doi:10.1016/j.jmb.2006.01.009 J. Mol. Biol. (2006) 357, 1306–1321

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

A Molecular Dynamics Approach to the Structural Characterization of Amyloid Aggregation

M. Cecchini[1], R. Curcio[1], M. Pappalardo[2], R. Melki[3] and A. Caflisch[1]*

[1]*Department of Biochemistry University of Zurich, Winterthurerstrasse 190 CH-8057 Zurich, Switzerland*

[2]*Dipartimento di Scienze Chimiche, Università di Catania Viale Andrea Doria 6, 95125 Catania, Italy*

[3]*Laboratoire d'Enzymologie et Biochimie Structurales, CNRS Avenue de la Terrasse, 91198 Gif-sur-Yvette, France*

**Corresponding author*

A novel computational approach to the structural analysis of ordered β-aggregation is presented and validated on three known amyloidogenic polypeptides. The strategy is based on the decomposition of the sequence into overlapping stretches and equilibrium implicit solvent molecular dynamics (MD) simulations of an oligomeric system for each stretch. The structural stability of the in-register parallel aggregates sampled in the implicit solvent runs is further evaluated using explicit water simulations for a subset of the stretches. The β-aggregation propensity along the sequence of the Alzheimer's amyloid-β peptide ($A\beta_{42}$) is found to be highly heterogeneous with a maximum in the segment $V_{12}HHQKLVFFAE_{22}$ and minima at S_8G_9, $G_{25}S_{26}$, $G_{29}A_{30}$, and $G_{38}V_{39}$, which are turn-like segments. The simulation results suggest that these sites may play a crucial role in determining the aggregation tendency and the fibrillar structure of $A\beta_{42}$. Similar findings are obtained for the human amylin, a 37-residue peptide that displays a maximal β-aggregation propensity at Q_{10}-$RLANFLVHSSNN_{22}$ and two turn-like sites at $G_{24}A_{25}$ and $G_{33}S_{34}$. In the third application, the MD approach is used to identify β-aggregation "hot-spots" within the N-terminal domain of the yeast prion Ure2p ($Ure2p_{1–94}$) and to design a double-point mutant (Ure2p-$N4748S_{1–94}$) with lower β-aggregation propensity. The change in the aggregation propensity of Ure2p-$N4748S_{1–94}$ is verified *in vitro* using the thioflavin T binding assay.

Keywords: amyloid; Alzhemer's disease; prion; protein aggregation; site-directed mutagenesis

Introduction

Protein folding and unfolding are the most sophisticated and specific ways of promoting and abolishing cellular activity. Failure to fold correctly, or to remain folded correctly, results in a plethora of diseases.[1–3] Some of these diseases originate from amyloidogenic polypeptides that have a high propensity to form ordered aggregates in the extracellular space,[4] e.g. Alzheimer's and Parkinson's diseases, type II diabetes, systemic amyloidosis, and transmissible spongiform encephalopathies.[5,6] Understanding the molecular determinants that cause soluble proteins to aggregate into insoluble fibrils and plaques is therefore an important challenge.

The soluble precursors of amyloid deposits do not share any sequence homology or common fold. However, amyloid aggregates have common structural features: (i) they show the same optical behavior (such as birefringence) on binding certain dye molecules such as Congo red; (ii) present very similar morphologies (long, unbranched and often twisted fibrillar structures a few nanometers in diameter); and (iii) display the characteristic cross-β X-ray diffraction pattern,[7,8] which indicates that the "core" structure is composed of β-sheets running perpendicular to the fibril axis.[9] Hence, regardless of the sequence, the key steps in the aggregation process may be common to all amyloidogenic polypeptides. The ability of a polypeptide chain to self-assemble is not restricted to disease-related proteins. A large number of non-pathogenic peptides and proteins have been shown to form amyloid fibrils under particular solvent, pH and temperature conditions.[10–12] Taken together, these

Abbreviations used: MD, molecular dynamics; APP, amyloid precursor protein; REMD, replica exchange MD; ThT, thioflavin T; EPR, electron paramagnetic resonance; SS, secondary structure; hIAPP, human islet amyloid polypeptide; Aβ42, Alzheimer's amyloid-β peptide.

E-mail address of the corresponding author: caflisch@bioc.unizh.ch

observations indicate that amyloid propensity is a general property of the polypeptide backbone,[13] thus suggesting that under certain conditions any protein above a critical concentration will eventually assemble into ordered aggregates. On the other hand, several studies have shown that the aggregation propensity depends dramatically on amino acid composition and that side-chains influence the kinetics and stability of amyloid fibrils enormously.[14–17] To shed light into the molecular mechanism of amyloid formation and the nature of the energetic contributions that stabilize these structures for an extremely diverse class of polypeptides, atomic-resolution three-dimensional structures are required. As non-crystalline solid material, amyloid aggregates are strongly incompatible with high-resolution techniques for protein structure determination, i.e. X-ray crystallography and liquid state NMR, and, with the exception of the amyloid-like spine formed by a seven-residue peptide,[18] no structure of an amyloid fibril has yet been determined at an atomic level of detail. To obtain structural information, more sophisticated approaches such as solid state NMR,[19–24] site-directed spin labeling,[25–27] cryo-electron microscopy[28,29] and proline-scanning mutagenesis[30] have been used.

