[Frontiers in Bioscience 13, 4495-4516, May 1, 2008]

Exploring energy landscapes of protein folding and aggregation

Normand Mousseau[1], Philippe Derreumaux[2]

Departement de physique, Universite de Montreal, Case postale 6128, succursale centre-ville, Montreal, (Quebec), Canada H3C 3J7, 2 Laboratoire de Biochimie Theorique, UPR 9080, Institut de Biologie Physico-Chimique and Universite Paris 7 Denis Diderot, 13 rue Pierre et Marie Curie, 75005 Paris, France

TABLE OF CONTENTS

1. ABSTRACT

Human diseases, such as Alzheimer's and Creutzfeldt-Jakob's are associated with misfolding and aggregation of specific proteins into amyloid fibrils sharing a generic cross-beta structure (1). The self-assembly process is complex, but once a nucleus is formed, rapid fibril formation occurs. Insight into the structures of the oligomers during the lag phase, varying between hours and days, is very difficult experimentally because these species are transient, and numerically using all-atom molecular dynamics because the time scale explored is on the order of 10-100 ns (2, 3). It is therefore important to develop simplified protein models and alternative methods to sample more efficiently the conformational space. In the past few years, we have developed the activation-relaxation technique (ART nouveau) coupled to the OPEP coarse-grained force field. This review reports the application of ART-OPEP on protein folding and aggregation.

2 INTRODUCTION

A remarkable property related to amyloid plaque formation is that this process is shared by related-disease proteins (e.g. the Aβ protein for Alzheimer's disease and the prion protein for Creutzfeldt-Jakob's disease) and non-homologous peptides of variable length as short as four amino acids. We can therefore attack the complexity of the aggregation process on two fronts by simulating either low-order species (e.g. monomers and dimers) of full-length proteins or higher-order species (e.g. 10-mers) of short peptides (4). In this review, we report the application of the activation-relaxation technique (ART nouveau) coupled to the OPEP force field on both types of systems and focus specifically on the knowledge that can be obtained regarding the energy landscape of protein aggregation. Because ART-OPEP has advantages and limitations as any method, it is important to show that this protocol provides pertinent information on the structures and folding

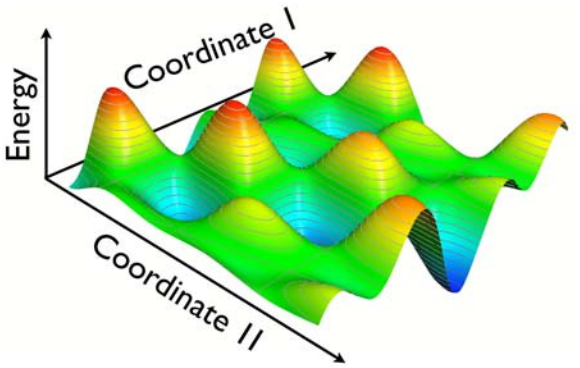

Figure 1. Two-dimensional example of an energy landscape. Each point on this surface corresponds to a unique configuration.

pathways of monomeric peptides studied by more standard methods.

In what follows, we first describe the concept of energy landscape and the details behind the ART-OPEP methodology. Then we focus on the folding simulations of the most studied model, namely a β-hairpin of 16 amino acids in solution. The final part of this review is essentially centered on ART-OPEP simulations of amyloid-forming peptides in settings varying from monomers to dodecamers and the monomer of the full-length Aβ protein. In the following article of this issue (5), we report on the applications of molecular dynamics and replica exchange molecular dynamics coupled to the OPEP force field in order to extract thermodynamics information.

3. THE ENERGY LANDSCAPE PICTURE

3.1. The configurational energy landscape

The concept of energy landscape takes its root in the transition state theory (TST) developed by Eyring (6) and Wigner (7) in the 1930's that associates the kinetic of chemical reactions with the relative free energy of various conformations. It was only 20 years later, however, in the 1960's, that this concept was defined as such. Goldstein, for example, used the energy landscape picture as a tool for understanding glasses, a disordered system characterized by a high level of frustration (8). While the idea of energy landscape can appear as somewhat abstract, it was considerably refined in the 1980's by Stillinger and Weber (9) and really took off, in the 1990, when it was adopted by a large fraction of the glass and the protein communities (for a more complete historical description, see Ref. (10)).

The energy landscape picture represents the projection of the total phase space onto the spatial coordinates as a function, generally, of the configurational energy (Figure 1). The advantage of this representation is that it encompasses not only the structure of a given configuration but how it is connected to other configurations nearby. This gain is in large part abstract, though, since it implies a high-dimensional representation that is not really available to our limited brains.

The difficulty of representing the energy landscape limits considerably the usefulness of this picture for understanding the dynamical properties of complex systems such as proteins and glasses, and most efforts in the last two decades have focused on developing clever projections that offer a usable description of the dominant properties of the energy landscape. Of course, no unique projection of a high-dimensional problem onto a one or two-dimensional space can provide the full picture; some can even lead to conflicting understanding if inappropriate coordinates are selected.

In spite of the challenge associated with finding a meaningful representation, the concept of energy landscape provides a very convenient basis for discussing the kinetic and dynamical properties of complex systems, such as proteins, that cannot easily be analyzed using more standard representations. This was the case, for example, for the solution of Levinthal's paradox, stated in the 1960's (11): how can a protein always fold into its native state if the number of conformations available reaches astronomical proportion? A first convincing answer was provided by Wolynes and collaborators (12), based in large part on theoretical studies of proteins on lattices: natural selection has chosen amino acid sequences that have a strong bias toward folding into their native state or, more concisely, proteins have a funnel-shape energy landscape that brings any initial state rapidly into the same minimum free-energy structure.

The standard funnel picture, presented in the top panel of Figure 2, shows a very deep and steep landscape. While the protein can be trapped briefly in some metastable states, that can be disordered or partially ordered (or even fully ordered but non-native (13)), the overall trend is strongly towards the native state. Of course, this representation is a strong simplification as in reality the funnel is 3N dimensional, where N is the number of atoms. For example, the two-dimensional funnel drawn in Figure 2 does not explain the two-state model observed in thermodynamical measurements of folding and unfolding of simple proteins. Moreover, pathways in a high-dimensional space are much more complex than in 2D and one would imagine an overall funnel where the peptide fails to reach a native state.

To circumvent this representation problem, Becker and Karplus (14) introduced the disconnectivity graph, a new projection from the energy landscape based on a distance both in configuration space (separated by a number of adjacent minima) and in energy. The typical funnel landscape can be represented here as in the bottom panel of Figure 2. The bottom of each leaf indicates the minimum-energy of a given configuration and the position of the branch, at the top of the leaf, the relative activation-barrier to go from one minimum to another connected to the same branch. Although clear spatial information is now lost, information regarding the dynamics is highlighted through the clear appearance of relative barriers, helping identify efficiently the nature of the dominant dynamics with a detailed knowledge of the various activated mechanisms.

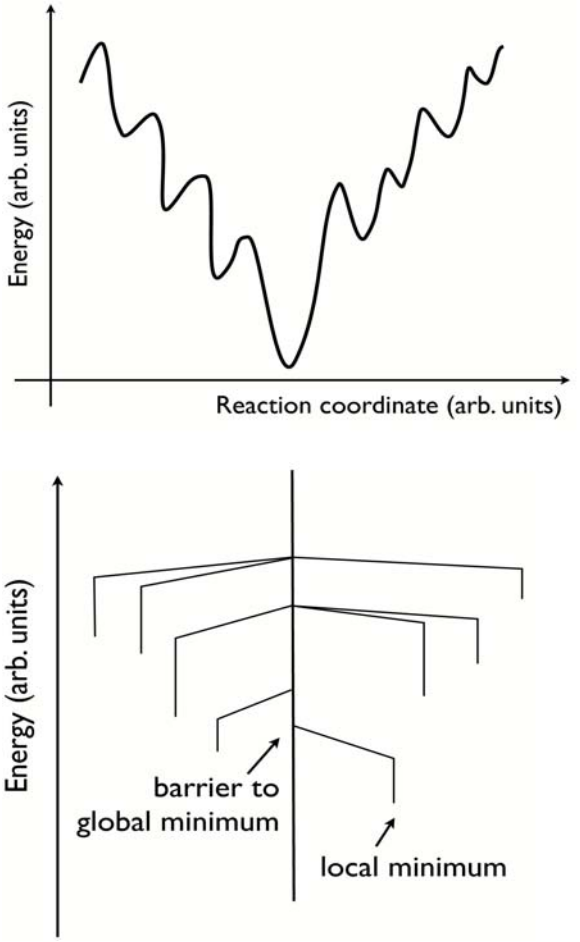

Figure 2. The bottom of each leaf indicates the minimum-energy of a given configuration and the position of the branch, at the top of the leaf, the relative activation-barrier to go from one minimum to another connected to the same branch. Although clear spatial information is now lost, information regarding the dynamics is highlighted through the clear appearance of relative barriers, helping identify efficiently the nature of the dominant dynamics with a detailed knowledge of the various activated mechanisms.

Recently, Rao and Caflisch introduced yet another graph (15) that goes beyond the funnel picture with a representation in terms of a network representing the connectivity between various basins, defined using a standard Cartesian metric. Instead of focusing solely on the native state, this representation helps recreate the ensemble of folding trajectories and emphasizes the existence of a number of quasi-essential intermediate states for folding. While this approach has not been used much since its introduction (16), it provides direct kinetic information that can be missed in simpler representations.

3.2. The free-energy landscape

It is possible to use similar concepts for the free-energy landscape. However, one must be careful because the free energy involves an integral to define the entropy. In the phase space, which defines the free energy, there is no landscape since every point has the same probability (17). As such, the concept of free-energy landscape is only valid when projected onto a relatively low-dimensional space, with the other degrees of freedom integrated, providing sufficient information for defining the entropy. As was pointed out by Krivov and Karplus (18), however, the relevant space of reaction coordinates can still be much larger than two dimensions so that the planar projection, necessary for the visualization, can introduce artifacts leading to an inaccurate description of the dominant free energy barriers.

4. EXPLORING THE LANDSCAPE

One of the major difficulties associated with the use of the concept of energy landscape is the need for constructing it, sampling at least the most important parts. For fast systems, such as liquids, molecular dynamics simulations are sufficient to sample the relevant configurational space. The problem is more challenging for proteins since the fastest ones fold on a μs timescale, and more sophisticated approaches are needed.

These methods can be classified into three groups: (1) biased sampling, (2) reweighting methods and (3) activated approaches. The first set of methods introduces biases in the sampling to force the system to visit less frequent parts of the phase space. These methods include transition path sampling (19), metadynamics (20) and milestoning (21), for example. The challenge here is to find the appropriate reaction coordinate to drive the bias efficiently. In the absence of such a coordinate, these methods become increasingly expensive with more complex systems, when the relevant part of the phase space becomes too small or when the barriers are important compared with the melting temperature. The second set of methods, such as Wang-Landau algorithm (22), replica exchange molecular dynamics (REMD) (23) and, to a lesser degree, the weighted-histogram method (WHAM) (24), focuses on producing a more balance sampling of the energy landscape or, in the case of WHAM, on extracting as much information as possible from independent simulations. While these approaches can help the sampling, they are limited by the efficiency of moving from one state to another and have not delivered as much as we could have expected initially for large and complex systems.

The final set of methods is appropriate for systems with a dynamics dominated by activated mechanisms (21, 25-31). This is the case, for example, of proteins, at least in implicit solvent. Activated methods generate pathways connecting adjacent local minima via first-order saddle points, reconstructing physical pathways as defined by the transition state theory (6). The phase space is therefore reduced to a discrete network composed of local minima connected via activated states, decreasing considerably the number of states to sample (see Figure 3). This network can be generated through continuous trajectories such as with the activation-relaxation technique (ART nouveau) (27) or assembled through a random and biased search (32). Activated methods are not sensitive to the height of the barriers and can therefore move rapidly

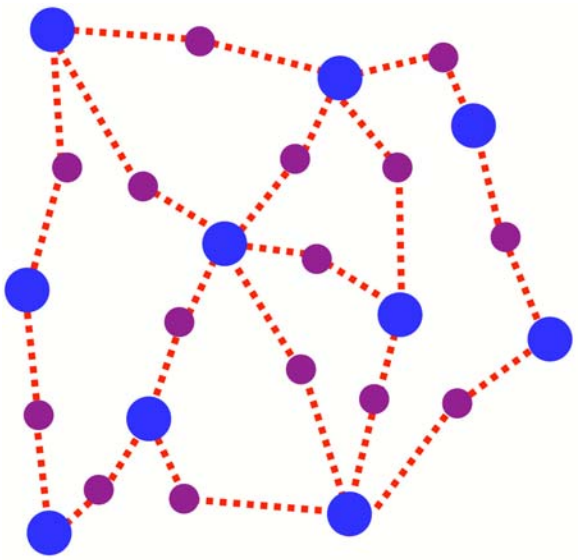

Figure 3. Activated methods, such as ART nouveau, sample a much reduced configurational space consisting only of local minima (large blue dots) connected by pathways defined by the position of the transition state (small purple dots) explaining, in part, the efficiency of these methods.

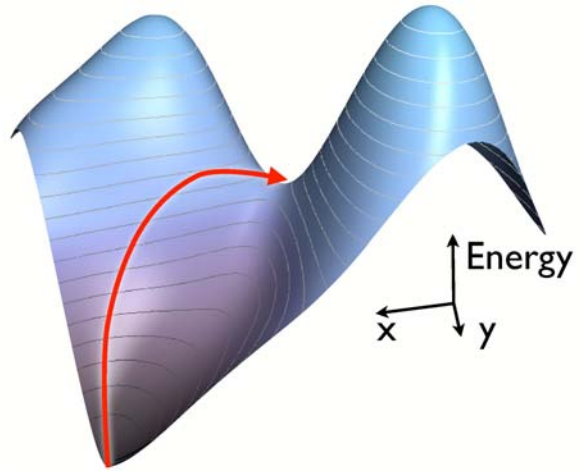

Figure 4. Cartoon representing the first steps of an ART nouveau event. Starting from a local minimum, the structure is pushed in a random direction until a direction of negative curvature appears. The structure is brought along this direction, minimizing its energy in the perpendicular hyperplane until it reaches an adjacent first-order saddle point.

through the configurational space. They lack a proper thermodynamical basis, however, which restricts their application. In spite of these limitations, activated approaches have provided many insights in protein folding.

