Exploring the early steps of aggregation of amyloid-forming peptide KFFE

2004 J. Phys.: Condens. Matter 16 S5047

(http://iopscience.iop.org/0953-8984/16/44/002)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **16** (2004) S5047–S5054 PII: S0953-8984(04)87012-7

Exploring the early steps of aggregation of amyloid-forming peptide KFFE

Guanghong Wei[1], **Normand Mousseau**[1] **and Philippe Derreumaux**[2]

[1] Département de Physique and Regroupement Québécois sur les Matériaux de Pointe, Université de Montréal, CP 6128, succursale centre-ville, Montréal, QC, H3C 3J7, Canada
[2] Laboratoire de Biochimie, Théorique, UPR 9080 CNRS, IBPC, Université Paris 7 Denis-Diderot, 13 rue Pierre et Marie Curie, 75005 Paris, France

E-mail: normand.mousseau@umontreal.ca

Received 30 September 2004
Published 22 October 2004
Online at stacks.iop.org/JPhysCM/16/S5047
doi:10.1088/0953-8984/16/44/002

Abstract

It has been shown recently that even a tetrapeptide can form amyloid fibrils sharing all the characteristics of amyloid fibrils built from large proteins. Recent experimental studies also suggest that the toxicity observed in several neurodegenerative diseases, such as Alzheimer's disease and Creutzfeldt–Jakob disease, is not only related to the mature fibrils themselves, but also to the soluble oligomers formed early in the process of fibrillogenesis. This raises the interest in studying the early steps of the aggregation process. Although fibril formation follows the nucleation–condensation process, characterized by the presence of lag phase, the exact pathways remain to be determined. In this study, we used the activation–relaxation technique and a generic energy model to explore the process of self-assembly and the structures of the resulting aggregates of eight KFFE peptides. Our simulations show, starting from different states with a preformed antiparallel dimer, that eight chains can self-assemble to adopt, with various orientations, four possible distant oligomeric well-aligned structures of similar energy. Two of these structures show a double-layer β-sheet organization, in agreement with the structure of amyloid fibrils as observed by x-ray diffraction; another two are mixtures of dimers and trimers. Our results also suggest that octamers are likely to be below the critical size for nucleation of amyloid fibrils for small peptides.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Several neurodegenerative diseases such as Alzheimer's disease, Creutzfeldt–Jakob disease and type II diabetes are associated with misfolded proteins that assemble into filament structures called amyloid fibrils [20, 7]. Recent data show that the ability to form amyloid fibrils is a

generic property of *any* polypeptidic chain under destabilizing conditions [18]; even peptides with as few as four to six residues can form well-defined fibrils showing all the characteristics of amyloid fibrils built from large proteins of 100 residues or more [21, 11]. Although a clear relation of causality between amyloid fibrils and associated diseases is still missing, mounting evidence indicates that soluble oligomers, representing the precursor stage in the formation of fibrils, are also toxic in cell cultures [8]. This finding underlines the importance of understanding the first steps in the assembly process of the misfolded proteins and the dynamics of protein folding in general, i.e., to go beyond structural prediction to describe properly the changes in protein conformations.

The assembly leading to fibril formation is well described by the nucleated polymerization model. In this model, the rate-limiting step is the formation of a stable nucleus or seed. This is discerned by the observation of a lag phase in polymer growth. Once a seed is present, the maturation into fibrils is rapid [7]. Understanding the details of growth is hindered by the lack of structural information. As fibrils are insoluble in water and do not yield good-quality crystals; high-resolution structural information is lacking for these structures. Similarly, structural and dynamical information on these soluble oligomers is difficult to obtain experimentally as they are typically short-lived and are present in a wide range of conformations. It is therefore necessary to turn to computer simulations to find the precise arrangement of these proteins. There has already been considerable effort in simulating these systems. For example, the aggregation of small oligomers (dimers, trimers, tetramers, and hexamers) of short amyloid-forming peptides has been studied by various groups by using different computational methods and protein energy models [15, 9, 6, 16, 5, 25, 14].