Given the difficulty of obtaining high-resolution structures, alternative theoretical and computational approaches have been followed to rationalize the physico-chemical principles of amyloidogenesis and understand the role of the sequence. Very efficient theoretical models to predict protein aggregation propensities from primary structures have been proposed.[31–33] At minimal computational cost, some of these models[32,33] determine putative aggregation-prone regions ("hot-spots") within a protein sequence. Despite remarkable correlation with experimental data, these methods do not provide detailed structural information. Experimental approaches on simplified amyloid systems have also been reported.[17] Remarkably, the full positional scanning mutagenesis of the amyloidogenic peptide STVIIE highlighted both sequence and position dependence of amyloid propensity, even for such a small peptide system. However, the lack of structural detail for the fibrils formed by these peptides has precluded a rational explanation for the origin of the observed mutational effects. Hence, to shed some light into the "hidden" link between protein sequence and amyloid propensity, computational studies providing structural information are required.[34,35]

Here, a novel approach to structurally characterize the propensity towards ordered aggregation of amyloid polypeptides is presented. The procedure is based on the decomposition of a polypeptide chain into overlapping segments and equilibrium molecular dynamics (MD) simulations of a small number of copies of each segment. An efficient implicit solvent model (based on the solvent-accessible surface area[36]) is used to obtain a statistically significant sampling for the trimeric and hexameric systems of each peptide segment. It is important to note in this context that the in-register parallel packing of a seven-residue peptide from the yeast prion protein Sup35 determined by this implicit solvent model[16] is in remarkable agreement with the X-ray microcrystal structure of the cross-β spine (see Materials and Methods).[18] The computational strategy has been designed to predict the position dependence of β-aggregation propensity along the sequence, i.e. the amyloidogenicity profile. From the shape of the profile, amyloidogenic stretches can be discriminated from regions with scarce β-aggregation propensity. Moreover, the atomic detail provided by the MD simulations allows us to interpret the shape of the profile on a structural basis. The method has been tested on three amyloid sequences: the amyloid-β peptide ($A\beta_{42}$), the human amylin and the N-terminal domain of the yeast prion protein Ure2 ($Ure2p_{1–94}$). $A\beta_{42}$ is a product of the proteolytic cleavage of the 695-residue amyloid precursor protein (APP) accomplished by the β and γ-secretases.[37] Amyloid fibrils and plaques formed by full-length Aβ are associated with Alzheimer's disease, which is the most common neurodegenerative disease and accounts for the majority of the dementia diagnosed after the age of 60.[38]

Human amylin, also known as islet amyloid polypeptide (hIAPP), is the major component of pancreatic amyloid deposits found in ~90% of patients with non-insulin-dependent (type 2) diabetes mellitus[39] of which there are about 150 million worldwide.[40] hIAPP is a peptide hormone of 37 amino acid residues produced by cleavage from a pro-amylin precursor protein. It has been shown by X-ray and electron diffraction that hIAPP fibrils are well-ordered cross-β structures.[41] However, a detailed understanding of the fibrillar structure and aggregation properties of full-length hIAPP ($hIAPP_{1–37}$) has yet to be achieved.

Ure2p is a prion from the yeast *Saccharomyces cerevisiae*[42] that acts as a negative regulator of nitrogen metabolism.[43] In its prion state, Ure2p is at the origin of the [URE3] phenotype.[44] *In vitro*, Ure2p aggregates into long, straight filaments that bind Congo red, show green-yellow birefringence and have an increased resistance to proteolysis.[45,46] The prion domain (residues 1–90)[46] is involved in filament formation[47] and contains an unusually high number of Asn, Gln, and Ser residues, i.e. 35%, 12%, and 10% of its 90 residues, respectively. The prion domain of Ure2p is protease-sensitive and poorly structured.[46] However, synthetic $Ure2p_{1–65}$ was shown to readily form fibrils with more than 60% β-sheet.[47]

The agreement between the amyloidogenicity profile obtained by MD simulations and experimental data on $A\beta_{42}$ and $hIAPP_{1–37}$ indicates that the computational approach is descriptive and can be applied to predict the aggregation properties of a polypeptide sequence. The β-aggregation profile highlights critical segments for β-sheet formation

and can be used to guide site-directed mutations that modulate the aggregation tendency. The predictive power of the computational approach is validated experimentally by a double-point mutant of $Ure2p_{1–94}$.