4.1. The activation-relaxation technique

There exist many methods for identifying transition states. Most of them, such as the rubber band method of Elber and Karplus (28), the variational Verlet algorithm (29), the nudged elastic band method (NEBM) (30) or milestoning (21) require the knowledge of both the initial and final state in addition to a rough guess of the transition. Moreover, the initial and final states must be nearby, separated by at most a few intermediate states. Longer pathways are unreliable since the resulting trajectory is strongly dependent on the initial guess.

Of course, for proteins, we do not know the intermediate states and we must turn to methods that can identify transition states without *a priori* knowledge of the final state, such as ART nouveau (27) or similar methods such as the dimer (31) and the eigenvector-following methods (26, 33). While these methods are basically equivalent (34), we focus here on ART nouveau that was developed by one of us.

An ART nouveau (simply called ART, for shortness, in the rest of this review) event can be divided in four steps (Figure 4):

(i) starting from a local minimum, the system is pushed slowly in a random direction; at each step, we evaluate the value of the direction of lowest curvature (the lowest eigenvalue of the Hessian matrix); when the lowest curvature becomes negative, we stop the push in this random direction;

(ii) the configuration is then displaced along the direction of negative value while the energy is minimized in the orthogonal directions;

(iii) when the total force reaches a value close to zero, the systems has converged onto a transition state (first-order saddle point);

(iv) the system is then moved over the saddle point and relaxed into a new minimum, completing the event defined by the initial minimum-transition state-final minimum.

Let us review these steps in some details. For a small peptide, up to about 15 amino acids, the initial displacement is taken as a random vector in the (3N-6) dimensional space where rotational and translationnal degrees of freedom are removed; all atoms are therefore displaced, each in a random direction. For a larger protein or ensemble of peptides, the initial random displacement is generally made on a subset of all atoms, to allow for a faster sampling. In all cases, the initial displacement excludes moves along the strong forces associated with covalent bonds and bond angles since these directions are strictly harmonic by definition.

Once the initial direction is set, the protein is then deformed very slowly, displacing all atoms by a total of about 0.4 Å at each step. After each displacement, the protein conformation is relaxed in the hyperplane perpendicular to the move away from the initial minimum.

This step prevents building too much stress in the system during the first phase of activation for a better control of the activation. After each move, the value of the lowest curvature is calculated, from the force field, using the Lanczós algorithm. Starting with a random vector, we apply the algorithm to construct a small tridiagonal matrix that we can diagonalize, obtaining value of the lowest curvature and the corresponding eigenvector. Reusing the result as a starting point for another application of the Lanczós method, we converge rapidly onto a stable eigenvector that evolves slowly as iterations progress.

The presence of a first-order saddle point is signaled by the appearance of a negative eigenvalue in the Hessian. To prevent the system from going back into the harmonic basin, with only positive curvature, we set a threshold below zero for starting the convergence to the transition state. This threshold depends on the potential and the system under study. The choice of this parameter is not very critical and it affects the success rate of the activation process but not the type of events found (27).

Once the edge of the harmonic basin is reached, the system is pushed along the direction of negative curvature, away from the initial minimum. While the initial push might involve only a subset of all the atoms, here, the whole system is free to move. After each step along the eigenvector, the Hessian is diagonalized again, with Lanczós algorithm, and the system is relaxed in the hyperplane perpendicular to this direction using an adaptative steepest-descent algorithm. Iterations stop when the absolute value of the force perpendicular to the ridge falls below typically 0.1 kcal/ (mol·Å) and the total force, below 1.0 kcal/ (mol·Å), meaning that a new transition state has been identified, or when the lowest curvature become positive, in which case we start again from the initial minimum.

From the saddle point, the structure is pushed away from the initial minimum along the direction of negative curvature and then relaxed, using damped molecular dynamics, into a minimum. Most of the time, the final state differs from the initial configuration and an event has been generated. In some case, however, the saddle point identified corresponds to a shoulder in the potential energy landscape and the structure goes back to the initial minimum.

Once an event is generated, it can be used in multiple ways. If one is interested in characterizing the energy landscape around a single minimum, then this event is simply stored and a new event is started from the same point (27, 35). It is also possible to perform an accept/reject, using the Metropolis criterion, for example,

$$p_{accept} = \min\left[1, \exp\left(-\Delta E / k_B T\right)\right]$$

with either the activation barrier (16) or the asymmetry energy (the energy of the final state minus that of the initial minimum) (27). In general, we have used the latter criterion as we would need to know the complete distribution of barriers around a local minimum in order to use the activation barrier meaningfully. By using only the asymmetry between adjacent minima, the method ensures that at least each configuration is weighted appropriately with respect to the others.

Wales and collaborators have used events generated with a similar method to construct a sparse network similar to that of Figure 3, from which it is possible to extract the fastest folding pathway, for example (36, 37). This approach, called *discrete path sampling,* has been used to study folding of a β-hairpin, for example (37). While discrete path sampling provides very useful kinetic information, the numerical effort needed to extract pathways makes it difficult to assess the impact of a missed barrier or varying prefactors that could modify the dominant folding pathway. Moreover, the quantity of events needed to construct a sufficiently connected network increases rapidly with the number of degrees of freedom so that this method is difficult to apply to larger systems.

In this Review, we focus on the more straightforward approach to sampling the landscape using a Metropolis criterion based on the energy difference between the initial and the final minima at each event. By doing so, we generate a fully connected walk through the energy landscape. Because this constructed landscape is reduced to a discrete set of local minima, ART does not include entropic contributions. The temperature used in the Metropolis accept/reject procedure is therefore unphysical and is generally set in such a way that the trajectories do not remain trapped in metastable states.

The main reserve about this approach is that the ART-constructed trajectories do not respect detailed balance nor belong to a well-defined thermodynamical ensemble. This is due to the fact that we do not know what the bias of ART is for finding one nearby saddle point instead of another. This criticism is valid. However, numerous tests and comparisons, many described below, find a strong similarity between the ART folding trajectories and those obtained from more standard techniques such as molecular dynamics.

We have investigated this question in a comparative ART study of the β-hairpin, using two different Metropolis criteria: one based on the barrier height and the other on the standard asymmetric energy (16). As can be seen from Figure 5, which superposes some of the intermediates explored by two runs using different Metropolis criteria, we find results in agreement with what can be expected from the network picture of folding observed by Rao and Caflisch (15). Specifically, there exist a number of well-connected intermediate structures, or hubs, that seem to connect various parts of the configurational space. Irrespective of the method used, these intermediate structures will be found as long as the pathway generated is continuous. Even though the kinetic varies considerably from one specific folding trajectory to another, the overall picture remains unchanged.

ART-generated trajectories are possible pathways that should capture the qualitative feature of folding or

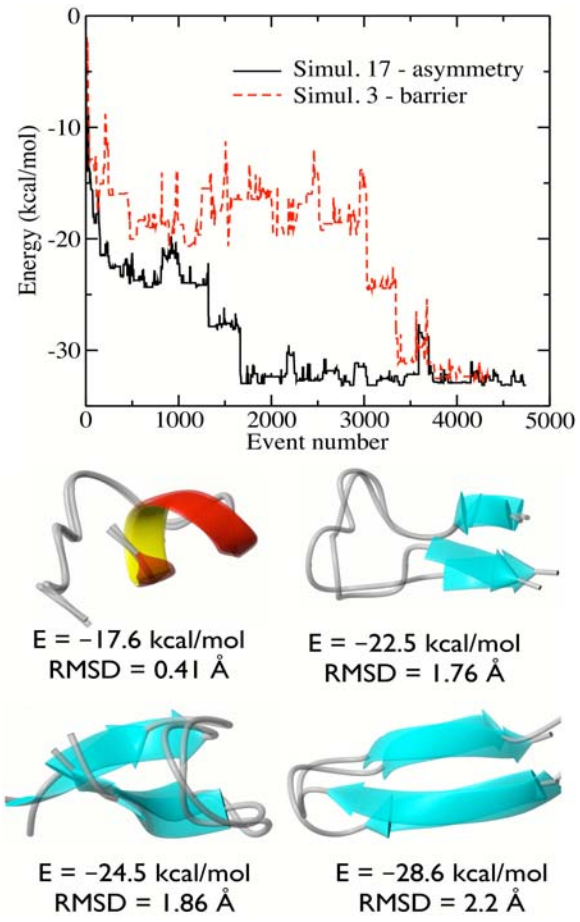

Figure 5. Top: Two folding trajectories using a different Metropolis accept/reject criterion based on either the energy barrier or the energy difference between the final and the initial minima. Bottom: Superposition of conformations found in these two sets of simulations. For each pair of conformations, we indicate the energy and the RMSD. Even though the detail of folding depends on the criterion used, we find the same intermediate structures.

aggregation since these are dominated by well-connected hubs. Because the relative weight of one hub over the other is subtle, the relative probability of the various folding mechanisms identified by ART cannot be used. While this might be construed as somewhat of a drawback, comparisons with various experiments and other simulations show rather that the trajectories identified with this method cover a wide range of specific conditions and can therefore be used to understand experimental setups that cannot easily be simulated, for example.

4.2. The OPEP potential

The OPEP (Optimized Potential for Efficient prediction of Protein Structure) is a generic force field that can be used for any amino acid sequence, with L- or D-amino acids. It is based on a coarse-grained representation of the amino acids, where all backbone atoms are included and all side chains are represented by a specific bead (38). While such a reduction in the number of degrees of freedom was proposed in 1976 by Levitt and Warshel (39), the originality of OPEP lies in a good compromise between energy accuracy, structural precision and computational cost (40). A force field with implicit solvent and coarse-graining of the side-chains cannot capture all the details of the interactions between the side-chains and between the side chains and the solvent, as the all-atom molecular mechanics AMBER force field (41) or all-atom spectroscopic force fields (42). However, we find it possible to design an energy function that discriminates native from non-native structures on an ensemble of 30 proteins (43), and predict, using Monte Carlo simulations, lowest-energy structures consistent with experimental data (44, 45). Full details of the energy function balancing side-chain – side-chain interactions (with both hydrophobic and hydrophilic components), dihedral angles of the backbone and hydrogen-bonding interaction can be found in Ref. (43). As discussed in this review and the next article of this series, OPEP has been coupled to ART nouveau (46), molecular dynamics simulations (47) and REMD simulations (48). Of interest for this review is that the OPEP-generated free energy surface of small proteins is fully consistent with experimental data, independently of the starting conformation used, providing further support of the capability of the OPEP force field to reproduce thermodynamical data as well (48).

5. THE β-HAIRPIN

The first step towards developing a full understanding of protein folding is to identify the mechanisms by which the secondary elements, such as α-helices and β-sheets, form. With dominant local interactions, α-helices form rapidly, within a few hundred ns for a 21-residue model (49-51). Beta structures are more difficult to create as they involve non-local interactions. This difficulty is observed even in β-hairpins, where the folding time can be 10 or more times slower than the helices, reaching many μs.

One of the most studied model is the second β-hairpin (GEWTYDDATKTFTVTE) of the domain B1 of protein G , which has be the subject of an extended range of theoretical (52, 53) and experimental studies (54, 55). With a melting point near room temperature, this peptide is not stable enough in solution for high-resolution NMR determination but it adopts, with a significant probability, a β-hairpin conformation near that found within the full-length protein G (54). Moreover, early fluorescence experiments show that this β-hairpin folds, in isolation, with a time constant of 6 μs with a kinetic consistent with that of the two-state folding model, i.e., with only two dominant states: folded and unfolded (55).

If experiments can often offer the details of the native structure, it is much more difficult to obtain information about the folding mechanisms and, from there, a clear explanation of the important folding time difference between the hairpin and a helix (56, 57). There has been significant progress, however, in the recent years. Ising-like models derived from NMR and temperature-jump

experiments (55, 56), that consider native interactions only, suggest that the zipper-mechanism explains much of the experimental data (56). In agreement with this picture, mutations and ϕ-value analysis also point to the importance of forming a loop as the rate-limiting step in the folding of a β-hairpin (58, 59). Dyer *et al.*, using FTIR and T-jump relaxation, conclude that the loop rearrangement could be the rate-limiting (60).

Much of the experimental picture depends finely on the models used to explain the data. To validate those, we must turn to simulations in order to follow precisely the sequence of events leading to the folded state. But 6 μs is a long time for computer simulations, and many new approaches, attempting to overcome the time limitation, have been used to try to solve this problem, transforming this fundamental question into a test-bed of new computational methods applied to the protein folding problem. The hairpin was therefore studied under a wide range of representations that go from all-atomic empirical force-field with explicit (61-73) and implicit solvent (18, 71, 74-76) to coarse-grained off-lattice potentials with implicit solvent (52, 53, 77-79), biased-potentials (80, 81) and minimal models (82).

The simulation methods are as varied as the molecular descriptions with straightforward unfolding (83-85) and folding molecular dynamics (66, 74, 86), activated methods (52, 53, 73, 75), replica exchange molecular dynamics (64, 72, 87-91), transition path sampling (92), distributed computing (71), self-guided molecular dynamics (93), multicanonical Monte Carlo (79) and other free-energy calculation methods (65, 94, 95), and statistical models (82).

Initial simulation work had identified two dominant folding mechanisms: the *zipper out* and the *zipper in* models. In the zipper out mechanism, the turn forms first, with the H-bonds propagating from there. This mechanism was first proposed based on an Ising-like model (55, 82) and observed in lattice Monte Carlo simulations as the dominant model (80). A similar mechanism, controlled by the hydrophobic core near the turn, which forms first and propagates from there, was identified by the same group (80) as well as by off-lattice model Langevin (81) and all-atom MD simulations (96).

The zipper in mechanism is dominated by the hydrophobic core: the two extremities come in proximity pulled in by the hydrophobic core, the H-bonds form nearby and propagate inward, towards the turn. This mechanism was identified in multicanonical MC (79), all-atom unfolding MD (83-85), Gō-based (86), REMD (64) and distributed MD simulations with implicit solvent (71).

ART-OPEP, while suffering from the limitations described in section 4.1, could provide a more comprehensive picture of folding (53, 97). Eighty-two folding simulations were launched at 300 K, from a fully extended conformation, using different initial random seeds. Three sets of parameters were used: the standard OPEP force field (52 runs), a modified OPEP potential (20 runs) and a Gō-like potential (10 runs). From all these 82 trajectories, counting between 4000 and 9000 trial events, 36 reached the native state, including 16 with the standard OPEP parameters. Here the definition of the native structure is very strict: less than 2.0 Å RMSD from the hairpin structure within protein G (PDB code 2GB1), a well defined hydrophobic core and six H-bonds formed. Focusing on the trajectories leading to the native state, Wei and collaborators identified three folding pathways (53, 97): the two pathways already observed, the *zipper out* (see Figure 6 (a) and *zipper in (*6b) mechanisms, and a third one, the *reptation mechanism*. This latter move occurs when a loop forms at the wrong place and the peptide chain is characterized by a network of non-native H-bonds. Slowly, however, following fluctuations, one strand "walks" over the other, in a reptation move, until the two strands are aligned and the peptide forms the native state (6c).

