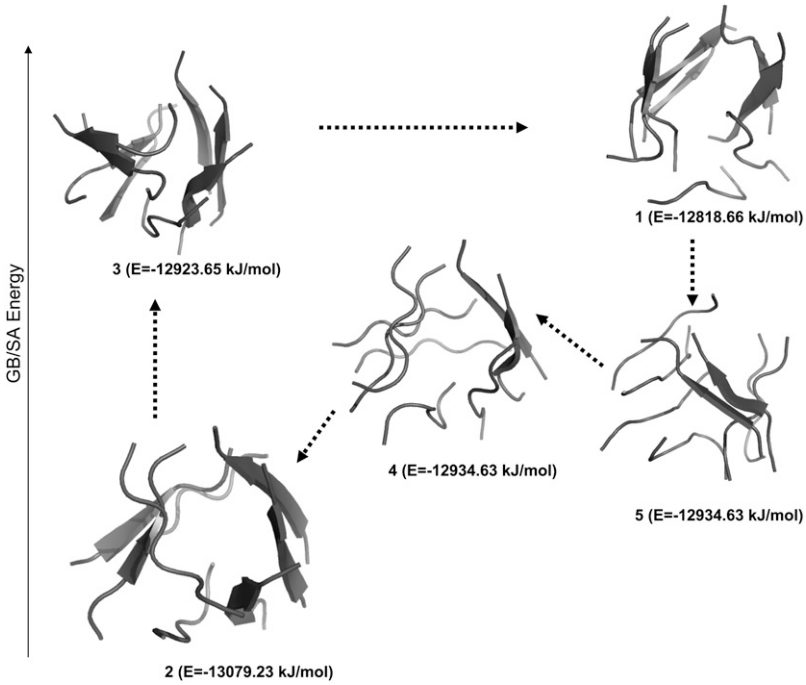

FIGURE 6 Representative structures of the most populated clusters from all-atom MD of the aggregation simulation for two preformed four-stranded β-sheets (preformed tetramers). The number identifying each structure represents the cluster rank (Number 1 is the most populated cluster). The value of the GB/SA energy in water of the complex is reported in parentheses. It is apparent that the barrel-like structure is an energy minimum. Dashed arrows represent transitions between clusters, indicating possible paths between cluster structures.

our previous experience on protein-folding studies in trying to extend the reach of all-atom MD simulations by a physically based preselection of starting conformations. This selection is carried out through the use of a preliminary Monte Carlo exploration of a simplified representation of the free-energy landscape of a system consisting of multiple copies (two to eight in this study) of the peptide evolving on a discrete grid. The simplification has to be introduced to adequately sample the otherwise enormously vast conformational space available to the system. The necessary simplifications also prevent the structures obtained from being as accurate as desired. Thus, the idea of feeding them hierarchically into a more accurate method that samples less effectively but with a higher degree of chemical and physical detail can be exploited. In this respect, all-atom, explicit, solvent MD simulations started from the sampled conformations are expected to have significant advantages over, for example, those starting from fully random peptide orientations, which can be affected by insurmountable (on the MD time scale) lag-phases. As a caveat, one should consider that, despite starting from time-advanced metastable structures, the picture provided by the simulations may still be limited because of the difficulties of extracting quantitative thermodynamic and kinetic data from limited sampling. A second caveat is that, at this stage, in parallel to the protein-folding case, it is not possible to provide the actual quantification of the time-advancement in the absence of clear, general, experimentally derived characterizations of the fibrillation mechanisms and in the absence of the definition of a general reaction coordinate. The benefits of this strategy can be ascertained only through the comparison of the degree of order, β-sheet content, 3D organization, and side-chain interaction patterns of the explored trajectories comparing them to the experimentally determined structure of the GNNQQNY fibril. The final results of our simulations are consistent with the structures and with the experimentally derived mechanistic considerations on fibril formation based on these structures.