Results

Polypeptide decomposition into segments

Here, the aggregation properties of three amyloid sequences are investigated by the MD computational approach: the human $A\beta_{42}$, $hIAPP_{1–37}$, and the central part of the N-terminal domain of the yeast prion protein Ure2 ($Ure2p_{20–70}$). Due to the large size of the polypeptide chains (42, 37 and 51 residues, respectively) their oligomeric systems cannot be effectively studied by all-atom MD simulations. Therefore, the polypeptide sequence was decomposed into overlapping stretches. By systematically applying a two-residue shift along the sequence, 18 seven-residue and 16 11-residue peptide segments span the 1–41 region of $A\beta_{42}$ (Tables 1 and S2), 16 seven-residue peptide segments span $hIAPP_{1–37}$ (Table S3), and 23 seven-residue peptide segments cover the 20–70 region of Ure2p (Table S4). Each peptide segment was both N-acetylated and C-amidated to reproduce the original context in the full-length sequence. The considerable overlap between neighboring segments allows us to extrapolate the simulation results from the stretches to the polypeptides.

Human amyloid-β peptide ($A\beta_{42}$)

β-Aggregation propensity

The β-aggregation profile of $A\beta_{42}$ was determined by first performing implicit solvent MD simulations of a trimeric system for each of the seven-residue peptide segments (see Table 1). For each segment, the β-aggregation propensity was obtained by averaging the time series of the nematic order parameter $\bar{P}_2$ along the trajectory (see Materials and Methods). At both temperatures of 310 K and 330 K the average value of $\bar{P}_2$ reaches convergence, on a time-scale faster than 1 μs, as indicated by the small standard deviations from three independent runs at 310 K (Figure 1, top). At 330 K, β-aggregation propensity values ranged from 0.51 to 0.89. According to our previous analysis of the amyloid-forming peptides GNNQQNY and QQQQQQQ and nonamyloidogenic peptides SQNGNQQRG and AAAAAAA,[48] these $\bar{P}_2$ values indicate the presence of both aggregation-prone and non aggregation-prone stretches along the $A\beta_{42}$ sequence. Interestingly, the most aggregation-prone segments are not distributed uniformly along the $A\beta_{42}$ sequence but tend to cluster in the region 12–22. The heterogeneity in the aggregation properties of the $A\beta_{42}$ segments is reflected in the free-energy projections along $\bar{P}_2$ (Figure 1, bottom left). Two main scenarios emerge: the first, described by a free-energy profile with a broad minimum at $\bar{P}_2 \sim 0.5$ (broken line),

Table 1. $A\beta_{42}$: seven-residue stretches simulations

Segment	Peptide sequence	Central Residue	Three peptides (μs) 310 K	Three peptides (μs) 330 K	Six peptides (μs) 330 K
$A\beta_{1-7}$	**DAEFRHD**SGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA	4	3 × 1.0	1 × 1.4	1 × 1.5
$A\beta_{3-9}$	DA**EFRHDSG**YEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA	6	3 × 1.0	1 × 1.3	1 × 1.6
$A\beta_{5-11}$	DAEF**RHDSGYE**VHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA	8	3 × 1.0	1 × 1.1	1 × 1.6
$A\beta_{7-13}$	DAEFRH**DSGYEVH**HQKLVFFAEDVGSNKGAIIGLMVGGVVIA	10	3 × 1.0	1 × 1.6	1 × 1.8
$A\beta_{9-15}$	DAEFRHDS**GYEVHHQ**KLVFFAEDVGSNKGAIIGLMVGGVVIA	12	3 × 1.0	1 × 1.4	1 × 1.7
$A\beta_{11-17}$	DAEFRHDSGY**EVHHQKL**VFFAEDVGSNKGAIIGLMVGGVVIA	14	3 × 1.0	1 × 1.3	1 × 1.7
$A\beta_{13-19}$	DAEFRHDSGYEV**HHQKLVF**FAEDVGSNKGAIIGLMVGGVVIA	16	3 × 1.0	1 × 1.9	1 × 1.6
$A\beta_{15-21}$	DAEFRHDSGYEVHH**QKLVFFA**EDVGSNKGAIIGLMVGGVVIA	18	3 × 1.0	1 × 1.7	1 × 1.8
$A\beta_{17-23}$	DAEFRHDSGYEVHHQK**LVFFAED**VGSNKGAIIGLMVGGVVIA	20	3 × 1.0	1 × 1.3	1 × 2.0
$A\beta_{19-25}$	DAEFRHDSGYEVHHQKLV**FFAEDVG**SNKGAIIGLMVGGVVIA	22	3 × 1.0	1 × 1.3	1 × 2.0
$A\beta_{21-27}$	DAEFRHDSGYEVHHQKLVFF**AEDVGSN**KGAIIGLMVGGVVIA	24	3 × 1.0	1 × 1.5	1 × 2.3
$A\beta_{23-29}$	DAEFRHDSGYEVHHQKLVFFAE**DVGSNKG**AIIGLMVGGVVIA	26	3 × 1.0	1 × 1.3	1 × 2.3
$A\beta_{25-31}$	DAEFRHDSGYEVHHQKLVFFAEDV**GSNKGAI**IGLMVGGVVIA	28	3 × 1.0	1 × 1.0	1 × 2.5
$A\beta_{27-33}$	DAEFRHDSGYEVHHQKLVFFAEDVGS**NKGAIIG**LMVGGVVIA	30	3 × 1.0	1 × 1.0	1 × 2.5
$A\beta_{29-35}$	DAEFRHDSGYEVHHQKLVFFAEDVGSNK**GAIIGLM**VGGVVIA	32	3 × 1.0	1 × 1.1	1 × 2.8
$A\beta_{31-37}$	DAEFRHDSGYEVHHQKLVFFAEDVGSNKGA**IIGLMVG**GVVIA	34	3 × 1.0	1 × 1.1	1 × 2.8
$A\beta_{33-39}$	DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAII**GLMVGGV**VIA	36	3 × 1.0	1 × 1.2	1 × 3.2
$A\beta_{35-41}$	DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGL**MVGGVVI**A	38	3 × 1.0	1 × 1.2	1 × 3.2