In this paper, we extend our investigation of the process of amyloid aggregation of peptides and explore the structural properties and the self-assembly process of an octamer of KFFE, a tetrapeptide known to form amyloid fibrils *in vitro* [21]. This is the largest oligomer self-assembled in a non-biased simulation using an explicit main-chain protein model. To do so, we use the activation–relaxation technique (ART) [2, 12] in combination with the optimized potential for efficient peptide structure prediction (OPEP) [3, 4]. The ART-OPEP protocol has already been applied to a number of systems, including hexamers of KFFE [25], and dimers [15, 16] and trimers of the 16–22 fragment of the amyloid protein (Aβ).

We find that eight tetrapeptides assemble into several oligomeric metastable organizations of similar energy, including a double-layer antiparallel β-sheet structure matching the experimental characteristics of the generic cross-β structure. This finding shows that full structural order in fibril requires larger aggregates.

2. Methods

The atomic interactions are described using OPEP. This potential uses a simplified chain representation with all backbone atoms included (i.e., N, H, C_α, C and O) and each side chain modelled by one bead [3, 4] with an appropriate van der Waals radius and geometry. The OPEP energy function, which includes solvent effects implicitly, is expressed as a function of three types of interactions: harmonic potentials for maintaining the geometry of peptides; pairwise contact potential between main chain–main chain, side chain–main chain and side chain–side chain particles (considering all 20 amino acid types); backbone two-body and four-body (cooperative) hydrogen bonding interactions [3, 22].

For our system, the N and C termini of each KFFE chain are neutralized using acetyl and amine groups, respectively. Without special parametrizing, OPEP properly describes the KFFE peptide both in monomeric and oligomeric conformations. For example, we find that 72% of all monomeric structures adopt a random coil conformation and 28% a β-strand

structure [25]. This finding is consistent with the CD analysis in solution at pH 6.0, with β-strands and random coil structures coexisting [21]. We also find that the low-energy structures of KFFE hexamers correspond to well-ordered amyloid-like structures [25], and that OPEP does not lead to well-ordered structures for the mutated peptide KPGE [13].

The sampling of the configurational space is done using ART. A detailed description of this algorithm can be found in [2, 12, 23]. In brief, ART is a generic method for exploring the energy landscape of complex systems. An ART event consists of four basic steps. Starting from a minimum, a randomly chosen chain of the system is pushed outside the harmonic well until a negative eigenvalue of the Hessian matrix appears. The system is then pushed along the eigenvector associated with the negative eigenvalue until the total force approaches zero, indicating a saddle point. The first two steps constitute the activation process. Following this, the system is pushed slightly over the saddle point and is relaxed to a new local minimum. This step is called the relaxation process. Finally, the new configuration is accepted or rejected using the Metropolis criterion at the desired temperature.

ART-OPEP has been used successfully to study the process of folding/aggregation of several small peptides. These include a 14-residue α-helix [23], the second β-hairpin of protein G [24], the Alzheimer's β-amyloid Aβ_{16-22} dimer [15, 16] and trimer [14], and the hexamer of KFFE [25]. Interestingly, we found that the folding trajectories generated by ART are very similar to those obtained by molecular dynamics (MD) and ensemble dynamics for the α-helix [23] and β-hairpin models [24]; the global energy minimum of the Alzheimer's Aβ_{16-22} dimer [16] identified by ART matches exactly the solid state NMR model proposed by Balbach *et al* [1]. The impact of the force field on the ART trajectories and the energy barriers is discussed in [24].