These results were important for two reasons. First, they demonstrated, for the first time, the predictive power of ART-OPEP, showing that even though the algorithm does not have detailed balance, the folding trajectories are physically possible and overlap with other better characterized algorithms such as molecular dynamics. Second, the reptation mechanism underlines the importance of non-native interactions during folding, and puts into question the use of unfolding and Gō-models where non-native interactions play either a small role or none at all.

In the recent years, experimental and numerical analysis of the folding pathway of various β-hairpins has continued. All-atom MD and REMD simulations found that asymmetric hairpins, stabilized by non-native H-bonds, are common in hairpins varying in length between 9 and 16 residues (73, 93, 95, 98, 99). The identification of the reptation move is more difficult, especially with techniques such as REMD that provide only thermodynamical information. Many works have studied shorter β-hairpins, counting as little as 9 residues. Because of the tight turn and the shorter branches, these hairpins seem to show a narrower range of mechanisms. A detailed analysis of an extensive set of MD folding trajectories of the trpzip2 peptide (12 residues) failed to reveal any reptation moves (70). Chen and Xiao, however, studied the 9-residue peptide designed by Blanco *et al.* (100) with an implicit solvent and found folding mechanisms that incorporated a mixture of the mechanisms discussed above, including reptation (74). Imamura and Chen studied the impact of the position of the hydrophobic interactions on the various folding mechanisms using a coarse-potential with MC (101). They found that if the hydrophobic core is placed at the turn, then the dominant mechanisms should be symmetric, including the zipping out, zipping in and middle out. According to these authors, the reptation move, which they found in their simulations, does not depend on the position of the hydrophobic core.

It is not clear how to reconcile all these different results. Restricting the discussion to the β-hairpin of protein G, the exact nature of the folding pathways remains debated today in part because it is still difficult to obtain

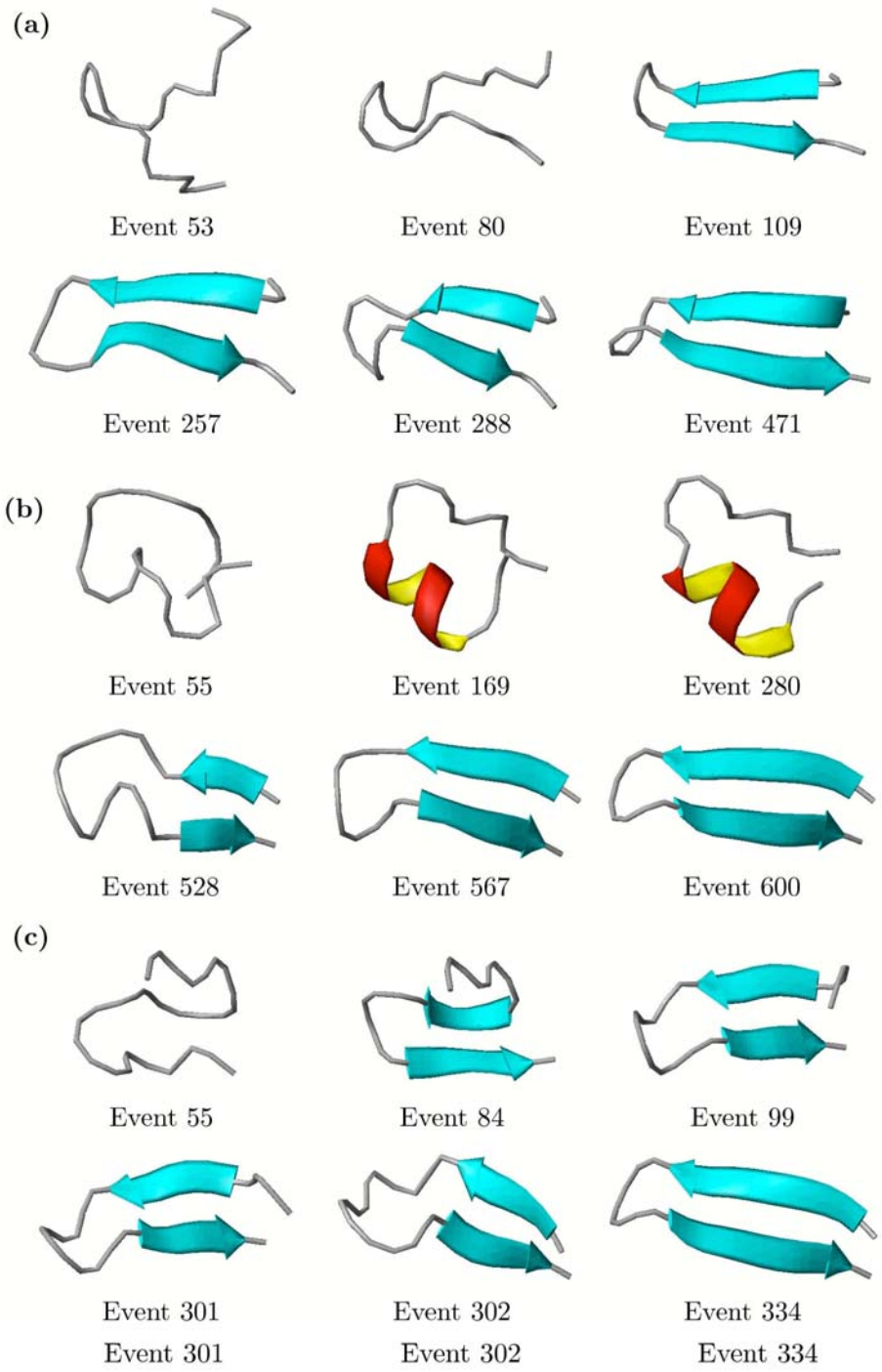

Figure 6. (a) A representative folding trajectory of the *zipper out* mechanism (mechanism I) (only accepted events are counted): The turn forms first (event 53), which forces the two branches to come together (event 80) and allows the formation of the H-bonds from the center to the end points (events 109 to 471). (b) A representative folding trajectory of the *zipper in* mechanism (mechanism II): the peptide collapses in a random coil (event 55), then forms partially a helix (events 169 and 280), as it stretches out, the end points meet and form native H-bonds (events 528), the H-bond network propagates towards the turn (events 567 and 600). (c) A representative folding trajectory of the *reptation* mechanism (mechanism III): the peptide collapses into a random coil with two turns (event 55), a short non-native hairpin forms (event 84), one branch slides over the other, moving the turn in the direction of the center (events 99 and 301), the two branches align themselves (event 302) and the native hairpin forms (event 334).

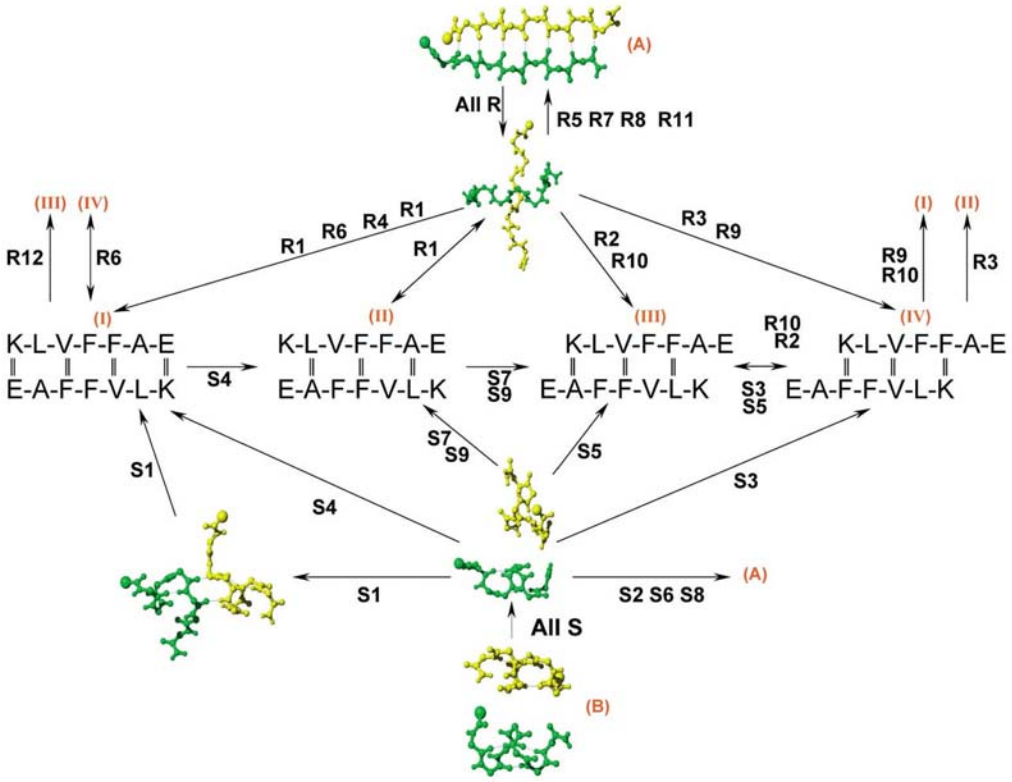

Figure 7. The various pathways identified for a dimer of Aβ (16-22). The arrows indicate the direction of the pathway and the double lines indicate hydrogen bonds. The runs in which a given transition observed are indicated by either Rx or Sx, with x a number. The initial states are indicated in (a) (a parallel β-sheet) and (b) (two anti-parallel α-helices). Except for the α-helical conformations, which are never revisited, pathways can enter and leave all states.

good statistics with all-atom potentials and explicit solvent. For example, a recent all-atom MD simulation, using GROMOS96, found only 3 folding events in a 278-ns simulation (72) while a recent REMD study with AMBER and implicit Poisson-Boltzmann solvent model identified a single pathway corresponding to the zipper in mechanism (76). While ART-OPEP could not provide the definitive answer, its efficiency allowed enough statistics to identify a new folding mechanism that was observed directly and indirectly afterwards by standard simulation techniques. These results on the β-hairpin demonstrate that the network nature of the protein energy landscape makes ART-OPEP valuable to generate relevant folding and certainly aggregation pathways.

6. AGGREGATION OF SHORT PEPTIDES

The discovery of short peptides with 4-8 amino acids forming amyloid fibrils *in vitro* (102-104) was a major advance for the computational community as it allows the study of the first steps of amyloid-forming protein aggregation. Simulations remain challenging but we are beginning to see various approaches converging towards a unified picture.

The favorite peptide for computational studies is Aβ (16-22), of sequence KLVFFAE, corresponding to the central hydrophobic core of the amyloid-β protein associated with Alzheimer's disease. Following the pioneering solid-state NMR study carried out by Balbach *et al.* (105), this peptide was studied using MD with explicit (106-109) and implicit (110) solvent, REMD with explicit solvent (111), Monte Carlo with reduced representations (87, 112, 113) and ART-OPEP (114, 115). Stability MD simulations support the experimentally-derived anti-parallel organization (108). This orientation is also confirmed in the biased MD aggregation study of three monomers (109). Interestingly, the latter study suggests the obligatory presence of α-helical intermediates on the aggregation pathways (109).

We have generated ART-OPEP simulations on the monomer, dimer (115) and trimer of Aβ (16-22) (114), without introducing any bias towards a specific structure. While the structure of the monomer has not been characterized experimentally, ART-OPEP simulations suggest a random coil structure, in agreement with other simulations (109, 113). For the dimer, simulations are started from either a parallel β-sheet or two α-helices with a Metropolis temperature selected such that the ordered structures are visited while ensuring an overall satisfactory sampling of the conformational space. The lowest-energy minimum is anti-parallel in character with a H-bond network matching the solid-state NMR pattern of the fibril (pattern I in Figure 7) Analysis of the low energy structures shows that the system can populate many local minima

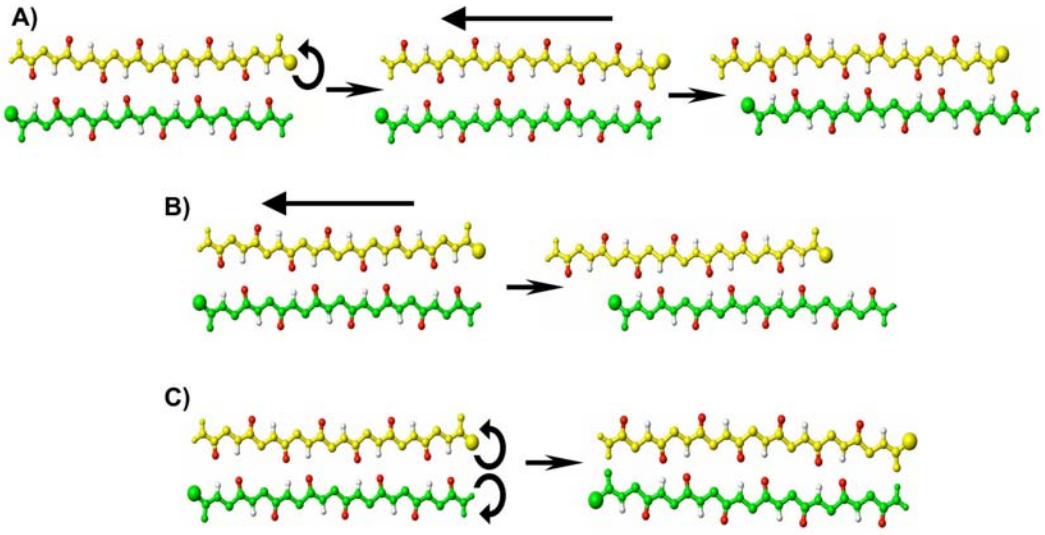

Figure 8. Basic rearrangement mechanisms observed by ART-OPEP for motion without full detachment. (a) The transition between patterns I and III (previous figure) takes place by a rotation of one chain with respect to the other following by a reptation step. (b) The transition between patterns I and IV was seen to occur by a two-step reptation move. (c) We also observe pathway connecting pattern I to pattern II going through a rotation of both chains. Figure 9. All-atom β-sheets of higher energies formed by the trimer of Aβ (16-22), as determined by ART simulations. Registries A to E are antiparallel in character, whereas registry F mixes parallel and antiparallel strands. The superposed structures in A, B, and F result from independent runs indicating that theses structures can be found through many trajectories.

with anti-parallel orientations of the chains and distinct H-bonding patterns as is shown in Figure 7, and notably the second NMR pattern of Balbach *et al.* (82). The existence of these metastable dimeric states with various H-bond registries was confirmed by all-atom MD simulations (110).