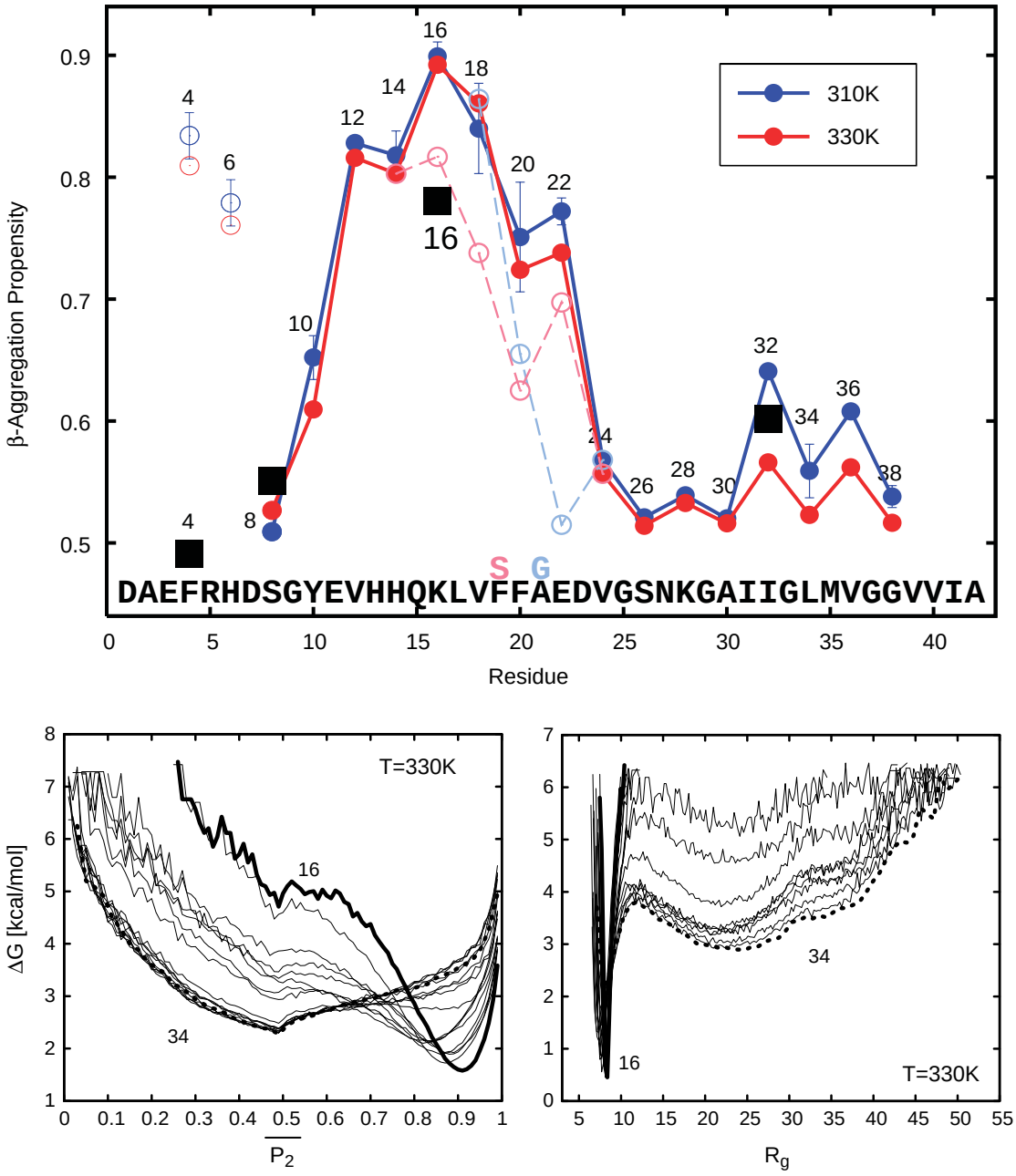

Figure 1. Aβ$_{42}$. Results of constant temperature MD simulations of trimeric seven-residue peptide systems. Top: β-aggregation propensity averaged along the implicit solvent simulations at 310 K (blue), 330 K (red), and the explicit solvent runs (black squares). The data points represent the values of the nematic order parameter $\bar{P}_2$ averaged over the canonical ensemble. The continuous and broken lines are drawn to help the eye for the wild-type and single-point mutants, respectively. The segment identification number corresponds to the position of the central residue in the Aβ$_{42}$ full-length sequence (see Table 1). Error bars on the data points at 310 K represent standard deviations of average β-propensities computed over three independent runs. For some data points at 310 K, e.g. 32 and 36, the error bar is smaller than the symbol. Pink and cyan open circles show the effect of the single-point mutations, F19S[51] and A21G,[52] respectively. Bottom: free-energy projections along $\bar{P}_2$ (left) and the radius of gyration (right) of the oligomeric system at 330 K. Thick and broken lines for segments 16 and 34, respectively, show the emergence of two distinct scenarios in both aggregation (left) and condensation (right) properties. Thin continuous lines represent the 16 remaining segments.