The analysis of our mixed MC-MD trajectories indicates a hierarchical picture of the assembly process, similar to what Nelson and co-workers hypothesized based on structural analysis of the fibril (47). Parallel β-sheets form easily in many instances and, in general, are not disrupted by the high temperature conditions. Moreover, they are mostly characterized by an in-register organization of interchain hydrogen bonds. Preformed parallel β-sheets can act as templates in the aggregation process to accelerate fiber formation. The stability of the parallel arrangement was observed by other groups using either different simulation strategies (46), different force fields with explicit solvent (63), or different solvents with implicit solvent representations (45). To ascertain that the stability of the representative structures is a property of the specific sequence and not an artifact of the protocol used, we have investigated the aggregation properties of a scrambled sequence of the same peptide, namely QNYGNQN, and of a mutated sequence GNNAANY (see Supplementary Material, Data S1). Interestingly, structural analysis of the structure at Tc in the preliminary MC provides only two major clusters that are significantly populated. In the case of one of the two (cluster 2), the similarity to experimentally determined geometries is much poorer than in all previously observed cases thus demonstrating that the coarse-grained Hamiltonian captures some of the main features that guide the aggregation process. In the other main cluster (cluster 1), the peptides are in a parallel orientation. Once the starting structures are reconstructed at the all-atom level and

simulated at 300 K in explicit water, an ordered β-sheet structure appears at the beginning of the simulation, but it is not stable and is lost after ~20 ns. The all-atom simulation started from the representative of cluster 2 immediately leads to totally disordered secondary structures during the simulation time. In the case of sequence GNNAANY, two main structural clusters result from the MC analysis. The analysis of the corresponding all-atom MD trajectories shows that stable β-sheet conformations cannot be found in either of the two simulations. The peptides, despite being spatially close and in contact, lack the necessary side chain contacts and conformational properties to assemble in a fibril-favorable conformation. These results show that the properties of the structures of the oligomers simulated in this work are a result of the properties of the sequence and not of the simulation protocol.

Once the first seed is formed, more strands can pack on it mainly through van der Waals and H-bonding interactions involving the N and Q side chains, giving rise to the second level of organization. Optimal packing is achieved when the peptides forming the second plane are antiparallel to the first one. This mechanism leads to the formation of the dry interface, with exclusion of water molecules from the space between the two layers. We observed that the pattern of interactions in the second stage is less specific than the one determined by the formation of in-register hydrogen bonds observed in the first level, as can be expected considering the flexibility of the amidic side chains of the interdigitating residues. The nonspecificity of van der Waals interactions in the dry interface leads to polymorphism in the structures of the starting aggregates, which can also be reflected in the amyloid fibril polymorphism observed for several sequences and is strictly dependent on the solution conditions (47).

The third level of organization would involve pair-of-sheets structures interacting to form the fibril, but our simulations cannot reach such a degree of complexity at this level. However, the docking of multiple preformed, double-layer structures with a dry interface would represent a natural consequence of the first two levels of events.

It is also interesting to note that the ensemble of structures observed is extremely dependent on the number of molecules simulated, especially in the cases of two, three, or four peptide chains in the simulation box. In the case of three peptides, in particular, we could observe a high diversity of aggregate conformations consistent with the fact that ordering three peptide molecules can determine a substantial kinetic barrier to fibril formation. In the case of four peptides in solution, we observed the easy formation of two stable, parallel β-sheets, which can subsequently dock as two rigid, preordered bodies to form the four-stranded nucleus. It is worth noting at this point that our simulations, starting from random placements of peptide chains, directly probe a mechanistic hypothesis based on structural considerations suggested by Nelson et al. (47).

The ease of formation of a four-stranded structure from two preformed, parallel β-sheet dimers supports a hierarchical picture for the fibril formation mechanism. Qualitatively speaking, this type of mechanism parallels the formation of local elementary structures in protein-folding and should reduce the entropic costs for the ordered organization of multiple chains compared to the cost of sequential addition and ordering of single chains. The stability of the preformed tetramer was noticed by Zheng and co-workers using a different force field (63). In this case, the authors performed several 10 ns long simulations with all-atom, explicit, solvent molecular simulations of various sizes and arrangements of preformed models of GNNQQNY oligomers (and mutants) to study its stability and dynamics. Their results also suggest that the minimal nucleus for fibril formation could be small and likely to consist of three or four peptides.

Using a preconstituted, four-stranded, parallel β-sheet as an aggregation seed in the presence of free peptides, it is possible to obtain an arrangement of the remaining unconstrained peptides in the simulation box, which is reminiscent of that observed in the x-ray structures of the fibril. Isolated peptides can dock on top of the plane of four-stranded parallel β-sheets in an antiparallel orientation with respect to the peptides of the seed, identifying a possible mechanism for the formation of the second layer of peptides and of the dry interface. Interestingly, the N and Q side chains of the seed and those of the free peptides docking on it end up forming a strongly intertwined configuration, excluding the presence of water from the region between the preformed and forming β-sheets.