3. Results and discussion

Ten simulations were performed with open boundary conditions; each run generates 9000 events, starting from two different initial states and using different random number seeds, with an acceptance rate of around 50%, and takes about two weeks on a single-processor SGI Altix. These two initial states include two monomers randomly oriented with respect to a preformed tetramer–dimer structure (IN1, 3 runs) in a sphere of diameter 50 Å, an antiparallel dimer and six monomers with random orientations (IN2, 7 runs). IN1 was chosen because a double-layer tetramer–dimer structure was found to be one of the most populated structures for the KFFE hexamer [25]. IN2 was chosen because the predicted aggregation process for trimers and hexamers of amyloid-forming peptides seems to start always with the formation of an antiparallel dimer [16, 14, 25]. In order to sample better the configurational space and identify the lowest-energy conformations, all our runs were performed at a Metropolis temperature of 600 K. In this study, all the structures were produced using the MolMol [10] software, the N-terminal end of each chain is located by a sphere, and the orientation (d_{ij}) between chains i and j is calculated using the scalar product between the end-to-end unit vectors of each chain.

3.1. Structural characterization of the four well-organized lowest-energy arrangements

The ordered structures of lowest energy in our simulations can be classified as belonging to one of four well-organized types (T1–T4 see figure 1). T1 is a double-layer four-stranded parallel–antiparallel β-sheet (figure 1(a)) where the first three strands are antiparallel and the fourth one parallel. T2 is a double-layer four-stranded antiparallel β-sheet (figure 1(b)), with perpendicular inter-sheet β-strands. These two first structures ((a) and (b)) show the cross-β structural characteristics of fibrils, with an inter-sheet distance of about 10–12 Å, in

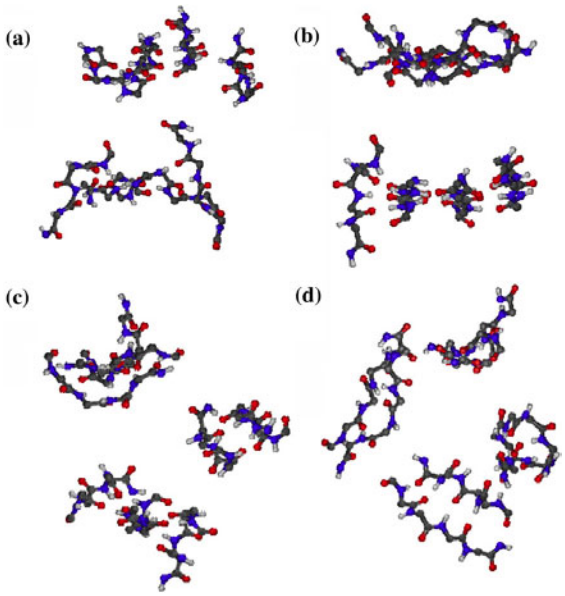

Figure 1. Four well-organized lowest-energy structures generated with ART-OPEP. (a) A double-layer four-stranded parallel–antiparallel β-sheet, (b) a double-layer four-stranded antiparallel β-sheet with perpendicular inter-sheet β-strands, (c) a mixture of two three-stranded antiparallel β-sheet trimers and an antiparallel β-sheet dimer, (d) a mixture of four antiparallel β-sheets dimers.

agreement with the x-ray diffraction measured distances of 10–11 Å between two β-sheets in fibrils [19, 1]. T3 is a mixture of two three-stranded antiparallel β-sheets and an antiparallel β-sheet dimer (figure 1(c)), while T4 consists of four antiparallel β-sheet dimers (figure 1(d)). Here, we discuss the ART-OPEP predicted structures T1 and T2, which are very similar to the experimentally measured fibril structures.