indicates scarce propensity for β-aggregation; the second, described by a steep downhill profile toward a minimum with high orientational order (thick line), highlights amyloid-like sequences.

To obtain insights into the relatively weak β-aggregation propensity detected at the C terminus, the radius of gyration of the oligomeric system R_g was monitored along the implicit solvent trajectories. Again, the free-energy projections along R_g (Figure 1, bottom right) show two different scenarios. The first scenario, described by a free-energy profile with a unique and narrow minimum at $R_g < R_g^C$ (see Materials and Methods), indicates that the simulated peptides are very likely

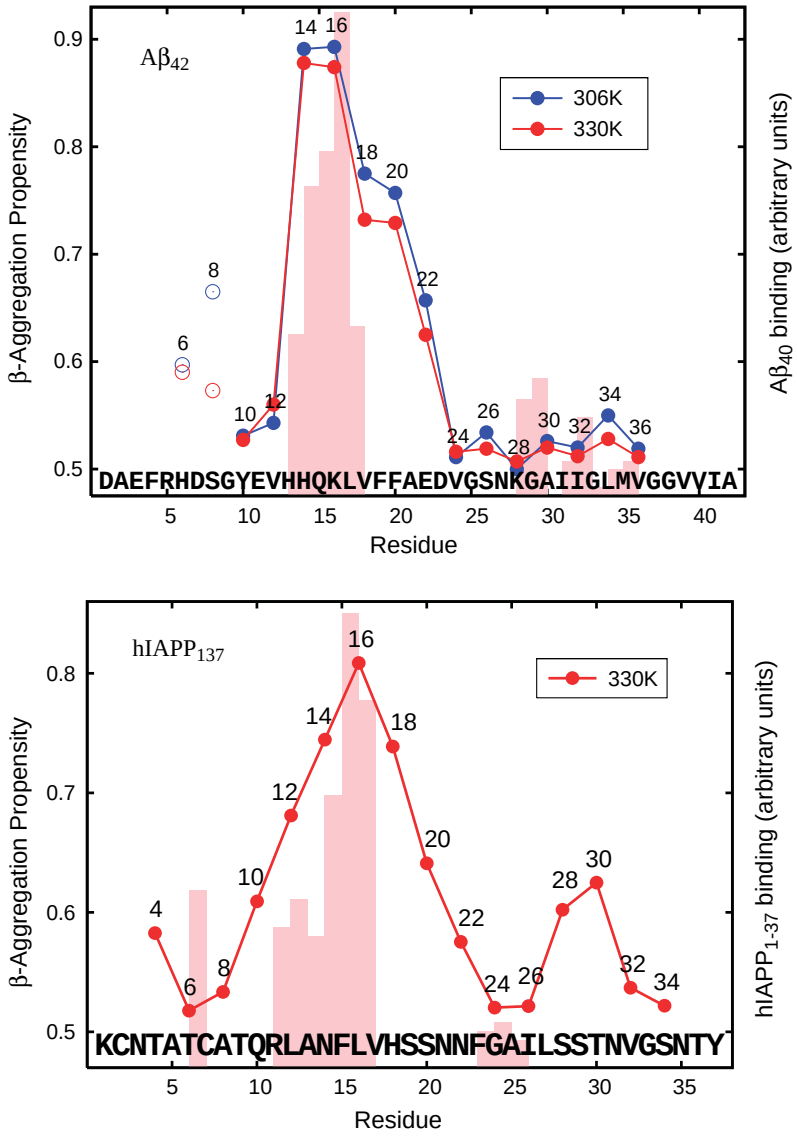

Figure 2. Comparison between β-aggregation propensities from MD simulations and experimental data. See the legend to Figure 1 for the meaning of data points and connecting lines. Top: 11-residue peptide segments of the amyloid-β peptide. The pink bars quantify the binding of the full-length $A\beta_{40}$ to each of 31 overlapping decapeptides (corresponding to residues 1–10 up to 31–40) as measured by radioligand experiments.[49] Bottom: seven-residue peptide segments of the human amylin. The pink bars quantify the binding of the full-length $hIAPP_{1-37}$ to each of 28 over lapping decap peptides (corresponding to residues 1–10 up to 28–37) as measured by immunoblotting experiments.[55]