As seen in Figure 7, ART-OPEP trajectories provide information on the various pathways. The dimer moves through intermediate anti-parallel states with various registries as well as from anti-parallel to parallel organization without the need for α-helical intermediates. Even though the parallel β-sheet, corresponding to the conformation adopted by the full-length Aβ in the fibril, has an energy 3 kcal/mol higher than the anti-parallel sheet, it is sampled in many runs, independently of the initial conformation. This result has also been observed on the same dimer using REMD simulations with the AMBER94 force field (111) and on the trimer of GNNQQNY, using CHARMM 19 where various orientations appear as metastable states (116). The reference in the orientation is therefore simply a matter of balancing the various contributions to the configurational energy.

Analyzing the rearrangement mechanisms within the two-stranded β-sheet, we find that the two chains can detach and reattach. But they can also move by a reptation mechanism similar to that of the β-hairpin, a rotation of one chain with respect to the other or by a combination of both rotation and reptation moves (see Figure 8). Interestingly, the reptation mechanism was also observed on aggregates of Aβ (16-22), by FTIR experiments (117), confirming our numerical results.

ART-OPEP simulations of the Aβ (16-22) trimer support the results observed on the dimer: the NMR-derived anti-parallel β-sheet with the 16+k↔22-k registry is one of the three lowest-energy structures, (see Figure 9 (a)). The two other lowest-energy structures have one strand shifted with respect to the native state (Figure 9 (b) and (c)). (105). It is important to note that these three predicted conformations are also found to be stable in explicit solvent MD simulations at 330 K for at least 7 ns, using the GROMOS96 force field.

Other groups have also obtained results similar to those of ART-OPEP. Favrin *et al.* manage to observe the aggregation of dimers, trimers and hexamers using Monte Carlo with moves restricted to ϕ and φ angles, reducing considerable the phase space available, but by a method completely different from ART-OPEP, and yet obtaining very similar results (112). The structures observed for the dimer and trimer are not confined to Aβ (16-22), however. Lei *et al.*, in all-atom MD simulations of the dimer of NVHTLSQ, a seven-residue peptide from human β2-microglobulin, also find many metastable β-sheets with various orientations and bonding networks (118).

But ART-OPEP simulations also sample higher energy structures, characterized by very diverse H-bonding patterns and monomer orientations (see Figure 9 d-f). Since the fibril-compatible state is already populated to some extent within a trimer, it is possible that some of the other higher energy trimeric states might share structural similarities with the off-pathway intermediates observed experimentally This point remains to be determined, however.

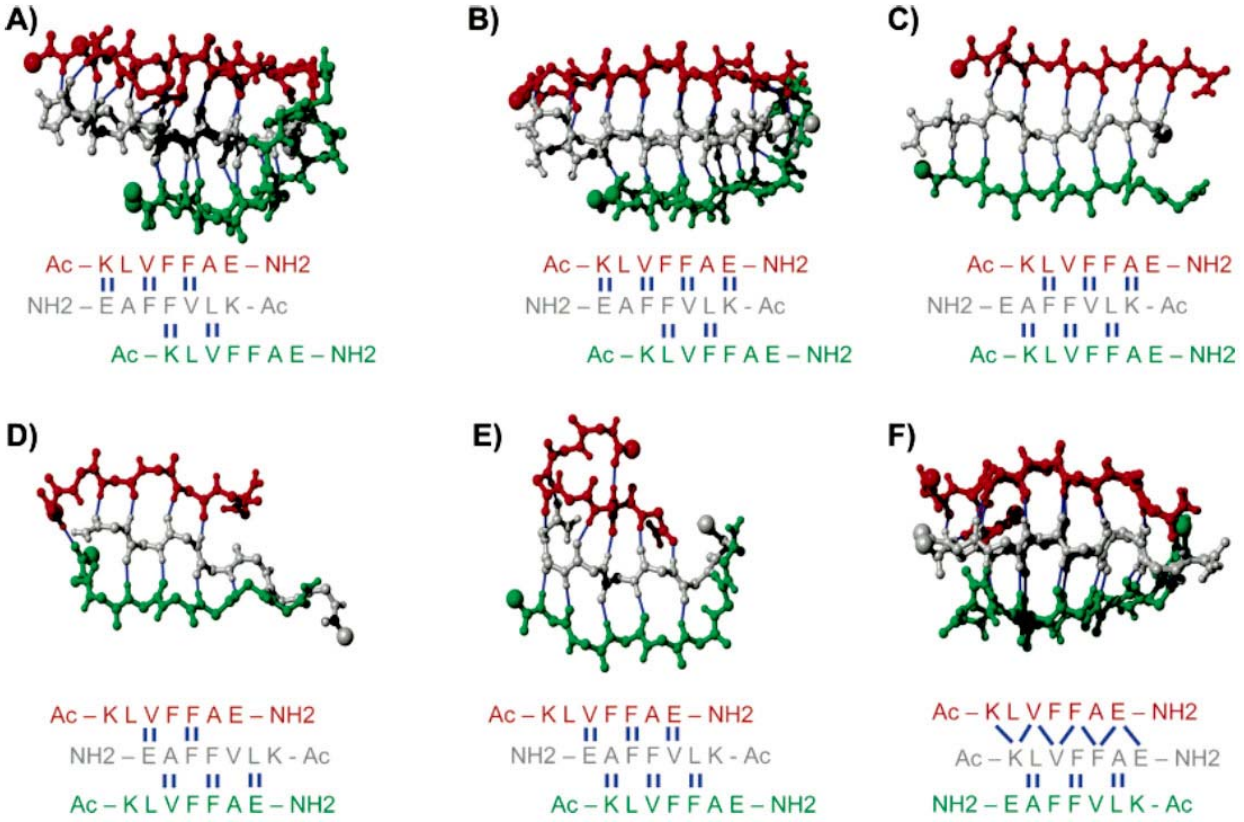

Figure 9. All-atom β-sheets of higher energies formed by the trimer of Aβ (16-22), as determined by ART simulations. Registries A to E are antiparallel in character, whereas registry F mixes parallel and antiparallel strands. The superposed structures in A, B, and F result from independent runs indicating that theses structures can be found through many trajectories.

As more chains are added it becomes much more difficult to generate structured aggregates for this sequence. For example, recent replica exchange Monte Carlo (REMC) simulations using an all-atom implicit solvent potential on six Aβ (16-22) peptides only find amorphous aggregates with little β-sheet contents (113) even though ordered structures remain stable over many hundred ns (106, 107).

Even with a coarse-grained force field such as OPEP, it is difficult to form well-ordered structures for six Aβ (16-22) peptides because the number of attractive intermolecular interactions (due essentially to the presence of five hydrophobic amino acids) can stabilize many disordered structures. To reduce this number, we have turned to KFFE, the shortest peptide known to form amyloid fibrils *in vitro* (102), and studied tetramers (119), hexamers (120), heptamers (121), and octamers (122). As we discuss below, the characterization of these four oligomeric sizes provide a fairly complete picture of the first steps of aggregation for short chains, up to maybe 10 residues.

For simplicity, we focus here on the hexamer of KFFE. As for Aβ (16-22), unbiased ART-OPEP simulations are launched from a random orientation of isolated, unfolded monomers, ensuring that the initial state does not play any role in the formation of specific ordered structures. Since the peptide is very short, it cannot collapse upon itself, leaving the side-groups and H-bonds available for intermolecular interactions. This accelerates the formation of ordered structures. For example, the hexamer assembles into ordered structures within typically less than 20 000 attempted events, fewer than 10 000 accepted events.

All 10 runs locate ordered structures which can be categorized into three generic families, following relatively well-defined aggregation pathways: 1) Two runs lead to a double-layer three-stranded β-sheet (see Figure 10, panel a). 2) Four runs lead to a tetramer-plus-dimer structure with the dimer often positioned in a direction orthogonal to the axis of the tetramer (panel b). 3) Four runs go to a six-stranded β-sheet folding onto itself into an open β-barrel (panel c and d). Although the detailed structure of the KFFE amyloid fibril has not been characterized experimentally, it is interesting to note that structure 1, a double-layer three-stranded β-sheet, displays intra-sheet and inter-sheet distances, respectively 4.5. Å and 10-12 Å (Figure 11), in agreement with the x-ray diffraction measured distances of 4.7. Å between strands and 10-11 Å between sheets in amyloid fibrils (105). While the intra-sheet orientation is mixed, suggesting that more ordered states are possible, these results indicate that

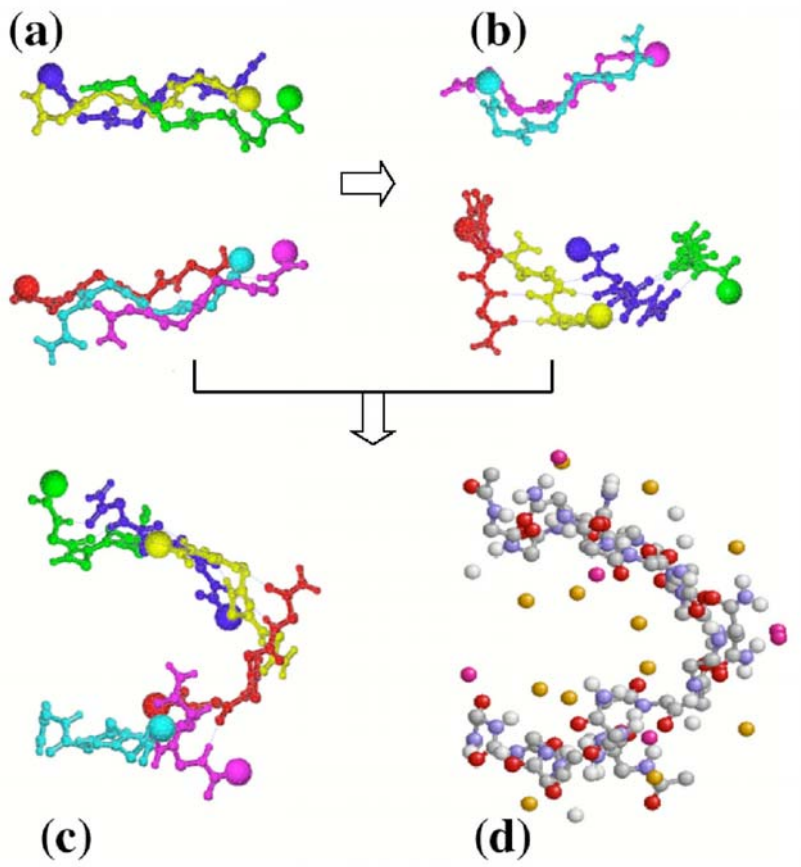

Figure 10. The three lowest-energy structures found from unbiased ART-OPEP simulation of a hexamer of KFFE. (a) The double-layer three-stranded β-sheet; (b) the double-layer tetramer-dimer structure; and (c-d) the β-barrel hexamer. In (c), only the main chains are indicated while in (d), a ball-and-stick representation with the Phe residues, labelled F, indicated as spheres. The arrows indicate pathways observed in the simulations.

even the hexamer of KFFE is compatible with the amyloid structure. This is not so surprising given the presence of the two phenylalanine amino acids (104).

While ART does not provide any information regarding the dynamics and does not have detailed balance, our experience with the hairpin and the Aβ (16-22) dimer and trimer shows that the aggregation trajectories generated with ART-OPEP are qualitatively representative of what is observed by other methods, when these results are available. For the hexamer of KFFE, these trajectories are most interesting. Figures 12 and 13 show the aggregation trajectory leading to the formation of the double-layer three-stranded β-sheet. Figure 12a follows the energy as a function of the number of accepted events. The energy drops rapidly, within 1000 events, before reaching a plateau (between event numbers 1000 and 2000) where a number of important rearrangements take place, as can be observed in the the relative orientation of a few dimers (panel b). Once all chains are well-stretched, they only have to align themselves to fall into place and the energy drops rapidly between events 2000 and 3000. Because ART continues to search for new local conformations, however, the system is not confined to this low-energy structure and we see that it continues to evolve after reaching the run's energy minimum. Figure 13 shows the same trajectory through a sequence of cartoons that help picture the aggregation process. In particular, we see that once the two three-stranded β-sheets are formed, there is still some considerable motion of one sheet with respect to the other, including a orthogonal positioning (panel e) before the two fall into place.

Similar intermediates were observed by our group in the simulations of tetramers (119), heptamers (121) and octamers (122) of KFFE, and can be reduced to three dominant ordered structures in addition to the amorphous state: (1) the bi-layer β-sheet; (2) the β-barrel, which protects its hydrophobic residues from contact with the solvent; and (3) the single-layer β-sheet stabilized by a monomer, dimer or more, laying perpendicular to the sheet, such as is observed in Figure 13 (e). These three structures have been observed, all or partially, in other simulations. Stability studies, by Nussinov *et al.* and other groups, showed that the multi-layer β-sheet is very stable for a large number of sequences (108, 123-126). Many other

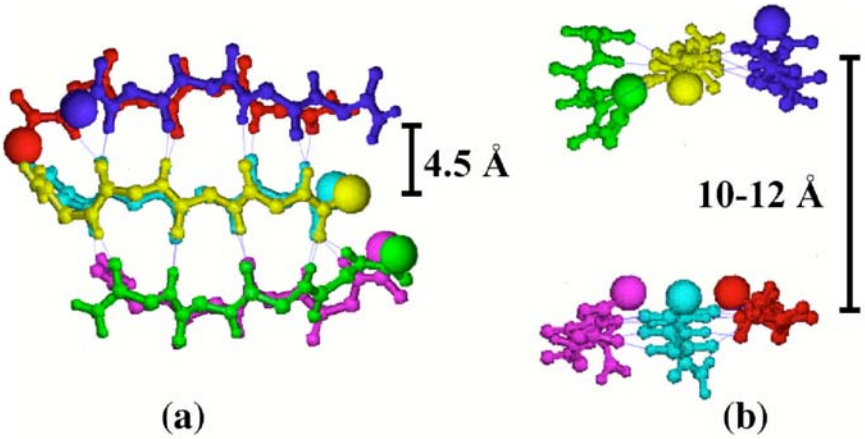

Figure 11. Two different views of one double-layer three-stranded β-sheet hexamer of KFFE generated from a random conformation using ART-OPEP. The two layers are perfectly parallel and show a parallel inter-plane organisation with a mixed in-plane orientation. (a) View perpendicular ot the fibril axis; (b) view perpendicular to the fibril axis. The dotted lines indicate the position of H-bonds. The N-terminal of each chain is indicated by a larger sphere. The distances indicated are in full agreement with the x-ray diffracted characteristic lengths associated with an amyloid fibril.