3.2. Self-assembly starting from a preformed tetramer–dimer β-sheet

In order to verify whether and determine how a tetramer–dimer structure can serve as a template for the growth of an octamer, we ran simulations from a tetramer–dimer structure (antiparallel intra-sheet β-strands and perpendicular inter-sheet β-strands) and two monomers (chains 7 and 8) with random orientations (IN1). The chains 4, 2, 1 and 3 are within the initial tetramer, whereas chains 5 and 6 are within the initial dimer. Starting from this conformation, one out of three runs generated a double-layer four-stranded mixed parallel–antiparallel β-sheet (figure 1(a)). The evolution of the relative orientations between two selected pairs of chains (monomer 7 and one strand (chain 6) of the dimer, monomer 8 and another strand (chain 5) of the dimer) is given in figure 2. From this figure, we can see that both monomers 7 and 8 are very flexible as they change their conformations several times during the simulation. The orientation between chains 8 and 5, for example, changes three times during the run, and that between chains 7 and 6 five times, alternating between parallel and antiparallel states. The rest of the structure does not remain frozen as the two monomers move around: the two layers rotate with respect to each other and move by steps from a mostly perpendicular to a parallel state.

During the octamer self-assembly process for this run, not only do monomers 7 and 8 move around and sample their conformation space rapidly, but the two layers rotate about each other and organize into different orientations. Figure 3 shows the inter-sheet orientation

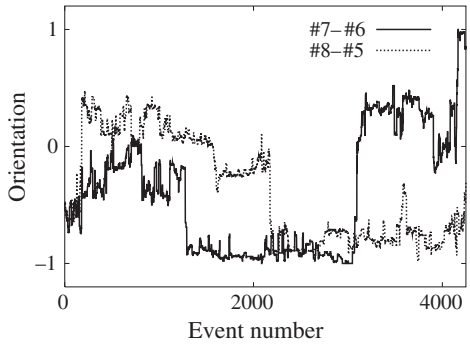

Figure 2. Evolution of the relative orientations for two pairs of chains (monomer 7 and one strand (chain 6) of the dimer, monomer 8 and another strand (chain 5) of the dimer) during the process of formation of the double-layer four-stranded parallel–antiparallel β-sheet. The dimer consists of chains 6 and 5.

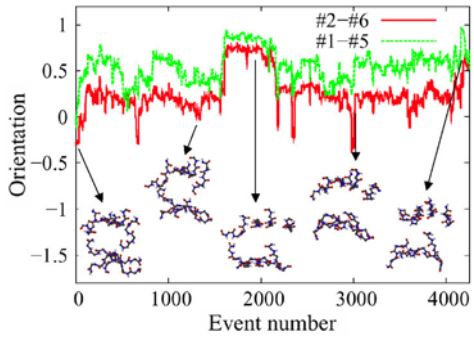

Figure 3. Inter-sheet orientation for two selected pairs 2–6 and 1–5 and five snapshots (cartoon representation of Rasmol) during the self-assembly process of a double-layer four-stranded parallel–antiparallel β-sheet.

for two selected pairs 2–6 and 1–5 and five snapshots (only main chain shown) during the assembly process. In the beginning of the assembly, the two sheets are almost perpendicular ($d_{15} = 0.06$) and chains 7 and 8 have random coil conformations. At event 1314, chain 8 adopts a β-strand conformation and aligns itself with chain 5 to form a β-sheet. The two layers have already started to rotate and have a quasi-parallel orientation ($d_{15} = 0.53$). At event 1900, the two sheets organize to a parallel orientation and monomer 7 starts to align with 6. After an additional 2000 events (event 4200), chains 7, 6, 5, and 8 assemble into a single layer and the two tetramers adopt a parallel-like orientation ($d_{15} = 0.82$).

3.3. Self-assembly starting from a preformed antiparallel dimer

Here, simulations are started from an antiparallel dimer (between chains 1 and 2) and six monomers with random orientations (IN2). Two of the seven runs starting from IN2 aggregate into a double-layer four-stranded antiparallel β-sheet within the number of events allotted. The other five runs converge towards various low-energy conformations, most of them consisting of two three-stranded antiparallel β-sheet trimers and an antiparallel β-sheet dimer.