to aggregate. The second scenario, which shows a prominent broad minimum at $R_g < R_g^C$ on the free-energy profile (broken line), reports that conformations characterized by isolated peptides are favored and the occurrence of condensed states of the system (either ordered or disordered) is rather low. Remarkably, all the stretches compatible with the second scenario belong to the C-terminal part of the amyloid-β peptide (from $A\beta_{21-27}$ to $A\beta_{35-41}$). Since at elevated temperature entropic effects favor the uncondensed state, three additional 1 μs implicit solvent simulations were performed at 310 K for each segment. Interestingly, an aggregation-prone region in the C-terminal part of $A\beta_{42}$ (residues 32–36) emerges at 310 K (Figure 1, blue circles). These simulation results suggest that the condensation propensity, though not sufficient to describe amyloidogenicity, is a necessary condition for the formation of amyloid nuclei. It is not sufficient because amyloidogenic sequences must also have β-sheet propensity, which promotes the assembly into highly ordered structures (see also Supplementary Data).

A comparison of the β-aggregation propensity calculated from the implicit and explicit solvent simulations shows a good agreement except for the N-terminal segment $D_1AEFRHD_7$ (Figure 1, top). The much larger $\bar{P}_2$ value in the implicit solvent runs is a consequence of the approximations inherent to the treatment of charged groups,[36] which are neutralized to prevent vacuum-like artifacts like the excessive formation of salt-bridges. In the implicit solvent simulations of $D_1AEFRHD_7$, the lack of strong Coulombic repulsion between side-chains with the same charge does not prevent formation of in-register parallel β-sheets. The first two N-terminal segments $D_1AEFRHD_7$ and E_3-$FRHDSG_9$ contain four and three charged side-chains, respectively, whereas the remaining seven-residue segments have between zero and two formal charges. For this reason, the implicit solvent β-aggregation profiles are more reliable for the 8–42

region of the $A\beta_{42}$ peptide, where they also show good agreement with the explicit solvent simulation results.

Implicit solvent replica exchange MD (REMD)[48] simulations of 11-residue segments were also used to derive the $A\beta_{42}$ amyloidogenicity profile (Figure 2, top). At 306 K and 330 K, the profiles look similar with higher propensity for the lower temperature. For all the aggregation-prone segments (8, 14, 16, 18, 20 and 22), the analysis of the trajectories by a polar order parameter ($\bar{P}_1$; see Cecchini *et al.*[48]) revealed a statistically relevant predominance of in-register parallel β-sheets and, with the exception of stretch 22, negligible antiparallel arrangements (data not shown). The high-propensity region encompasses residues 14–22. Interestingly, the same region was identified by means of radioligand experiments[49] as the most prone to bind full-length $A\beta_{42}$ (pink bars in Figure 2). The radioligand experiments were carried out with overlapping ten-residue stretches, which is very close to the simulation systems. A second region located at the C terminus, which is missing from the 11-residue segments aggregation profile, was detected by the experiments. However, the binding of $A\beta_{42}$ to the C-terminal decapeptides was considerably less prominent and probably mediated by hydrophobic rather than specific interactions.

An in-depth comparison of the amyloidogenicity profiles obtained from seven and 11-residue peptide simulations (Figures 1 and 2, respectively) provides additional information. The profiles are qualitatively similar and display a prominent high-propensity region located in the central part of the sequence. However, the effect of the increased peptide length is not negligible; a considerably lower β-aggregation propensity is detected at the N terminus (residues 1–6). When the peptide length is increased, the resulting stretches are more likely to include both aggregation-prone and non-aggregation-prone segments. Therefore, the β-aggregation propensity decreases in regions of the sequence with mixed properties. To verify that the length of the segments identified on the $A\beta_{42}$ sequence had no impact on the results, the β-aggregation propensity profile was recalculated by considering seven-residue substretches along the 11-residue simulation trajectories (e.g. D_1AEFRHD$_7$, E_3-FRHDSG$_9$ and R_5HDSGYE$_{11}$ from D_1-AEFRHDSGYE$_{11}$). Both β-aggregation propensity and secondary structure profiles are in good agreement with those obtained from seven-residue peptide simulations (Figure S4 in Supplementary Data). Thus, the simulation results are robust with respect to the choice of the segment length.

Secondary structure analysis

To interpret the amyloidogenic trend in the central zone (stretches 14, 16, 18, 20 and 22), a secondary structure analysis of the conformations saved along the trajectories was performed. The analysis showed that β-aggregation propensity correlates with β-strand content, anticorrelates with α-helical content, and seems very sensitive to β-turn or bend propensity (Figure 3 and Figure S5 in Supplementary Data). The β-aggregation profile in the segment 14–22 is influenced by the location of a turn-like segment ($G_{25}S_{26}$) and the α-helical/β-strand equilibrium. The latter is consistent with NMR studies that highlighted a trend to helical structures for the segment 16–24 in aqueous solution.[50] The identification of four turn-like