A possible different hierarchy of structure formation was probed by simulating the possible aggregation of two seeds. Our results suggest that the structures formed by the docking of two seeds would have to undergo major structural rearrangements to obtain an antiparallel organization of the two facing β-sheets. Possible intermediates on this rearrangement pathway are represented by compact and barrel-like structures whose presence has already been observed in the aggregation pathways of other sequences, both at the simulation and at the experimental level. The conformational energy barriers opposing this reorganization must be expected to be quite high, considering that one should break and reform multiple energetically favorable H-bonding and side chain-side chain hydrophobic and aromatic interactions.

Although this latter mechanism cannot be excluded at this level of simulation, we think that the previous picture based on the formation of an ordered seed, followed by the juxtaposition on it of single chains with the appropriate orientations, should be regarded qualitatively as more accessible from the energetic point of view. The conformational reorganization of single peptide chains on the preformed template would in fact be more sterically and energetically favorable than the analogous reorganization of a larger, more complex substructure.

The GNNQQNY sequence has already been the subject of simulation studies from different groups (36,44–46). Gsponer and co-workers (45) studied the ab initio aggregation pathways of three copies of the peptide with MD using a simplified,

implicit solvent model that allowed long time scales to be reached with multiple simulations. They observed that the process was not purely downhill on the energy surface, due to the presence of enthalpic barriers that originated from out-of-register interactions. The parallel β-sheet arrangement was found to be favored over the antiparallel thanks to side chain contacts. Intermediates consisting of a mixed parallel-antiparallel geometry were found along these pathways. These results were substantially confirmed in a subsequent article by the same group through the use of replica exchange MD simulations (36) to sample conformational space and transitions more efficiently, using the same simplified, implicit, solvent approach. In general, considering a system consisting of three chains, the authors observed that, at higher temperatures, the peptides were more likely to form disordered aggregates characterized by nonspecific interactions. In this range of temperatures, the enthalpic contribution due to in-register backbone or side chain interactions did not dominate the entropic one, and the growth of ordered nuclei was found to be forbidden. However, when the temperature decreased, the entropic contribution became less important and ordered in-register aggregates started forming. The parallel arrangement also was found to be favored in these cases. The stability of the parallel arrangement in the initial stages of amyloid formation also emerged as an important feature in the MD study carried out by Zhang and co-workers (46). Therefore, the authors suggested that new strands might prefer to extend in a parallel arrangement to form oligomers, in agreement with x-ray–based structural information.

The results of our simulations, besides showing a substantial agreement with the other simulations involving up to four peptide chains, extend the possibility to study systems that are more complex and provide atomic-resolution insights into several aspects of peptide aggregation. We could directly examine different possible pathways leading to higher order aggregates using a realistic simulation temperature of 300 K, starting from stable structures obtained from preliminary unbiased MC simulations. A four-stranded parallel β-sheet can easily form through the docking of two preformed parallel dimers and act as a template for the docking of single chains or interact with other preformed seeds. In the former case, docking of single chains leads directly to the formation of the dry interface with the exclusion of water (a dock-and-lock type of mechanism). In the latter, the system forms barrel-like structures, trapping water molecules inside the barrel. In this case, the mechanism leading to the structure observed in the microcrystal would involve major rearrangements and exclusion of water from the core. Both the mechanisms may be energetically accessible in solution, with the one going through the barrel intermediate being characterized by higher energetic barriers due to the rearrangement and dewetting of the interface.

Moreover, through the all-atom refinement stage, we could observe different possible interactions and geometries for the amidic side chains, which form the cross-β-spine, stabilizing different structures and leading eventually to fibril polymorphism. Obtaining this type of insight at all-atom resolution without neglecting the presence of water would require extremely long simulation times, spent mainly in simulating peptide diffusion. A possible way to overcome this problem would be the use of extremely concentrated solutions as we proposed in a previous article (21). However, in that case, the formation of initial, very compact structures prevented the reorganization of the initial aggregates with the formation of specific H-bonding interactions typical of the β-sheet geometries.