The total energy, number of hydrogen bonds (H-bonds) and the two-end distances (from the nitrogen atom N of the first residue to the carbon atom C of the last residue of each chain)

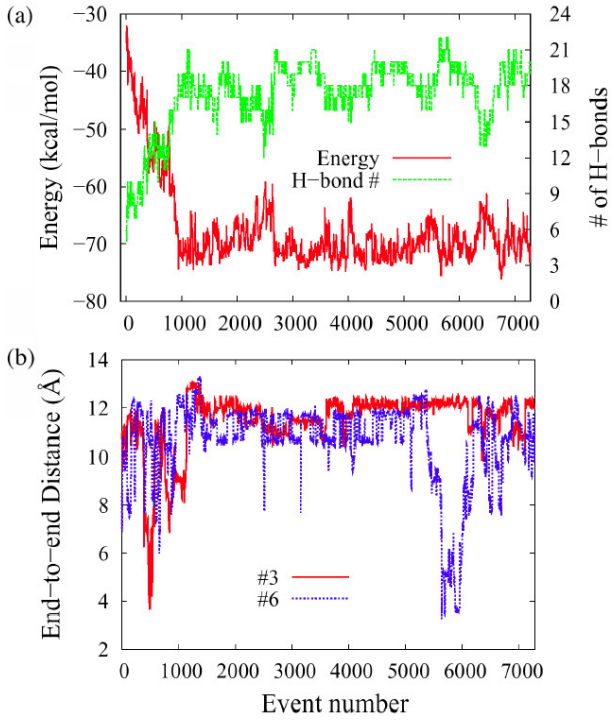

Figure 4. Characterization of the self-assembly pathway as a function of accepted event number leading to the formation of the double-layer four-stranded antiparallel β-sheet. Evolution of (a) the conformational energy and number of H-bonds; (b) end-to-end C α distance for the two chains 3 and 6.

of chains 3 and 6 as a function of event numbers are plotted in figure 4 for one of the runs leading to a fully-formed octamer. As can be seen from this figure, there is considerable energy relaxation (from -36 to -75 kcal mol^{-1}) in the first 1200 events, and the total number of H-bonds rapidly increases from 6 to 21 (see figure 4(a)). The organization of a single tetrameric sheet is already sufficient to stabilize greatly the structure of the octamer, as shown by the energy, which decreases only by an additional 2–3 kcal mol^{-1} with the formation of the second β-sheet. The energy should see a second large jump, however, as the sheets rotate into a parallel orientation.

During the first period, chains 3 and 6 become extended, with an end-to-end distance close to 12 Å, the fully extended value (see figure 4(b)); however, neither of them adopt β-strand conformations (see figure 5 which plots the secondary structure of each residue as a function of event number). By visualizing the structure, we find that a three-stranded antiparallel β-sheet forms within 1200 events (chain 4, 2, 1) from the preformed dimer and chain 3 comes close to chain 1. After 1500 events (at event 2700), chain 3 adopts a β-strand conformation allowing for the formation of a four-stranded antiparallel β-sheet (chains 4, 2, 1 and 3). Once this tetramer occurs, it is rather stable and it remains formed until the end of the simulation. Following the formation of this tetramer, the remaining four chains (6, 5, 8, 7) aggregate with two of them (7 and 8) adopting β-strand conformations. This is followed by a self-organization of the four chains, with chain 6 becoming extended, collapsing, and extending again, causing the second tetramer to form and break repeatedly. In the end, however, we assist at the formation of a two-layer β-sheet octamer, with perpendicular inter-sheet orientation. Based

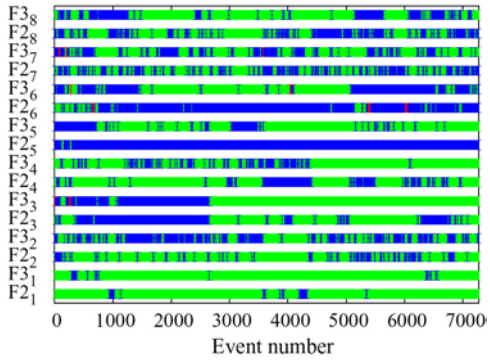

Figure 5. Secondary structure of each residue in the eight chains—the secondary structure is assigned according to the values of dihedral angles ϕ and ψ in [17]; dark stripes for the random coil, and light stripes for the β-strand; Fj_i is the jth ($j = 2$, 3, the two inner residues in each peptide) residue Phe in peptide i ($i = 1$–8). There is essentially no α-helix in this run.