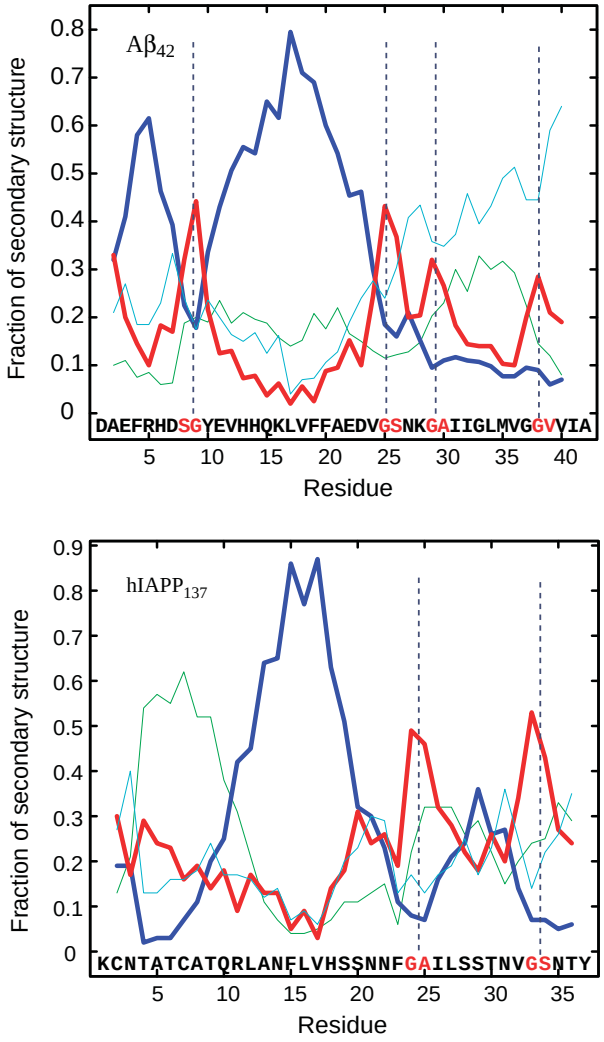

Figure 3. Secondary structure profiles. Single-residue secondary structure propensities have been extracted from REMD trajectory segments at 306 K of trimeric 11-residue systems for $A\beta_{42}$ (top) and from constant temperature MD trajectories at 330 K of trimeric seven-residue systems for $hIAPP_{1–37}$ (bottom). To obtain the propensity value of residue *i*, averages were taken over all stretches containing residue *i*. Green, blue, red, and cyan lines correspond to α-helical, β-strand, turn or bend, and random coil content, respectively. The simulation results indicate the presence of turn-like segments (red letters), which play a key role in determining amyloid aggregation properties. The vertical broken lines mark the borders of the regions identified by the specific location of the turns along the sequence.

peaks (S_8G_9, $G_{25}S_{26}$, $G_{29}A_{30}$, and $G_{38}V_{39}$) helps to interpret the shape of both seven and 11-residue profiles as well as the differences between the two. When $A\beta_{42}$ is dissected into 11-residue segments, the stretches at the N terminus always include the first potential turn (S_8G_9) in the middle of the sequence and therefore the amyloidogenic content is low. In contrast, the heptamer $A\beta_{3-9}$ has the S_8G_9 turn at the C terminus and shows a high β-aggregation propensity. Similar considerations explain the low propensity of the 11-residue profile in the C-terminal region of $A\beta_{42}$. With longer segments, either the third ($G_{29}A_{30}$) or the fourth ($G_{38}V_{39}$) turn-like site is always included in the stretches and the β-aggregation propensity is suppressed. On the other hand, the seven-residue stretches between $G_{29}A_{30}$ and $G_{38}V_{39}$ are responsible for the peak at residues 31–37 (Figure 1 and Figure S2 in Supplementary Data). Taken together, the simulation results show that the β-aggregation profile of $A\beta_{42}$ is strongly modulated by the position of four turn-like sites along the sequence.

Single-point mutants of $A\beta_{42}$

The aggregation properties of four familial disease-related variants of $A\beta_{42}$, i.e., the Arctic (E22G), Dutch (E22Q), Italian (E22K) and Flemish (A21G) mutants, and one non-pathological variant obtained by random mutation[51] (F19S) were investigated. *In vitro* studies[51–53] have shown that the E22G, E22Q and E22K mutations accelerate fibril formation while A21G and F19S decrease the fibrillogenesis rate with respect to wild-type $A\beta_{42}$. Starting from the β-aggregation profile of the wild-type sequence (Figure 1), mutational effects can be predicted at moderate computational cost, i.e. by repeating only the implicit solvent simulations of the seven-residue stretches affected by the mutation (Table S1 in Supplementary Data). Seven-residue segments are preferred to 11-residue segments because of the lower computational cost, which allows us to investigate a larger number of mutants.

As shown in Figure 1, the lower aggregation propensity of the F19S and A21G variants is reproduced correctly. Interestingly, the disease-related mutant E22G has a profile (data not shown) similar to that of wild-type $A\beta_{42}$, indicating that the simulation-based approach is able to distinguish the subtle difference between A21G and E22G. On the other hand, the fact that three disease-related variants E22G, E22Q and E22K have profiles similar to wild-type (data not shown) suggests that the approach is less sensitive in the very-high propensity region, which could be a consequence of the reduced dimensionality of the simulation system, i.e. number of peptides smaller than in the nucleus and/or short segment length. In this context it is important to note that previously published explicit water simulations of the monomeric $A\beta_{10-35}$ peptide and its E22Q mutant do not support the hypothesis that the Dutch E22Q variant leads to a higher β-structure propensity,[54] in agreement with the present implicit solvent results.