The approach we propose here simulates the initial formation of unordered aggregate structures, which could be the rate-limiting step of the process, by the preliminary coarse-grained MC investigation. Hydrophobic interactions, reproduced by the residue-specific statistical potential, could play a key role at the beginning of the process by disfavoring organized aggregates formation. Interestingly, these structures appear to be partially stable once they are brought into the realm of all-atom molecular mechanics. Reorganization of this initial aggregate with the formation of more specific H-bonding and cross-β-spine interactions determines, in most cases, the formation of stable parallel dimers. The latter determine the formation of larger structures—ordered, four-stranded, parallel β-sheets, in particular—that can drive the process of fibril growth favoring the ordered docking of other monomers to reach the final amyloid state.

In conclusion, we have followed the all-atom dynamical evolution of the initial stages of the aggregation process of the GNNQQNY peptide starting from several different configurations obtained by a preliminary coarse exploration of the complex free-energy landscape. Starting from completely unbiased placements and orientations of peptides (up to eight chains), the simulated trajectories highlight the formation of ordered, mainly parallel, β-sheet aggregates. In particular, once the dimeric structures are formed, they display a high structural stability, even at high temperatures. The analysis of the chemical interactions, secondary structure evolution, and 3D conformational evolution, and the ability to include explicit solvent to monitor the effect of water on the self-organization process has resulted in dynamical information on the first stages of the aggregation process at the all-atom level of resolution that is still impossible to obtain at the experimental level. The results of our simulations suggest the presence of multiple aggregation pathways and indicate a hierarchical mechanism of fibril formation based on the preliminary formation of organized elementary structures reducing the entropic cost of bringing multiple chains together. Interestingly, the stable structures emerging from the simulations reproduce the main structural features observed for the peptides in the x-ray structures obtained by Nelson and co-workers (47).

The protocol used here has been kept as general and unbiased as possible and has shed light on different possible organization mechanisms. Therefore, the proposed strategy might be applicable to other instances of short proteins with significant advantages over the use of only MD simulations.

SUPPLEMENTARY MATERIAL

To view all of the supplemental files associated with this article, visit www.biophysj.org.

The authors acknowledge Dr. Giacomo Carrea for discussion and support. The authors thank LITBIO Consortium for generous allocation of computer time.

This work was supported by funding under the Sixth Research Framework Programme of the European Union (ref. LSHB-CT-2006-037325); under the FIRB program: Folding e aggregazione di proteine: metalli e biomolecole nelle malattie conformazionali (RBNE03PX83); and under support of the Ministero degli Esteri exchange program: understanding the molecular determinants of amyloid fibril formation in human neurodegenerative diseases.