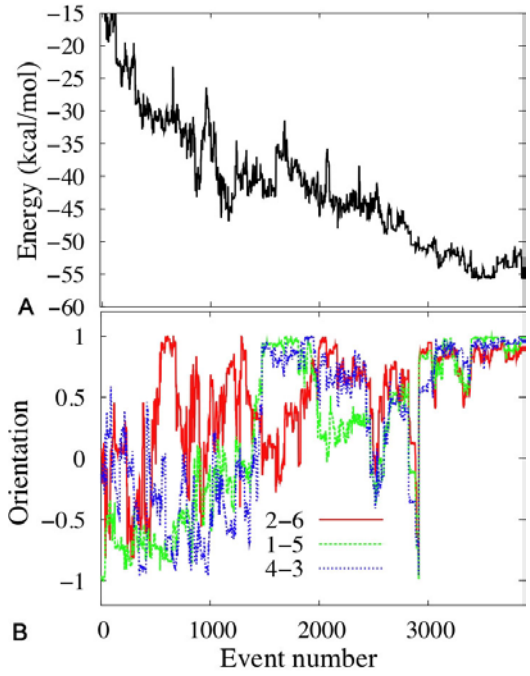

Figure 12. Characterization of the self-assembly pathway as a function of accepted event leading to the formation of the double-layer hexamer shown in Figure 16. (a) Evolution of the energy as a function of accepted event. (b) The intersheet orientation , calculated using the scalar product between vectors defined between the two end points of each monomer, for three pairs of chains found in the final structure.

groups observed structure, often a single sheet for 3 or 4 chains, starting from random conformations. For example, Favrin *et al.* observed structure 3, in two forms, in their simulation of Aβ (16-22) using a simplified model (112). Röhrig *et al.* also find, in simulations of Aβ (16-22) ranging from a dimer to 32-mer, that the tetramer adopts structures similar to structures 1 and 3 identified by ART-OPEP (the single-layer tetramer cannot bend into a barrel, of course) (107).

Comparing various simulations, it is clear that although similar ordered structures are found, the probability of these varies strongly with the sequence, but also the force field and the simulation conditions. For

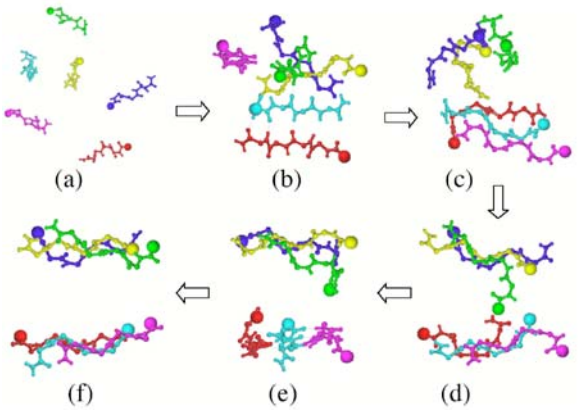

Figure 13. A generic assembly pathway constructed from the aggregation trajectories observed in the ART-OPEP simulation of an hexamer of KFFE. (a) Starting from six monomers with random orientations, (b) a first dimer forms, with the four other chains in a random conformation, (b) a third chain joins the original dimer, (c) while the three remaining random chains also assemble into a trimer, (d) the two trimers rotate and move around each other, adopting a perpendicular orientation, (e) finally, the two trimers organize themselves into a parallel interstrand arrangement with a mixed β-sheets. Dotted lines represents H-bonds between chains.

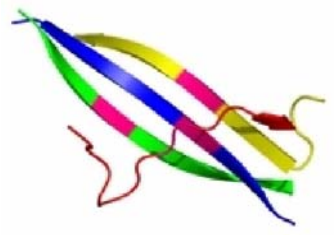

Figure 14. Low-energy structure of tetramer of Aβ (11-25). This structure is obtained through an ART-OPEP simulations starting from a well-formed antiparallel dimer and two monomers placed at 30 Å from the dimer. This structure, a three-stranded β-sheet stabilized by a fourth strand laying across the sheet is typical of what is observed in these simulations even though its lies more than 25 kcal/mol above the native state.

example, while the tetramer of Aβ (16-22) has a non-negligible probability to visit a four stranded β-sheet (47), Wu *et al.* and Colombo *et al.* could not manage to obtain ordered β-sheet of NFGAIL using MD with AMBER (127) or GROMOS force fields (128).

In spite of the variability in the ART-OPEP trajectories (see also Wei *et al.* in this issue) seen between different sequences, we start to have a good understanding of the structure of small oligomers formed by short peptides and the paths leading to ordered β-sheet structures. We find that various ordered topologies with β-sheet content are in equilibrium with amorphous aggregates, the early steps of aggregation are dominated essentially by side-chain - side-chain interactions, while the late steps are dominated essentially by H-bonds, through, notably, reptation moves. These findings are consistent with IR spectroscopy (117) and were confirmed numerically by a recent MD study on the dimer of human transthyretin (105-115) peptide (129).

We must now turn to studying larger oligomers of the same short peptides and understanding the oligomeric structure of longer chains, more relevant from a biological point of view. ART-OPEP trials on dodecamers of NFGAIL have led mostly to amorphous structures, even using fast activated methods (130) and new approaches are clearly needed. As for longer chains, they are the topic of the next section.

7. MONOMERS AND DIMERS OF LONGER PEPTIDES

Longer chains of 10 residues or more show a much richer aggregation behavior than the short peptides discussed in the previous section. They are also much more difficult to simulate as the space of conformation becomes very costly to explore in detail for most methods. Activated techniques, such as ART-OPEP, can therefore play an important role in identifying possible intermediates and aggregation pathways, where sampling with standard methods becomes problematic. However, even with an activated approach, the aggregation of long chains remains challenging.

This is clearly observed in the aggregation simulations of a tetramer of the 15-residue fragment Aβ (11-25) which provides an ideal length midway between the short fragments presented in the previous section and the full-length Aβ peptide. Various experiments indicate that this peptide assembles *in vitro* into amyloid fibril with anti-parallel orientation of the chains (131, 132). ART-OPEP simulations of the monomer (133) suggest, in agreement with experiments on closely related sequence (134), that random coil, α-helices and β-sheets are present in similar proportion excepts in very polar solvent. Stability ART-OPEP study of the four-stranded β-sheet (131, 132) also fits experimental results. However, ART-OPEP simulations are unable to lead to ordered structures when starting from initial random structures. The monomers come to together and form amorphous aggregates with 30% β-sheet content, at best. Even simulations starting from a preformed dimer fail to reach the ordered anti-parallel tetramer, converging instead onto a well-formed anti-parallel three-stranded beta-sheet stabilized by a fourth chain laying across the β-sheet (see Figure 14). These results indicate that the enhanced flexibility of monomers of about 10 residues or more, compared to that of shorter chains, changes qualitatively the aggregation process. If a tetramer of Aβ (16-22) explores well-formed β-sheets, a longer peptide will favor less-ordered structures, mostly for entropic reasons, as it can adopt many more disordered conformations while protecting its hydrophobic core.

It appears therefore that the current numerical tools are not powerful and rapid enough to simulate, from a random solution, the aggregation process of peptides of 10

Table 1. Interproton distances for Aβ (21-30).

ROEs	21α-23N	22α-24N	24α-26N	28α-30N	22α-30N
SC1-ART	3.12	4.04	4.46	4.42	4.77
SC2-ART	3.31	3.59	4.64	4.94	**7.42**
SC3-ART	4.12	**5.03**	4.69	4.69	4.22
NMR-1	4.79	3.64	4.63	4.55	**5.38**
NMR-2	4.92	3.54	4.33	**5.07**	4.84
C1-REMD	4.97	**6.36**	**5.38**	4.36	**6.53**

Weighted-averaged interproton distances for the three families of structures of the fragment Aβ (21-30) identified in ART-OPEP simulations (140) compared to NMR-derived distances (136) and the best structure of all-atom REMD simulations with implicit solvent (138). Distances larger than 5Å, which are considered to violate the constraints, are indicated in bold.

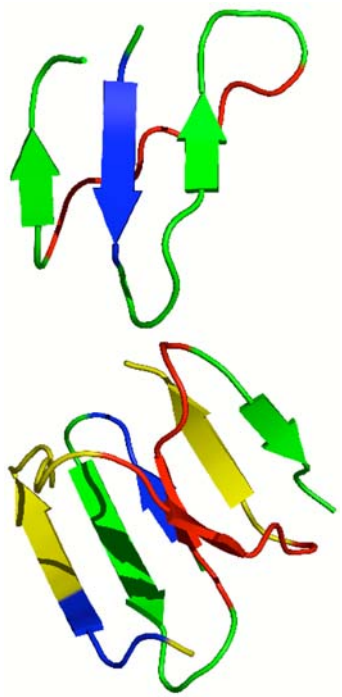

Figure 15. Top: The lowest-energy structure identified in ART-OPEP simulations of the monomer of Aβ (1-28). Bottom: Lowest-energy structure for the dimer of Aβ (1-28). Residues A2-R5 are colored in blue and Q16-A22, in read. Chain 1 is colored in green and chain 2 in yellow.

residues or more with reasonable atomic details. Before facing this question, we must therefore characterize better the configurational space of a monomer and dimer of long peptides and see how it is possible to restrict the space of possible structures as chains are added.

A first test of the approach was performed on the monomer of Aβ (21-30) (AEDVGNKGA), a segment of the amyloid-β peptide which is known to be part of a loop in the fibril and was thought to play a nucleation role in the Aβ aggregation process (135). The interest for this fragment came from a partial proteolysis experiment by Lazo *et al.*, who identified he 21-30 as protease resistant for both Aβ (1-40) and Aβ (1-42) sequences (136). The high-resolution NMR analysis of the peptide Aβ (21-30) that followed identified two dominant structural families for the 10-residue fragment both characterized by a turn at residues Val24- Lys28 (136). Such a structuration is very rare for 10-amino acid sequences and is a useful test for various simulations methods and potentials. These included standard all-atom explicit solvent MD (137), all-atom REMD simulations with an implicit solvent (138), coarse-grained implicit-solvent discontinuous MD (139) and ART-OPEP simulations (140).

As usual, the ART-OPEP simulations (using the OPEP version 3.0 parameters (43)) where started from a fully stretched conformation and run with a Metropolis accept/reject criterion based on the energy difference between the final and initial minimum for each event. In agreement with the other simulations, clustering analysis revealed that the peptide displays a strong bias towards forming a loop at Val24-Lys28. There are many ways to form such a loop and ART-OPEP identified three dominant clusters, but only one overlapped closely with a structure found in the REMD simulations (138). While ART-OPEP results differ in the details from other simulations, they provide the best agreement of all simulations with NMR distances (140): two of the three structural families identified in the simulation respect the 5 NMR-derived constraints while the third family violates only one (See Table 1). In addition to providing new atomistic models respecting the NMR data, this study also demonstrates that ART-OPEP can also provide quantitative and not only qualitative information on the structural properties of monomers in solution.

The theoretical work on longer peptides is much more diverse and so it is not easy to compare the various methods. There has been a number of MD and REMD studies on the 11-residue peptide Aβ (25-35) (141), and the longer peptides Aβ (10-35) (142-144) and Aβ (12-36) (145, 146). Due to sampling limitations, these works tend to focus on the monomer, providing mostly hints as to the dominant contacts observed on these essentially disordered, but not fully random structures.

In this context, we have launched ART-OPEP simulations on both the monomer and dimer of Aβ (1-28), a peptide that has been extensively studied experimentally (147-152). This peptide has received very little attention numerically (153-155). ART-OPEP simulations starting from a fully stretched state show, in agreement with experiments, that the monomer is more than 50 % in

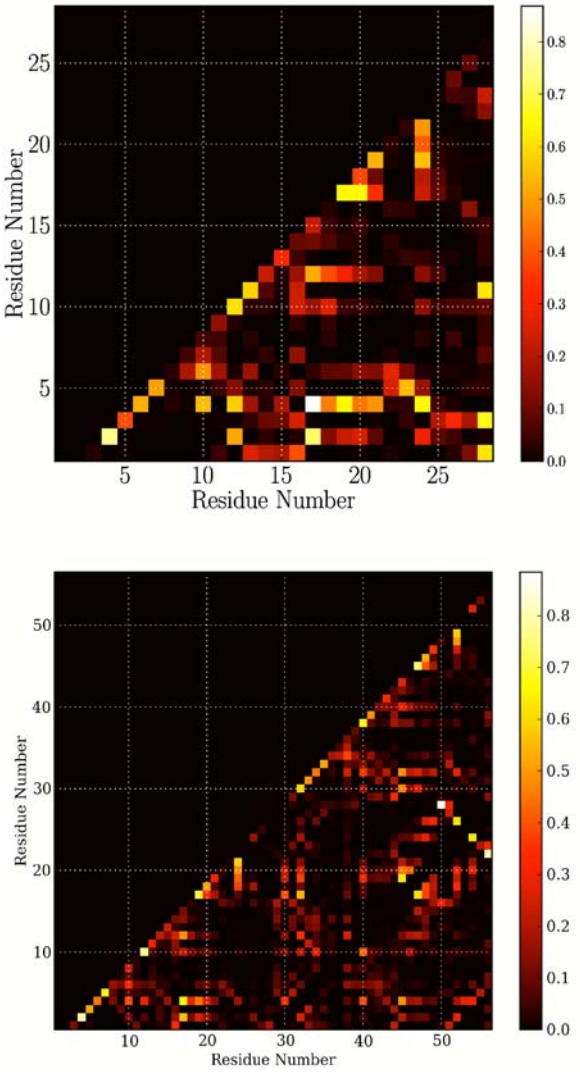

Figure 16. Contact maps for the monomer (top) and dimer (bottom) of Aβ (1-28). In both cases, only the lower-energy structures are considered, i.e. below a standard deviation from the averaged energy sampled.

random coil structure (156, 157). However, its overall shape is relatively well-defined, very often forming a double-hairpin, stabilized by an anti-parallel β-sheet between Ala2-Phe4 and Tyr10-Val12 (Figure 15a). As is generally the case with ART which provides structures at their minimum, the structures tend to be more ordered than those explored experimentally, or by MD. However, ART simulations have the advantage of providing a much clearer picture of the dominant structural features.