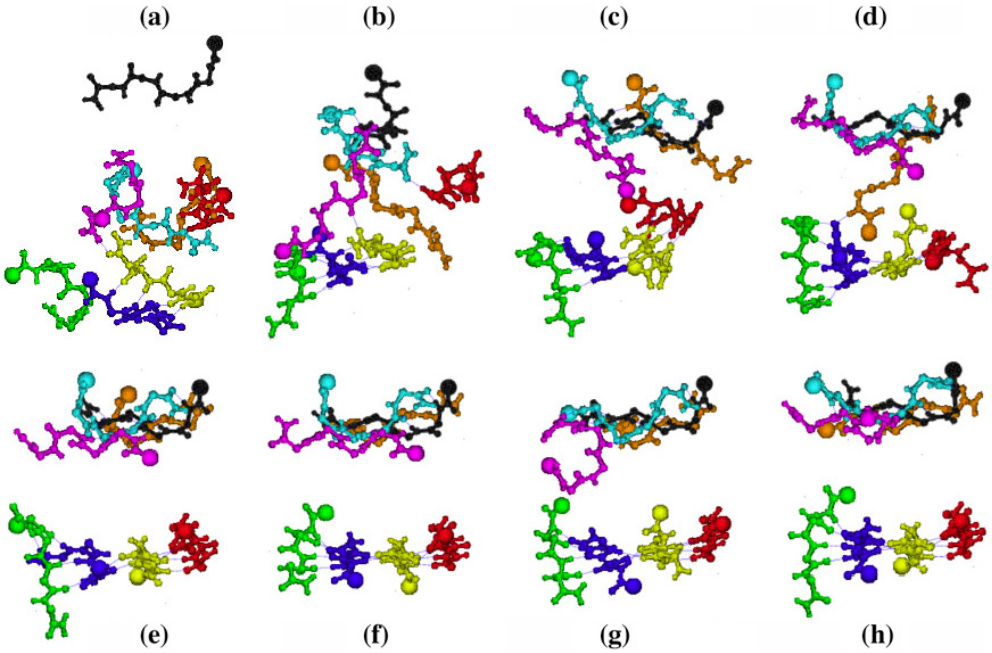

Figure 6. Eight snapshots showing the self-assembly process of a fibril-like octamer with a four-stranded antiparallel β-sheet structure: (a) a preformed dimer and six monomers with random orientations; (b) a trimer and five monomers; (c) a not well-aligned tetramer and four monomers; (d) a tetramer and four-monomer aggregate; (e) a well-aligned tetramer and a not well-aligned tetramer with perpendicular inter-strand arrangement; (f) a well-aligned tetramer and another tetramer; (g) a well-aligned tetramer, a trimer and a monomer; (f) two well-aligned tetramers. The solid line represents H-bonds formed between the carbonyl oxygen and amide hydrogen.

on previous simulations on hexamers [25] and all-atom molecular dynamics simulations of KFFE heptamers in solution at 310 K (unpublished results), we believe that a two-layer four-stranded antiparallel β-sheet can adopt various inter-sheet orientations at room temperature. The assembly process into a fibril-like octamer is summarized in figure 6.