REFERENCES

1. Chiti, F., and C. M. Dobson. 2006. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* 75:333–366.
2. Berriman, J., L. C. Serpell, K. A. Oberg, A. L. Fink, M. Goedert, and R. A. Crowther. 2003. Tau filaments from human brain and from in vitro assembly of recombinant protein show cross-β structure. *Proc. Natl. Acad. Sci. USA.* 100:9034–9038.
3. Serpell, L. C. 2000. Alzheimer's amyloid fibrils: structure and assembly. *Biochim. Biophys. Acta.* 1502:16–30.
4. Makin, O. S., and L. C. Serpell. 2005. Structures for amyloid fibrils. *FEBS J.* 272:5950–5961.
5. Lansbury, P. T., and H. A. Lashuel. 2006. A century-old debate on protein aggregation and neurodegeneration enters the clinic. *Nature.* 443:774–779.
6. Cohen, E., J. Bieschke, R. M. Perciavalle, J. W. Kelly, and A. Dillin. 2006. Opposing activities protect against age-onset proteotoxicity. *Science.* 313:1604–1610.
7. Jarrett, J. T., and P. T. L. Jr. 1992. Amyloid fibril formation requires a chemically discriminating nucleation event: studies of an amyloidogenic sequence from the bacterial protein OsmB. *Biochemistry.* 31: 12345–12352.
8. Guijarro, J. I., M. Sunde, J. A. Jones, I. D. Campbell, and C. M. Dobson. 1998. Amyloid fibril formation by an SH3 domain. *Proc. Natl. Acad. Sci. USA.* 95:4224–4228.
9. Zurdo, J., J. I. Guijarro, and C. M. Dobson. 2001. Preparation and characterization of purified amyloid fibrils. *J. Am. Chem. Soc.* 123: 8141–8142.
10. Ellis-Behnke, R. G., Y. X. Liang, S. W. You, D. K. Tay, S. Zhang, K. F. So, and G. E. Schneider. 2006. Nano neuro knitting: peptide nanofiber scaffold for brain repair and axon regeneration with functional return of vision. *Proc. Natl. Acad. Sci. USA.* 103:5054–5059.
11. Smith, J. F., T. P. Knowles, C. M. Dobson, C. E. Macphee, and M. E. Welland. 2006. Characterization of the nanoscale properties of individual amyloid fibrils. *Proc. Natl. Acad. Sci. USA.* 103:15806–15811.
12. Nilsson, M. R. 2004. Techniques to study amyloid fibril formation in vitro. *Methods.* 34:151–160.
13. Walsh, D. M., I. Klyubin, J. V. Fadeeva, W. K. Cullen, R. Anwyl, M. S. Wolfe, R. J. Rowan, and D. J. Selkoe. 2002. Naturally secreted oligomers of amyloid β protein potently inhibit hippocampal long-term potentiation in vivo. *Nature.* 416:535–539.
14. Colombo, G., P. Soto, and E. Gazit. 2007. Peptide self-assembly at the nanoscale: a challenging target for computational and experimental biotechnology. *Trends Biotechnol.* 25:211–218.
15. Holmes, T. C., S. de Lacalle, X. Su, G. Liu, A. Rich, and S. Zhang. 2000. Extensive neurite outgrowth and active synapse formation on self-assembling peptide scaffolds. *Proc. Natl. Acad. Sci. USA.* 97: 6728–6733.
16. Silva, G. A., C. Czeisler, K. L. Niece, E. Beniash, D. A. Harrington, J. A. Kessler, and S. I. Stupp. 2004. Selective differentiation of neural progenitor cells by high-epitope density nanofibers. *Science.* 303:1352–1355.
17. Baumketner, A., and J. E. Shea. 2007. The structure of the Alzheimer amyloid β 1035 peptide probed through replica-exchange molecular dynamics simulations in explicit solvent. *J. Mol. Biol.* 366:275–285.
18. Kamiya, N., D. Mitomo, J. E. Shea, and J. Higo. 2007. Folding of the 25 residue Aβ (12–36) peptide in TFE/water: temperature-dependent transition from a funneled free-energy landscape to a rugged one. *J. Phys. Chem. B.* 111:5351–5356.
19. Daidone, I., F. Simona, D. Roccatano, R. A. Broglia, G. Tiana, G. Colombo, and A. D. Nola. 2004. β-hairpin conformation of fibrillogenic peptides: structure and α-β transition mechanism revealed by molecular dynamics simulation. *Proteins..* 57:198–204.
20. Lopez de la Paz, M., G. M. S. de Mori, L. Serrano, and G. Colombo. 2005. Sequence dependence of amyloid fibril formation: insights from molecular dynamics simulations. *J. Mol. Biol.* 349:583–596.
21. Colombo, G., I. Daidone, E. Gazit, A. Amadei, and A. Di Nola. 2005. Molecular dynamics simulation of the aggregation of the core-recognition motif of the islet amyloid polypeptide in explicit water. *Proteins.* 59:519–527.
22. Zanuy, D., Y. Porat, E. Gazit, and R. Nussinov. 2004. Peptide sequence, molecular assembly and amyloid formation: molecular simulations and experimental study of a human islet amyloid polypeptide fragment and its analogues. *Structure.* 12:439–455.
23. Ma, B., and R. Nussinov. 2002. Stabilities and conformations of Alzheimer's β-amyloid peptide oligomers (Aβ 16–22, Aβ 16–35, and Aβ 10–35): sequence effects. *Proc. Natl. Acad. Sci. USA.* 99:14126–14131.
24. Zanuy, D., B. Ma, and R. Nussinov. 2003. Short peptide amyloid organization: stabilities and conformations of the islet amyloid peptide NFGAIL. *Biophys. J.* 84:1884–1894.
25. Klimov, D. K., J. E. Straub, and D. Thirumalai. 2004. Aqueous urea solution destabilizes Aβ (16–22) oligomers. *Proc. Natl. Acad. Sci. USA.* 101:14760–14765.
26. Nguyen, P. H., M. S. Li, G. Stock, J. E. Straub, and D. Thirumalai. 2007. Monomer adds to preformed structured oligomers of A β-peptides by a two-stage dock-lock mechanism. *Proc. Natl. Acad. Sci. USA.* 104: 111–116.
27. Tarus, B., J. E. Straub, and D. Thirumalai. 2005. Probing the initial stage of aggregation of the A β (10–35)-protein: assessing the propensity for peptide dimerization. *J. Mol. Biol.* 345:1141–1156.
28. Bitan, G., B. Tarus, S. S. Vollers, H. A. Lashuel, M. M. Condron, J. E. Straub, and D. B. Teplow. 2003. A molecular switch in amyloid assembly: met(35) and amyloid β-protein oligomerization. *J. Am. Chem. Soc.* 125:15359–15365.
29. Thirumalai, D., D. K. Klimov, and R. I. Dima. 2003. Emerging ideas on the molecular basis of protein and peptide aggregation. *Curr. Opin. Struct. Biol.* 13:146–159.
30. Klimov, D. K., and D. Thirumalai. 2003. Dissecting the assembly of A β (16–22) amyloid peptides into antiparallel β sheets. *Structure.* 11: 295–307.
31. Caflisch, A. 2006. Computational models for the prediction of polypeptide aggregation propensity. *Curr. Opin. Chem. Biol.* 10:437–444.
32. Cecchini, M., R. Curcio, M. Pappalardo, R. Melki, and A. Caflisch. 2006. A molecular dynamics approach to the structural characterization of amyloid aggregation. *J. Mol. Biol.* 357:1306–1321.
33. Marchut, A. J., and C. K. Hall. 2007. Effects of chain length on the aggregation of model polyglutamine peptides: molecular dynamics simulations. *Proteins.* 66:96–109.
34. Dokholyan, N. V. 2006. Studies of folding and misfolding using simplified models. *Curr. Opin. Struct. Biol.* 16:79–85.
35. Nguyen, H. D., and C. K. Hall. 2004. Molecular dynamics simulations of spontaneous fibril formation by random-coil peptides. *Proc. Natl. Acad. Sci. USA.* 101:16180–16185.