While results on the monomer are in qualitative agreement with experiment, it is the conformational change from the monomer to the dimer that is most interesting. The bottom panel of Figure 15 presents one of the lowest-energy structures obtained for the dimer. As we can see, dimerization, that leads to 50 kcal/mol gain in energy over two well-relaxed isolated monomers, produces a much more ordered structure. While the dominant contacts remain in the same regions as for the isolated monomer, the two chains are now fully intertwined with a locally *anti-parallel* organization but a globally *parallel* alignment. This is more evident by looking at the contact map (Figure 16) calculated only on the lower-energy configurations. Since ART does not sample within any well-specified thermodynamical ensemble, an exact weighing of the various conformations is not possible. However, we can focus on the lower-energy structures, defined here as those with an energy below one standard deviation under the average energy of the whole set of visited structure (<E> – σ). This low-energy subset provides considerable information on the dominant structures that should be visited by the peptides. In the top panel, for example, we see that residue Phe4 has a strong probability to form a contact with residue Leu17 in the monomer. The same is true for the dimer (bottom panel), where residue Phe4 of the first chain shows a large propensity to form both an intramolecular contact with residue Leu17 and an intermolecular contact with residue Leu17 of the second chain (residue 45). A careful study of the various contacts reveals a similar pattern of contacts.

The ART-OPEP study of the monomer and dimer of Aβ (1-28) indicates a strong propensity of the N-terminal, in particular the amino acids E3-F4-R5, to form β-sheets, stabilizing the locally anti-parallel structure. Clearly, the structural properties of the dimer are not consistent with that of the fully formed fibril, indicating that a structural change must take place as the fibril grows, a result that has not yet been observed by other simulation techniques at the moment.

The real goal of simulations is, of course, to reach biologically-relevant systems, such as full-length Aβ. Numerical work on the full sequence has proceeded on three fronts. First, it focused on the stability, at least within the MD-reachable time scale of the various proposed fibril models (158, 159). While providing a useful verification of the NMR-derived structure, they cannot really offer any indication of the aggregation pathway. Second, all-atom MD and REMD simulations have focused on the monomer of Aβ40 or Aβ42 (2, 3, 160, 161) since these two sequences, that differ by only two amino acids, show qualitatively different aggregation pathways. Results from Baumketner *et al.* show a Aβ42 monomer that is mostly random coil (160), while Sgourakis *et al.* find much more structuring for both Aβ40 and Aβ42 (161). The third approach, finally, is focused on the description of the aggregation process using simplified potentials and discrete MD (99, 162-164). These simulations can certainly indicate possible intermediates but they tend to have strong bias towards aggregation and lack the sequence specificity. They represent currently, nevertheless, the only possible approach for the study of large-scale aggregation.

ART-OPEP has also been applied to the full Aβ sequence, generating 10 runs of 30 000 events for both Aβ40 and Aβ42. Preliminary results are available at the moment. As with all-atom MD simulations, we find that both monomers are mostly random coil and sample a wide

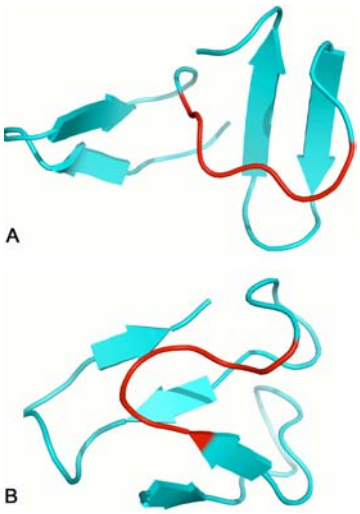

Figure 17. Low energy structures for the monomer of (a) Aβ40 and (b) Aβ42 from runs performed with ART-OPEP starting from fully extended state. The residues L17- E22, corresponding to the hydrophobic central core, are marked in red.

range of structures. In some cases, we recover the C-terminal hairpin observed by Skourakis *et al.* (161) for Aβ42 (see Figure 17), but we also observe a tendency for the N-terminal to assemble into a hairpin, a behavior similar to that observed in the monomer of Aβ (1-28), albeit with a lower probability. If we only consider the lowest-energy state for each of the runs, we see that the monomer of Aβ40 is almost as likely to form β-strands as Aβ42, with a percentage of β-strand, calculated over the lowest-energy state of each of 10 runs, of about 25 % for the former, compared to about 28 % for the latter. Representative structures for both Aβ40 and Aβ42 are shown in Figure 17. We see that both sequences visit some very similar structures. We also note also that the central hydrophobic core, residues 17-21, marked in red, is very disordered in agreement with NMR study (135). Clearly, the conformations sampled by the monomer do not seem sufficient to explain the kinetic differences observed in the aggregation of Aβ40 and Aβ42 (165).

8. SUMMARY AND PERSPECTIVES

The simulation of protein aggregation is a major challenge that cannot be tackled heads on with current computing facilities. It is necessary to approach this problem from multiple angles making various approximations, in order to develop a complete and coherent picture of the aggregation process.

In this review, we have focused on an activated approach, ART nouveau, coupled with the generic OPEP coarse-grained force field. This approach can go beyond qualitative prediction for the monomer of Aβ (21-30) (141). This demonstrates that in spite of its limitations, ART-OPEP can also provide quantitative match to experiments, with structures in full agreement with NMR measurements. Both components were developed for protein folding of non-amyloidogenic sequences and are therefore not biased for following the aggregation of amyloid-forming proteins. ART-OPEP has proven especially useful in identifying the richness in the structure of the small aggregates of short chains, a richness that was confirmed by many other works. Of course, any simulation method is only as good as its two components: the model — determined by the force field — and the sampling method — MD (47), REMD, MC (38) or ART. Since ART and OPEP are independent, both can be used with other methods. OPEP, for example, was also used with MC, MD and REMD. On proteins, ART has also been used with internal coordinates using FLEX and AMBER (166). While ART can work with any derivable potential, it must be used with implicit solvents since it only samples local-energy minima.

Activated simulations on amyloid-forming peptides of various length indicate that there is an important difference in the early steps of aggregation between short (less than 10 amino acids) and long peptides. Results on the monomer and dimer of Aβ (1-28), for example, indicate that their populated structures are very far from the amyloid fibril structure, while the dimer of KFFE, Aβ (16-22) and other related sequences already visit fibril-competent structures, albeit with a probability of 10-20%. Similar conclusions can be drawn for the monomer (this work) and the dimer of Aβ40 and Aβ42 (in preparation). At this point, as was shown in this review, activated methods with unbiased coarse-grained potentials, i.e., not biased towards the formation of amyloid fibrils, are certainly a very promising approach to explore these questions.

9. ACKOWLEDGEMENTS

The results discussed in this review were obtained in collaboration with a number of people, including G. Boucher, W. Chen, X. Dong, A. Melquiond, S. Santini, and G.-H. Wei. We thank them for their essential contributions to this project over the years. We would also like to acknowledge support from the Alzheimer Society of Canada (2005-2007). N.M. is grateful to NSERC, the Canada Research Chair Foundation and the Fonds québécois de recherche sur la nature et les technologies. P.D. acknowledges financial support from the Centre National de la Recherche Scientifique and the University of Paris 7 Denis Diderot.

10. REFERENCES

1. Lashuel, H. A. & P. T. Lansbury, Jr.: Are amyloid diseases caused by protein aggregates that mimic bacterial pore-forming toxins? *Quarterly reviews of biophysics*, 39, 167-201 (2006)
2. Flock, D., S. Colacino, G. Colombo & A. Di Nola: Misfolding of the amyloid beta-protein: a molecular dynamics study. *Proteins*, 62, 183-92 (2006)
3. Xu, Y., J. Shen, X. Luo, W. Zhu, K. Chen, J. Ma & H. Jiang: Conformational transition of amyloid beta-peptide. *Proc Natl Acad Sci U S A*, 102, 5403-7 (2005)
4. Ma, B. & R. Nussinov: Simulations as analytical tools to understand protein aggregation and predict amyloid

conformation. *Current opinion in chemical biology*, 10, 445-52 (2006)
5. Wei, G., W. Song, P. Derreumax & N. Mousseau: Self-assembly of amyloid-forming peptides using molecular dynamics simulations and the OPEP coarse-grained force field *Frontiers in biosciences*, this issue,
6. Eyring, H.: The Activated Complex and the Absolute Rate of Chemical Reactions. *Chem. Rev.*, 17, 65-77 (1935)
7. Wigner, E.: The transition state method. *Trans. Faraday Soc.*, 34, 29-41 (1938)
8. Goldstein, M.: Viscous Liquids and the Glass Transition: A Potential Energy Barrier Picture. *J. Chem. Phys.*, 51, 3728 (1969)
9. Stillinger, F. H. & T. A. Weber: Dynamics of structural transitions in liquids. *Phys. Rev. A*, 28, 2408 (1983)
10. Wales, D. J.: Energy landscapes. Cambridge University Press, Cambridge (2004)
11. Levinthal, C.: *Mössbauer Spectroscopy in Biological Systems, Proceedings of a Meeting Held at Allerton House, Monticello*, J. T. P. DeBrunner, E. Munck, Ed., University of Illinois Press, 22 (1969).
12. Onuchic, J. N., P. G. Wolynes, Z. Lutheyschulten & N. D. Socci: Toward an Outline of the Topography of a Realistic Protein-Folding Funnel. *Proc. Natl. Acad. Sci. U. S. A.*, 92, 3626-3630 (1995)
13. St-Pierre, J.-F., N. Mousseau & P. Derreumaux: The complex folding pathways of protein A suggest a multiple-funnelled energy landscape *J. Chem. Phys.*, (in press),
14. Becker, O. M. & M. Karplus: The topology of multidimensional potential energy surfaces: Theory and application to peptide structure and kinetics. *J. Chem. Phys.*, 106, 1495-1517 (1997)
15. Rao, F. & A. Caflisch: The protein folding network. *J Mol Biol*, 342, 299-306 (2004)
16. Mousseau, N., P. Derreumaux & G. Gilbert: Navigation and analysis of the energy landscape of small proteins using the activation-relaxation technique. *Physical biology*, 2, S101-7 (2005)
17. Wales, D. J. & T. V. Bogdan: Potential energy and free energy landscapes. *J. Phys. Chem. B*, 110, 20765-76 (2006)
18. Krivov, S. & M. Karplus: Hidden complexity of free energy surfaces for peptide (protein) folding. *Proc Natl Acad Sci USA*, 101, 14766-70 (2004)
19. Dellago, C., P. G. Bolhuis, F. S. Csajka & D. Chandler: Transition path sampling and the calculation of rate constants. *J. Chem. Phys.*, 108, 1964-1977 (1998)
20. Laio, A. & M. Parrinello: Escaping free-energy minima. *Proc Natl Acad Sci USA*, 99, 12562-6 (2002)
21. Faradjian, A. & R. Elber: Computing time scales from reaction coordinates by milestoning. *J. Chem. Phys.*, 120, 10880-9 (2004)
22. Wang, F. & D. P. Landau: Determining the density of states for classical statistical models: A random walk algorithm to produce a flat histogram. *Phys. Rev. E*, 64, 056101 (2001)
23. Sugita, Y. & Y. Okamoto: Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett*, 314, 141-151 (1999)
24. Kumar, S., J. M. Rosenberg, D. Bouzida, R. H. Swendsen & P. A. Kollman: The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.*, 13, 1011-1021 (1992)
25. Barkema, G. T. & N. Mousseau: Event-Based Relaxation of Continuous Disordered Systems. *Phys. Rev. Lett.*, 77, 4358-4361 (1996)
26. Doye, J. P. K. & D. J. Wales: Surveying a potential energy surface by eigenvector-following - Applications to global optimisation and the structural transformations of clusters. *Z. Phys. D-Atoms Mol. Clusters*, 40, 194-197 (1997)
27. Malek, R. & N. Mousseau: Dynamics of Lennard-Jones clusters: A characterization of the activation-relaxation technique. *Phys. Rev. E*, 62, 7723-8 (2000)
28. Elber, R. & M. Karplus: A method for determining reaction paths in large molecules: Application to myoglobin. *Chem Phys Lett*, 139, 375-380 (1987)
29. Gillilan, R. E. & K. R. Wilson: Shadowing, rare events, and rubber bands. A variational Verl*et al*gorithm for molecular dynamics. *J. Chem. Phys.*, 97, 1757-1772 (1992)
30. Jónsson, H., G. Mills & K. W. Jacobsen: Nudged elastic band method for finding minimum energy paths of transitions. In: Classical and Quantum Dynamics in Condensed Phase Simulations. Eds: B. J. Berne, G. Ciccotti&D. F. Coker. World Scientific, Singapore (1998)
31. Henkelman, G. & H. Jonsson: A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives. *J. Chem. Phys.*, 111, 7010-7022 (1999)
32. Doye, J. P. K., D. J. Wales & M. A. Miller: Thermodynamics and the global optimization of Lennard-Jones clusters. *J. Chem. Phys.*, 109, 8143-8153 (1998)
33. Munro, L. J. & D. J. Wales: Defect migration in crystalline silicon. *Phys. Rev. B*, 59, 3969 (1999)
34. Olsen, R., G. Kroes, G. Henkelman, A. Arnaldsson & H. Jónsson: Comparison of methods for finding saddle points without knowledge of the final states. *J. Chem. Phys.*, 121, 9776-92 (2004)
35. El-Mellouhi, F., N. Mousseau & P. Ordejon: Sampling the diffusion paths of a neutral vacancy in silicon with quantum mechanical calculations. *Phys. Rev. B*, 70, 9 (2004)
36. Miller, M. A., J. P. K. Doye & D. J. Wales: Structural relaxation in atomic clusters: Master equation dynamics. *Phys. Rev. E*, 60, 3701-3718 (1999)
37. Evans, D. A. & D. J. Wales: Folding of the GB1 hairpin peptide from discrete path sampling. *J Chem Phys*, 121, 1080-90 (2004)
38. Derreumaux, P.: Folding a 20 amino acid alpha beta peptide with the diffusion process-controlled Monte Carlo method. *J. Chem. Phys.*, 107, 1941-1947 (1997)
39. Levitt, M. & A. Warshel: Computer simulation of protein folding. *Nature*, 253, 694-8 (1975)
40. Derreumaux, P.: From polypeptide sequences to structures using Monte Carlo simulations and an optimized potential. *J. Chem. Phys.*, 111, 2301-2310 (1999)
41. Case, D., T. Cheatham, T. Darden, H. Gohlke, R. Luo, K. Merz, A. Onufriev, C. Simmerling, B. Wang & R. Woods: The Amber biomolecular simulation programs. *J Comp Chem*, 26, 1668-88 (2005)
42. Derreumaux, P., K. Wilson, G. Vergoten & W. Peticolas: Conformational studies of neuroactive ligands.