4. Conclusions

The aggregation trajectories presented here for the octamer of KFFE confirm our previous results on the hexamer. Eight chains spontaneously self-assemble into molecular arrangements related to the experimental structure for fibrils. The double-layer structure persists as monomers are added to a preformed hexamer—this growth process minimizes the number of monomers at the surface and hence it maximizes the number of hydrogen bonds while keeping a large proportion of the hydrophobic side groups buried in between the layers. The double-layer four-stranded β-sheet is easily reached starting from a preformed antiparallel dimer, but our energy surface makes it clear that the critical nucleus for the tetrapeptide KFFE likely goes beyond an octamer, maybe between 10 and 20. To fully resolve this issue, however, will require extensive free energy calculation. Much remains to be done to solve the structure and the dynamics of amyloid aggregates.

Acknowledgments

GW and NM are supported in part by the *Fonds québécois de la recherche sur la nature et les technologies* and the *Natural Sciences and Engineering Research Council* of Canada. NM also acknowledges funding from the Canada Research Chair Program. Most of the calculations were done on the computers of the *Réseau québécois de calcul de haute performance* (RQCHP). NM is a Cottrell Scholar of the Research Corporation.

References

[1] Balbach J J, Ishii Y, Antzutkin O N, Leapman R D, Rizzo N W, Dyda F, Reed J and Tycko R 2000 *Biochemistry* **39** 13748
[2] Barkema G T and Mousseau N 1996 *Phys. Rev. Lett.* **77** 4358
[3] Derreumaux P 1999 *J. Chem. Phys.* **11** 2301
[4] Derreumaux P 2000 *Phys. Rev. Lett.* **85** 206
[5] Emanuele P, Gsponer J, Xavier S and Michele V 2004 *J. Mol. Biol.* **340** 555
[6] Gsponer J, Haberthúr U and Caflisch A 2003 *Proc. Natl Acad. Sci. USA* **100** 5154
[7] Hardy J and Selkoe D J 2002 *Science* **297** 353
[8] Kirkitadze M D, Bitan G and Teplow D B 2002 *J. Neurosci. Res.* **69** 567
[9] Klimov D and Thirumalai D 2003 *Structure* **11** 295
[10] Koradi R, Billeter M and Wuthrich K 1996 *J. Mol. Graphics* **14** 51
[11] Lopez D L P M, Goldie K, Zurdo J, Lacroix E, Dobson C M, Hoenger A and Serrano L 2002 *Proc. Natl Acad. Sci. USA* **99** 16052
[12] Malek R and Mousseau N 2000 *Phys. Rev.* E **62** 7723
[13] Melquiond A, Boucher G, Mousseau N and Derreumaux P 2004 in preparation
[14] Santini S, Mousseau N and Derreumaux P 2004 *J. Am. Chem. Soc.* **126** 11509
[15] Santini S, Wei G, Mousseau N and Derreumaux P 2003 *Internet Electron. J. Mol. Des.* **2** 564
[16] Santini S, Wei G, Mousseau N and Derreumaux P 2004 *Structure* **12** 1245
[17] Srinivasan R and Rose G D 1995 *Proteins* **22** 81
[18] Stefani M and Dobson C M 2003 *J. Mol. Med.* **81** 678
[19] Sunde M, Serpell L, Bartlam M, Fraser P, Pepys M and Blake C 1997 *J. Mol. Biol.* **273** 729
[20] Taylor J P, Hardy J and Fischbeck K H 2002 *Science* **296** 1991
[21] Tjernberg L, Hosia W, Bark N, Thyberg J and Johansson J 2002 *J. Biol. Chem.* **277** 43243
[22] Wei G, Derreumaux P and Mousseau N 2003 *J. Chem. Phys.* **119** 6403
[23] Wei G, Mousseau N and Derreumaux P 2002 *J. Chem. Phys.* **117** 11379
[24] Wei G, Mousseau N and Derreumaux P 2004 *Proteins* **56** 464
[25] Wei G, Mousseau N and Derreumaux P 2004 *Biophys. J.* at press