36. Cecchini, M., F. Rao, M. Seeber, and A. Caflisch. 2004. Replica exchange molecular dynamics simulations of amyloid peptide aggregation. *J. Chem. Phys.* 121:10748–10756.

37. Baumketner, A., and J. E. Shea. 2006. Folding landscapes of the Alzheimer amyloid-β (12–28) peptide. *J. Mol. Biol.* 362:567–579.

38. Tsai, H. T., C. Tsai, K. Gunasekaran, E. Gazit, and R. Nussinov. 2005. Energy landscape of amyloidogenic peptide oligomerization by parallel-tempering molecular dynamics simulation: significant role of Asn ladder. *Proc. Natl. Acad. Sci. USA.* 102:8174–8179.

39. Mousseau, N., and P. Derreumaux. 2005. Exploring the early steps of amyloid peptide aggregation by computers. *Acc. Chem. Res.* 38:885–891.

40. Melquiond, A., J. C. Gelly, N. Mousseau, and P. Derreumaux. 2007. Probing amyloid fibril formation of the NFGAIL peptide by computer simulations. *J. Chem. Phys.* 126:065101.

41. De Mori, G. M. S., C. Micheletti, and G. Colombo. 2004. All-atom folding simulations of the villin headpiece from stochastically selected coarse-grained structures. *J. Phys. Chem. B.* 108:12267–12270.

42. De Mori, G. M. S., G. Colombo, and C. Micheletti. 2005. Study of the villin headpiece folding dynamics by combining coarse-grained Monte Carlo evolution and all-atom molecular dynamics. *Proteins.* 58: 459–471.

43. Colombo, G., and C. Micheletti. 2006. Protein folding simulations: combining coarse-grained models and all-atom molecular dynamics. *Theor. Chem. Acc.* 116:75–86.

44. Esposito, L., C. Pedone, and L. Vitagliano. 2006. Molecular dynamics analyses of cross-β-spine zipper models: β-sheet twisting and aggregation. *Proc. Natl. Acad. Sci. USA.* 103:11533–11538.

45. Gsponer, J., U. Haberthuer, and A. Caflisch. 2003. The role of side chain interactions in the early steps of aggregation: molecular dynamics simulations of an amyloid–forming peptide from yeast prion Sup35. *Proc. Natl. Acad. Sci. USA.* 100:5154–5159.