1. Force field and vibrational spectra of crystalline acetylcholine. *J. Phys. Chem.*, 93, 1338-1350 (1989)
43. Maupetit, J., P. Tuffery & P. Derreumaux: A coarse-grained protein force field for folding and structure prediction. *Proteins*, 69, 394-408 (2007)
44. Derreumaux: Generating ensemble averages for small proteins from extended conformations by Monte Carlo simulations. *Phys. Rev. Lett.*, 85, 206-209 (2000)
45. Floquet, N., S. Pasco, L. Ramont, P. Derreumaux, J. Laronze, J. Nuzillard, F. Maquart, A. Alix & J. Monboisse: The antitumor properties of the alpha3 (IV)- (185-203) peptide from the NC1 domain of type IV collagen (tumstatin) are conformation-dependent. *J. Biol. Chem.*, 279, 2091-100 (2004)
46. Mousseau, N., P. Derreumaux, G. T. Barkema & R. Malek: Sampling activated mechanisms in proteins with the activation-relaxation technique. *J. Mol. Graph. Model.*, 19, 78-86 (2001)
47. Derreumaux, P. & N. Mousseau: Coarse-grained protein molecular dynamics simulations. *J. Chem. Phys.*, 126, 025101 (2007)
48. Dong, X., R. Laghaei, Y. Chebaro, P. Derreumaux & N. Mousseau: Replica exchange molecular dynamics simulations using coarse-grained OPEP force field: application to peptides. *in preparation*
49. Munoz, V. & L. Serrano: Helix design, prediction and stability. *Current Opinion in Biotechnology*, 6, 382-386 (1995)
50. Williams, S., T. P. Causgrove, R. Gilmanshin, K. S. Fang, R. H. Callender, W. H. Woodruff & R. B. Dyer: Fast Events in Protein Folding: Helix Melting and Formation in a Small Peptide. *Biochemistry*, 35, 691-697 (1996)
51. Doig, A. J.: Recent advances in helix-coil theory. *Biophysical Chemistry*, 101-102, 281-293 (2002)
52. Wei, G., P. Derreumaux & N. Mousseau: Sampling the complex energy landscape of a simple beta-hairpin. *J. Chem. Phys.*, 119, 6403-6406 (2003)
53. Wei, G., N. Mousseau & P. Derreumaux: Complex folding pathways in a simple beta-hairpin. *Proteins*, 56, 464-74 (2004)
54. Blanco, F. J. & L. Serrano: Folding of protein-g b1 domain studied by the conformational characterization of fragments comprising its secondary structure elements. *Eur. J. Biochem.*, 230, 634-649 (1995)
55. Muñoz, V., P. Thompson, J. Hofrichter & W. Eaton: Folding dynamics and mechanism of beta-hairpin formation. *Nature*, 390, 196-9 (1997)
56. Muñoz, V., R. Ghirlando, F. Blanco, G. Jas, J. Hofrichter & W. Eaton: Folding and aggregation kinetics of a beta-hairpin. *Biochemistry*, 45, 7023-35 (2006)
57. Blanco, F., G. Rivas & L. Serrano: A short linear peptide that folds into a native stable beta-hairpin in aqueous solution. *Nat Struct Biol*, 1, 584-90 (1994)
58. Du, D., M. Tucker & F. Gai: Understanding the mechanism of beta-hairpin folding via phi-value analysis. *Biochemistry*, 45, 2668-78 (2006)
59. Wei, Y., B. Huyghues-Despointes, J. Tsai & J. Scholtz: NMR study and molecular dynamics simulations of optimized beta-hairpin fragments of protein G. *Proteins*, 69, 285-96 (2007)
60. Dyer, R., S. Maness, E. Peterson, S. Franzen, R. Fesinmeyer & N. Andersen: The mechanism of beta-hairpin formation. *Biochemistry*, 43, 11560-6 (2004)
61. Ma, B. & R. Nussinov: Explicit and implicit water simulations of a beta-hairpin peptide. *Proteins*, 37, 73-87 (1999)
62. Higo, J., N. Ito, M. Kuroda, S. Ono, N. Nakajima & H. Nakamura: Energy landscape of a peptide consisting of alpha-helix, 3 (10)-helix, beta-turn, beta-hairpin, and other disordered conformations. *Protein Sci*, 10, 1160-71 (2001)
63. Ma, B. & R. Nussinov: Energy landscape and dynamics of the beta-hairpin G peptide and its isomers: Topology and sequences. *Protein Sci*, 12, 1882-93 (2003)
64. Zhou, R., B. Berne & R. Germain: The free energy landscape for beta hairpin folding in explicit water. *Proc Natl Acad Sci USA*, 98, 14931-6 (2001)
65. Kamiya, N., J. Higo & H. Nakamura: Conformational transition states of a beta-hairpin peptide between the ordered and disordered conformations in explicit water. *Protein Sci*, 11, 2297-307 (2002)
66. Daidone, I., A. Amadei & A. Di Nola: Thermodynamic and kinetic characterization of a beta-hairpin peptide in solution: an extended phase space sampling by molecular dynamics simulations in explicit water. *Proteins*, 59, 510-8 (2005)
67. Bolhuis, P.: Kinetic pathways of beta-hairpin (un)folding in explicit solvent. *Biophys. J.*, 88, 50-61 (2005)
68. Zhou, R. & B. Berne: Can a continuum solvent model reproduce the free energy landscape of a beta -hairpin folding in water? *Proc Natl Acad Sci USA*, 99, 12777-82 (2002)
69. García, A. & K. Sanbonmatsu: Exploring the energy landscape of a beta hairpin in explicit solvent. *Proteins*, 42, 345-54 (2001)
70. Pitera, J., I. Haque & W. Swope: Absence of reptation in the high-temperature folding of the trpzip2 beta-hairpin peptide. *J. Chem. Phys.*, 124, 141102 (2006)
71. Zagrovic, B., E. Sorin & V. Pande: Beta-hairpin folding simulations in atomistic detail using an implicit solvent model. *J Mol Biol*, 313, 151-69 (2001)
72. Yoda, T., Y. Sugita & Y. Okamoto: Cooperative folding mechanism of a beta-hairpin peptide studied by a multicanonical replica-exchange molecular dynamics simulation. *Proteins*, 66, 846-59 (2007)
73. Santiveri, C., M. Jiménez, M. Rico, W. Van Gunsteren & X. Daura: beta-hairpin folding and stability: molecular dynamics simulations of designed peptides in aqueous solution. *J Pept Sci*, 10, 546-65 (2004)
74. Chen, C. & Y. Xiao: Molecular dynamics simulations of folding processes of a beta-hairpin in an implicit solvent. *Phys. Biol.*, 3, 161-71 (2006)
75. Evans, D. & D. Wales: Folding of the GB1 hairpin peptide from discrete path sampling. *J. Chem. Phys.*, 121, 1080-90 (2004)
76. Lwin, T. & R. Luo: Force field influences in beta-hairpin folding simulations. *Protein Sci*, 15, 2642-55 (2006)
77. Irbäck, A., B. Samuelsson, F. Sjunnesson & S. Wallin: Thermodynamics of alpha- and beta-structure formation in proteins. *Biophys. J.*, 85, 1466-73 (2003)

78. Ding, F., J. Borreguero, S. Buldyrey, H. Stanley & N. Dokholyan: Mechanism for the alpha-helix to beta-hairpin transition. *Proteins*, 53, 220-8 (2003)
79. Dinner, A., T. Lazaridis & M. Karplus: Understanding beta-hairpin formation. *Proc Natl Acad Sci USA*, 96, 9068-73 (1999)
80. Kolinski, A., B. Ilkowski & J. Skolnick: Dynamics and thermodynamics of beta-hairpin assembly: insights from various simulation techniques. *Biophys J*, 77, 2942-52 (1999)
81. Klimov, D. & D. Thirumalai: Mechanisms and kinetics of beta-hairpin formation. *Proc Natl Acad Sci USA*, 97, 2544-9 (2000)
82. Munoz, V., E. R. Henry, J. Hofrichter & W. A. Eaton: A statistical mechanical model for beta-hairpin kinetics. *Proc Natl Acad Sci U S A*, 95, 5872-9 (1998)
83. Pande, V. & D. Rokhsar: Molecular dynamics simulations of unfolding and refolding of a beta-hairpin fragment of protein G. *Proc Natl Acad Sci USA*, 96, 9062-7 (1999)
84. Ma, B. & R. Nussinov: Molecular dynamics simulations of a beta-hairpin fragment of protein G: balance between side-chain and backbone forces. *J Mol Biol*, 296, 1091-104 (2000)
85. Lee, J. & S. Shin: Understanding beta-hairpin formation by molecular dynamics simulations of unfolding. *Biophys. J.*, 81, 2507-16 (2001)
86. Zhou, Y. & A. Linhananta: Role of hydrophilic and hydrophobic contacts in folding of the second beta-hairpin fragment of protein G: molecular dynamics simulation studies of an all-atom model. *Proteins*, 47, 154-62 (2002)
87. Irback, A. & S. Mohanty: Folding thermodynamics of peptides. *Biophys. J.*, 88, 1560-9 (2005)
88. Nguyen, H. D. & C. K. Hall: Spontaneous fibril formation by polyalanines; discontinuous molecular dynamics simulations. *J. Am. Chem. Soc.*, 128, 1890-901 (2006)
89. Zhou, R.: Free energy landscape of protein folding in water: explicit vs. implicit solvent. *Proteins*, 53, 148-61 (2003)
90. Yang, S., J. Onuchic, A. García & H. Levine: Folding Time Predictions from All-atom Replica Exchange Simulations. *J. Mol. Biol,* 372, 756-63 (2007)
91. Baumketner, A. & J. Shea: The influence of different treatments of electrostatic interactions on the thermodynamics of folding of peptides. *J Phys Chem B*, 109, 21322-8 (2005)
92. Bolhuis, P.: Transition-path sampling of beta-hairpin folding. *Proc Natl Acad Sci USA*, 100, 12129-34 (2003)
93. Wu, X. & B. Brooks: Beta-hairpin folding mechanism of a nine-residue peptide revealed from molecular dynamics simulations in explicit water. *Biophys J*, 86, 1946-58 (2004)
94. Tyka, M., A. Clarke & R. Sessions: An efficient, path-independent method for free-energy calculations. *J Phys Chem B*, 110, 17212-20 (2006)
95. Gallicchio, E., M. Andrec, A. Felts & R. Levy: Temperature weighted histogram analysis method, replica exchange, and transition paths. *J Phys Chem B*, 109, 6722-31 (2005)
96. Tsai, J. & M. Levitt: Evidence of turn and salt bridge contributions to beta-hairpin stability: MD simulations of C-terminal fragment from the B1 domain of protein G. *Biophysical chemistry*, 101-102, 187-201 (2002)
97. Wei, G., P. Derreumaux & N. Mousseau: Sampling the complex energy landscape of a simple beta-hairpin. *J Chem Phys*, 119, 6403-6406 (2003)
98. Swope, W., J. Pitera, F. Suits, M. Pitman, M. Eleftheriou, B. Fitch, R. Germain, A. Rayshubski, T. Ward, Y. Zhestkov & R. Zhou: Describing Protein Folding Kinetics by Molecular Dynamics Simulations. 2. Example Applications to Alanine Dipeptide and a beta-Hairpin Peptide. *J Phys Chem B*, 108, 6582-6594 (2004)
99. Nguyen, P., G. Stock, E. Mittag, C. Hu & M. Li: Free energy landscape and folding mechanism of a beta-hairpin in explicit water: a replica exchange molecular dynamics study. *Proteins*, 61, 795-808 (2005)
100. Blanco, F. J., M. A. Jimenez, J. Herranz, M. Rico, J. Santoro & J. L. Nieto: NMR evidence of a short linear peptide that folds into a beta-hairpin in aqueous-solution. *Journal of the American Chemical Society*, 115, 5887-5888 (1993)
101. Imamura, H. & J. Chen: Dependence of folding dynamics and structural stability on the location of a hydrophobic pair in beta-hairpins. *Proteins*, 63, 555-70 (2006)
102. Tjernberg, L., W. Hosia, N. Bark, J. Thyberg & J. Johansson: Charge attraction and beta propensity are necessary for amyloid fibril formation from tetrapeptides. *J. Biol. Chem.*, 277, 43243-6 (2002)
103. Lopez De La Paz, M., K. Goldie, J. Zurdo, E. Lacroix, C. M. Dobson, A. Hoenger & L. Serrano: De novo designed peptide-based amyloid fibrils. *Proc Natl Acad Sci U S A*, 99, 16052-7 (2002)
104. Reches, M., Y. Porat & E. Gazit: Amyloid fibril formation by pentapeptide and tetrapeptide fragments of human calcitonin. *J. Biol. Chem.*, 277, 35475-80 (2002)
105. Balbach, J. J., Y. Ishii, O. N. Antzutkin, R. D. Leapman, N. W. Rizzo, F. Dyda, J. Reed & R. Tycko: Amyloid fibril formation by A beta 16-22, a seven-residue fragment of the Alzheimer's beta-amyloid peptide, and structural characterization by solid state NMR. *Biochemistry*, 39, 13748-59 (2000)
106. Takeda, T. & D. K. Klimov: Dissociation of A (beta)16-22 Amyloid Fibrils Probed by Molecular Dynamics. *J. Mol. Biol.*, 368, 1202-1213 (2007)
107. Rohrig, U. F., A. Laio, N. Tantalo, M. Parrinello & R. Petronzio: Stability and structure of oligomers of the Alzheimer peptide Abeta16-22: from the dimer to the 32-mer. *Biophys. J.*, 91, 3217-29 (2006)
108. Ma, B. Y. & R. Nussinov: Stabilities and conformations of Alzheimer's beta-amyloid peptide oligomers (A beta (16-22 ') A beta (16-35 ') and A beta (10-35)): Sequence effects. *Proc. Natl. Acad. Sci. U. S. A.*, 99, 14126-14131 (2002)
109. Klimov, D. & Thirumalai: Dissecting the assembly of Abeta16-22 amyloid peptides into antiparallel beta sheets. *Structure*, 11, 295-307 (2003)
110. Hwang, W., S. Zhang, R. D. Kamm & M. Karplus: Kinetic control of dimer structure formation in amyloid fibrillogenesis. *Proceedings of the National Academy of Sciences*, 101, 12916-12921 (2004)
111. Gnanakaran, S., R. Nussinov & A. E. Garcia: Atomic-level description of amyloid beta-dimer formation. *J. Am. Chem. Soc.*, 128, 2158-9 (2006)

112. Favrin, G., A. Irback & S. Mohanty: Oligomerization of amyloid Abeta16-22 peptides using hydrogen bonds and hydrophobicity forces. *Biophys. J.*, 87, 3657-64 (2004)
113. Meinke, J. H. & U. H. Hansmann: Aggregation of beta-amyloid fragments. *J. Chem. Phys.*, 126, 014706 (2007)
114. Santini, S., N. Mousseau & P. Derreumaux: In silico assembly of Alzheimer's A beta (16-22) peptide into beta-sheets. *J. Am. Chem. Soc.*, 126, 11509-11516 (2004)
115. Santini, S., G. H. Wei, N. Mousseau & P. Derreumaux: Pathway complexity of Alzheimer's beta-amyloid A beta (16-22) peptide assembly. *Structure*, 12, 1245-1255 (2004)
116. Gsponer, J., U. Haberthur & A. Caflisch: The role of side-chain interactions in the early steps of aggregation: Molecular dynamics simulations of an amyloid-forming peptide from the yeast prion Sup35. *Proc Natl Acad Sci U S A*, 100, 5154-9 (2003)
117. Petty, S. A. & S. M. Decatur: Experimental evidence for the reorganization of beta-strands within aggregates of the Abeta (16-22) peptide. *J. Am. Chem. Soc.*, 127, 13488-9 (2005)
118. Lei, H., C. Wu, Z. Wang & Y. Duan: Molecular dynamics simulations and free energy analyses on the dimer formation of an amyloidogenic heptapeptide from human beta2-microglobulin: implication for the protofibril structure. *J. Mol. Biol.*, 356, 1049-63 (2006)
119. Melquiond, A., G. Boucher, N. Mousseau & P. Derreumaux: Following the aggregation of amyloid-forming peptides by computer simulations. *J. Chem. Phys.*, 122, 174904 (2005)
120. Wei, G. H., N. Mousseau & P. Derreumaux: Sampling the self-assembly pathways of KFFE hexamers. *Biophys. J.*, 87, 3648-3656 (2004)
121. Melquiond, A., N. Mousseau & P. Derreumaux: Structures of soluble amyloid oligomers from computer simulations. *Proteins*, 65, 180-191 (2006)
122. Wei, G. H., N. Mousseau & P. Derreumaux: Exploring the early steps of aggregation of amyloid-forming peptide KFFE. *J Phys -Cond Mat*, 16, S5047-S5054 (2004)
123. Ma, B. & R. Nussinov: Molecular dynamics simulations of alanine rich beta-sheet oligomers: Insight into amyloid formation. *Protein Sci*, 11, 2335-50 (2002)
124. Zanuy, D. & R. Nussinov: The sequence dependence of fiber organization. A comparative molecular dynamics study of the islet amyloid polypeptide segments 22-27 and 22-29. *J. Mol. Biol.*, 329, 565-84 (2003)
125. Gazit, E., P. della Bruna, S. Pieraccini & G. Colombo: The molecular dynamics of assembly of the ubiquitous aortic medial amyloidal medin fragment. *J. Mol. Graph. Model.*, 25, 903-11 (2007)
126. Zanuy, D., Y. Porat, E. Gazit & R. Nussinov: Peptide sequence and amyloid formation; molecular simulations and experimental study of a human islet amyloid polypeptide fragment and its analogs. *Structure*, 12, 439-55 (2004)
127. Wu, C., H. Lei & Y. Duan: Elongation of ordered peptide aggregate of an amyloidogenic hexapeptide NFGAIL observed in molecular dynamics simulations with explicit solvent. *J. Am. Chem. Soc.*, 127, 13530-7 (2005)
128. Colombo, G., I. Daidone, E. Gazit, A. Amadei & A. Di Nola: Molecular dynamics simulation of the aggregation of the core-recognition motif of the islet amyloid polypeptide in explicit water. *Proteins*, 59, 519-27 (2005)
129. Li, D., L. Han & S. Huo: Structural and pathway complexity of beta-strand reorganization within aggregates of human transthyretin (105-115) peptide. *J Phys Chem B*, 111, 5425-33 (2007)
130. Melquiond, A., J. C. Gelly, N. Mousseau & P. Derreumaux: Probing amyloid fibril formation of the NFGAIL peptide by computer simulations. *J. Chem. Phys.*, 126, 065101 (2007)
131. Sikorski, E. Atkins & L. Serpell: Structure and texture of fibrous crystals formed by Alzheimer's abeta (11-25) peptide fragment. *Structure*, 11, 915-926 (2003)
132. Petkova, A., G. Buntkowsky, F. Dyda, R. Leapman, W. Yau & R. Tycko: Solid state NMR reveals a pH-dependent antiparallel beta-sheet registry in fibrils formed by a beta-amyloid peptide. *J. Mol. Biol.*, 335, 247-60 (2004)
133. Boucher, G., N. Mousseau & P. Derreumaux: Aggregating the amyloid A beta (11-25) peptide into a four-stranded beta-sheet structure. *Proteins*, 65, 877-888 (2006)
134. Juszczyk, P., A. Kołodziejczyk & Z. Grzonka: Circular dichroism and aggregation studies of amyloid beta (11-8) fragment and its variants. *Acta Biochim Pol*, 52, 425-31 (2005)
135. Hou, Shao, Zhang, Li, N. Menon, E. Neuhaus, J. Brewer, I. Byeon, D. Ray, M. Vitek, Iwashita, R. Makula, A. Przybyla & M. Zagorski: Solution NMR studies of the A beta (1-40) and A beta (1-42) peptides establish that the Met35 oxidation state affects the mechanism of amyloid formation. *J. Am. Chem. Soc.*, 126, 1992-2005 (2004)
136. Lazo, N., M. Grant, M. Condron, A. Rigby & D. Teplow: On the nucleation of amyloid beta-protein monomer folding. *Protein Sci*, 14, 1581-1596 (2005)
137. Cruz, L., B. Urbanc, J. Borreguero, N. Lazo, D. Teplow & H. Stanley: Solvent and mutation effects on the nucleation of amyloid beta-protein folding. *Proc Natl Acad Sci USA*, 102, 18258-63 (2005)
138. Baumketner, A., S. Bernstein, T. Wyttenbach, N. Lazo, D. Teplow, M. Bowers & J. Shea: Structure of the 21-30 fragment of amyloid beta-protein. *Protein Sci*, 15, 1239-47 (2006)
139. Borreguero, J., B. Urbanc, N. Lazo, S. Buldyrev, D. Teplow & H. Stanley: Folding events in the 21-30 region of amyloid beta-protein (Abeta) studied in silico. *Proc Natl Acad Sci USA*, 102, 6015-20 (2005)
140. Chen, W., N. Mousseau & P. Derreumaux: The conformations of the amyloid-beta (21-30) fragment can be described by three families in solution. *J. Chem. Phys.*, 125, 084911 (8 pages) (2006)
141. Wei, G. & J. E. Shea: Effects of solvent on the structure of the Alzheimer amyloid-beta (25-35) peptide. *Biophys. J.*, 91, 1638-47 (2006)
142. Tarus, B., J. Straub & D. Thirumalai: Dynamics of Asp23-Lys28 salt-bridge formation in Abeta10-35 monomers. *J. Am. Chem. Soc.*, 128, 16159-68 (2006)
143. Han, W. & Y. D. Wu: A strand-loop-strand structure is a possible intermediate in fibril elongation: long time simulations of amyloid-beta peptide (10-35). *J. Am. Chem. Soc.*, 127, 15408-16 (2005)

144. Baumketner, A. & J. Shea: The structure of the Alzheimer amyloid beta 10-35 peptide probed through replica-exchange molecular dynamics simulations in explicit solvent. *J. Mol. Biol.*, 366, 275-85 (2007)
145. Ikebe, J., N. Kamiya, J. Ito, H. Shindo & J. Higo: Simulation study on the disordered state of an Alzheimer's beta amyloid peptide Abeta (12 36) in water consisting of random-structural, beta-structural, and helical clusters. *Protein Sci*, 16, 1596-608 (2007)
146. Kamiya, N., D. Mitomo, J. Shea & J. Higo: Folding of the 25 Residue Abeta (12-36) Peptide in TFE/Water: Temperature-Dependent Transition from a Funneled Free-Energy Landscape to a Rugged One. *J Phys Chem B,* 19, 5351-6 (2007)
147. Mikros, Benaki, Humpfer, Spraul, Loukas, Stassinopoulou & Pelecanou: High-Resolution NMR Spectroscopy of the beta-Amyloid (1-28) Fibril Typical for Alzheimer's Disease. *Angew Chem Int Ed Engl*, 40, 3603-3605 (2001)
148. Talafous, K. Marcinowski, Klopman & M. Zagorski: Solution structure of residues 1-28 of the amyloid beta-peptide. *Biochemistry*, 33, 7788-7796 (1994)
149. Ma, Clancy, Zhang, Ray, Wollenberg & Zagorski: Residue-Specific pKa Measurements of the beta-Peptide and Mechanism of pH-Induced Amyloid Formation. *J. Am. Chem. Soc.*, 121, 8698-8706 (1999)
150. Eker, Griebenow & Schweitzer-Stenner: Abeta (1-28) fragment of the amyloid peptide predominantly adopts a polyproline II conformation in an acidic solution. *Biochemistry*, 43, 6893-6898 (2004)
151. Shen, C., G. Scott, F. Merchant & R. Murphy: Light scattering analysis of fibril growth from the amino-terminal fragment beta (1-28) of beta-amyloid peptide. *Biophys. J.*, 65, 2383-95 (1993)
152. Yang, D., J. McLaurin, K. Qin, D. Westaway & P. Fraser: Examining the zinc binding site of the amyloid-beta peptide. *Eur J Biochem*, 267, 6692-8 (2000)
153. Tiana, G., F. Simona, R. Broglia & G. Colombo: Thermodynamics of beta-amyloid fibril formation. *J. Chem. Phys.*, 120, 8307-17 (2004)
154. Kirshenbaum & Daggett: pH-dependent conformations of the amyloid beta (1-28) peptide fragment explored using molecular dynamics. *Biochemistry*, 34, 7629-7639 (1995)
155. Valerio, M., A. Colosimo, F. Conti, A. Giuliani, A. Grottesi, C. Manetti & J. Zbilut: Early events in protein aggregation: molecular flexibility and hydrophobicity/charge interaction in amyloid peptides as studied by molecular dynamics simulations. *Proteins*, 58, 110-8 (2005)
156. Danielsson, J., J. Jarvet, P. Damberg, A. Graslund, A. Univ Stockholm & N. d. A. s. d. r. c. peptides: Translational diffusion measured by PFG-NMR on full length and fragments of the Alzheimer A beta (1-40) peptide. Determination of hydrodynamic radii of random coil peptides of varying length. *Magn Reson Chem*, 40, Sp. Iss. SI (2002)
157. Danielsson, Jarvet, Damberg & Graslund: The Alzheimer beta-peptide shows temperature-dependent transitions between left-handed 3-helix, beta-strand and random coil secondary structures. *FEBS J*, 272, 3938-3949 (2005)
158. Buchete, N. & G. Hummer: Structure and dynamics of parallel beta-sheets, hydrophobic core, and loops in Alzheimer's A beta fibrils. *Biophys. J.*, 92, 3032-9 (2007)
159. Huet, A. & P. Derreumaux: Impact of the mutation A21G (Flemish variant) on Alzheimer's beta-amyloid dimers by molecular dynamics simulations. *Biophys. J.*, 91, 3829-40 (2006)
160. Baumketner, A., S. L. Bernstein, T. Wyttenbach, G. Bitan, D. B. Teplow, M. T. Bowers & J. E. Shea: Amyloid beta-protein monomer structure: a computational and experimental study. *Protein Sci*, 15, 420-8 (2006)
161. Sgourakis, N., Y. Yan, S. McCallum, C. Wang & A. Garcia: The Alzheimer's peptides abeta40 and 42 adopt distinct conformations in water: a combined MD / NMR study. *J. Mol. Biol.*, 368, 1448-57 (2007)
162. Nguyen, H. D. & C. K. Hall: Molecular dynamics simulations of spontaneous fibril formation by random-coil peptides. *Proc Natl Acad Sci U S A*, 101, 16180-5 (2004)
163. Urbanc, Cruz, Yun, S. Buldyrev, Bitan, D. Teplow & H. Stanley: In silico study of amyloid beta-protein folding and oligomerization. *Proc Natl Acad Sci U S A*, 101, 17345-17350 (2004)
164. Peng, S., F. Ding, B. Urbanc, S. Buldyrev, L. Cruz, H. Stanley & N. Dokholyan: Discrete molecular dynamics simulations of peptide aggregation. *Physical review E, Statistical, nonlinear, and soft matter physics*, 69, 041908 (2004)
165. Bitan, M. Kirkitadze, Lomakin, S. Vollers, G. Benedek & D. Teplow: Amyloid beta -protein (Abeta) assembly: Abeta 40 and Abeta 42 oligomerize through distinct pathways. *Proc Natl Acad Sci U S A*, 100, 330-335 (2003)
166. Yun, M. R., R. Lavery, Mousseau, K. Zakrzewska & P. Derreumaux: ARTIST: an activated method in internal coordinate space for sampling protein energy landscapes. *Proteins*, 967-75 (2006)

Key Words: Amyloids, Activation-Relaxation Technique, OPEP, Molecular Dynamics, Protein Dynamics, Beta-Amyloid, Alzheimer, Computer Simulation, Beta-Hairpin, Coarse-Grained Potential, Review

Send correspondence to: Dr Normand Mousseau, Departement de physique, Universite de Montreal, c.p. 6128, succ. centre-ville, Montreal (Quebec), Canada, H3C 3J7, Tel: 514-343-6614, Fax: 514-343-02071, E-mail normand.mousseau@umontreal.ca

http://www.bioscience.org/current/vol13.htm