46. Zhang, Z. Q., H. Chen, H. J. Bai, and L. H. Lai. 2007. Molecular dynamics simulations on the oligomer-formation process of the GNNQQNY peptide from yeast prion protein Sup35. *Biophys. J.* 93: 1484–1492.

47. Nelson, R., M. R. Sawaya, M. Balbirnie, A. O. Madsen, C. Riekel, R. Grothe, and D. Eisenberg. 2005. Structure of the cross-β spine of amyloid-like fibrils. *Nature.* 435:773–778.

48. Sawaya, M. R., S. Sambashivan, R. Nelson, M. I. Ivanova, S. A. Sievers, M. I. Apostol, M. J. Thompson, M. Balbirnie, J. J. W. Wiltzius, H. T. McFarlane, A. O. Madsen, C. Riekel, and D. Eisenberg. 2007. Atomic structures of amyloid cross-β spines reveal varied steric zippers. *Nature.* 447:453–457.

49. Daura, X., K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren, and A. E. Mark. 1999. Peptide folding: when simulation meets experiment. *Angew. Chemie. Intl. Ed.* 38:236–240.

50. Berendsen, H. J. C., J. R. Grigera, and P. R. Straatsma. 1987. The missing term in effective pair potentials. *J. Phys. Chem.* 91: 6269–6271.

51. Hess, B., H. Bekker, J. G. E. M. Fraaije, and H. J. C. Berendsen. 1997. A linear constraint solver for molecular simulations. *J. Comput. Chem.* 18:1463–1472.

52. Miyamoto, S., and P. A. Kollman. 1992. SETTLE: an analytical version of the SHAKE and RATTLE algorithms for rigid water models. *J. Comput. Chem.* 13:952–962.

53. Berendsen, H. J. C., J. P. M. Postma, W. F. van Gunsteren, A. Di Nola, and J. R. Haak. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81:3684.

54. van der Spoel, D., E. Lindahl, B. Hess, A. R. van Buuren, E. Apol, P. J. Meulenhoff, D. P. Tieleman, A. L. T. M. Sijbers, K. A. Feenstra, R. van Drunen, and H. J. C. Berendsen. 2004. Gromacs User Manual, version 3.2. http://www.gromacs.org/documentation/paper_manuals.php. Accessed April 2008.

55. Scott, W. R. P., P. H. Hunenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Kruger, and W. F. V. Gunsteren. 1999. The GROMOS biomolecular simulation program package. *J. Phys. Chem. A.* 103:3596–3607.

56. van Gunsteren, W. F., S. R. Billeter, A. A. Eising, P. H. Hunenberger, P. Kruger, A. E. Mark, W. R. P. Scott, and I. G. Tironi. 1996. Biomolecular Simulation: The GROMOS96 Manual and User Guide. vdf Hochschulverlag AG, Zurich, Switzerland.

57. Sambashivan, S., Y. Liu, M. R. Sawaya, M. Gingery, and D. Eisenberg. 2005. Amyloid-like fibrils of ribonuclease A with three dimensional domain-swapped and native like structure. *Nature.* 437:266–269.

58 Reference deleted in proof.

59. Melquiond, A., N. Mousseau, and P. Derreumaux. 2006. Structures of soluble amyloid oligomers from computer simulations. *Proteins.* 65: 180–191.

60. Lashuel, H. A., B. M. Petre, J. Wall, M. Simon, R. J. Nowak, T. Walz, and P. T. J. Lansbury. 2002. α-synuclein, especially the Parkinson's disease-associated mutants, forms pore-like annular and tubular protofibrils. *J. Mol. Biol.* 322:1089–1102.

61. Kwan, A. H., R. D. Winefield, M. Sunde, J. M. Matthews, R. G. Haverkamp, M. D. Templeton, and J. P. Mackay. 2006. Structural basis for rodlet assembly in fungal hydrophobins. *Proc. Natl. Acad. Sci. USA.* 103:3621–3626.

62. Thundimadathil, J., R. W. Roeske, H. Y. Jiang, and L. L. Guo. 2005. Aggregation and porin-like channel activity of a β sheet peptide. *Biochemistry.* 44:10259–10270.

63. Zheng, J., B. Y. Ma, C. J. Tsai, and R. Nussinov. 2006. Structural stability and dynamics of an amyloid-forming peptide GNNQQNY from the yeast prion sup-35. *Biophys. J.* 91:824–833.