Human amylin ($hIAPP_{1-37}$)

The β-aggregation profile of $hIAPP_{1-37}$ was determined using the same approach as for $A\beta_{42}$. Sixteen implicit solvent MD simulations of three seven-residue peptide segments were performed (Table S3 in Supplementary Data). At 330 K, β-aggregation propensity values ranged from 0.52 to 0.81 (Figure 2, bottom). The resulting profile highlights two well-defined hot-spots along the $hIAPP_{1-37}$ sequence, with the first (residues 10–22) more prominent than the second (residues 28–30).

A systematic mapping of the $hIAPP_{1-37}$ sequence for the identification of domains that can potentially mediate molecular recognition and lead to amyloid fibril formation has been performed by Gazit and co-workers.[55] Their *in vitro* immunoblotting experiments showed that the region most prone to bind the full-length $hIAPP_{1-37}$, i.e. the recognition domain, is located at the center of the sequence (residues 7–21). The simulation results are in good agreement with *in vitro* findings (Figure 2, bottom). In particular, the NFVLH pentapeptide suggested as the "core" of the recognition motif[55] is included in the seven-residue stretch $hIAPP_{13-19}$ (central residue 16) that showed the highest β-aggregation propensity *in silico*. The largest discrepancy is located at the region 24–26, which was not identified by the simulation-based approach. However, the immunoblotting signal is very weak in this region and likely to originate from hydrophobic rather than specific interactions.

The structural analysis performed on $A\beta_{42}$ was repeated on the $hIAPP_{1-37}$ sequence. Average single-residue secondary structure propensities were extracted from simulation trajectories and used to draw the profiles shown in Figure 3, which highlights two short turn-like segments ($G_{24}A_{25}$ and $G_{33}S_{34}$) corresponding to regions of reduced β-aggregation propensity. Again, the simulation results suggest that turn-like sites strongly modulate the aggregation propensity of amyloid polypeptides. It is worth noting, though that a third β-aggregation propensity minimum observed around Thr6 is due to a strong α-helical propensity (green line in Figure 3 bottom) and not to a turn- or bend-site. As mentioned above, the α-helical/β-strand equilibrium can modulate the β-aggregation propensity of a polypeptide chain.

N-terminal domain of the prion protein Ure2 ($Ure2p_{1-94}$)

As for $A\beta_{42}$ and $hIAPP_{1-37}$, the β-aggregation profile of $Ure2p_{20-70}$ was determined by performing implicit solvent MD simulations of a trimeric system for each of the seven-residue peptide segments (Table S4 in Supplementary Data). At 330 K, β-aggregation propensity values ranged from 0.51 to 0.76 (Figure 4, top). Three aggrega-

tion-prone regions intercalated with short non aggregation-prone segments are highlighted: residues 33–39, 45–49, and 55–61. It is worth noting that one of the two central regions with low β-aggregation propensity corresponds to three consecutive serine residues located at positions 51–53. Apparently, these serine residues reduce the local aggregation tendency by splitting a poly (N) stretch in two segments. To further investigate the role of serine residues, six variants of the stretch Ure2p$_{44–50}$ (NNNNNNN) were modeled by considering all possible single and double-point N-to-S mutants at positions 47, 48 and 49 (Table S5 in Supplementary Data). Six additional simulations were run at 330 K and β-aggregation propensities computed. As shown in Figure 4 (top, blue triangles), a strong position dependence on mutation is observed in agreement with recent experimental findings.[17] Furthermore, the central positions (residues 47 and 48) are more sensitive than the lateral ones (residue 49). The N-to-S mutations reduce the aggregation propensity of Ure2p$_{44–50}$ with the lowest tendency for the double mutant Ure2p-N4748S$_{44–50}$. To study the effect of this double-point mutation on Ure2p$_{1–94}$, aggregation simulations of all stretches affected by the mutations, i.e. Ure2p$_{42–48}$, Ure2p$_{44–50}$, Ure2p$_{46–52}$ and Ure2p$_{48–54}$, were carried out. The N4748S double-point mutation is responsible for the disappearance of the hot-spot at residues 45–49 (empty circles in Figure 4 top) and therefore is predicted to strongly affect the assembly behavior of the whole prion domain. To validate the *in silico* prediction, the assembly kinetics of the N-terminal domain of wild-type Ure2p (Ure2p$_{1–94}$) and double-point mutant (Ure2p-N4748S$_{1–94}$) were compared using the thioflavin T (ThT) binding assay. For the latter, a pronounced increase in the lag phase of the assembly reaction and a lower level of ThT fluorescence at steady state were observed (Figure 4, bottom). From this observation, we conclude that the substitution of asparagine by serine residues at position 47 and 48 hinders the

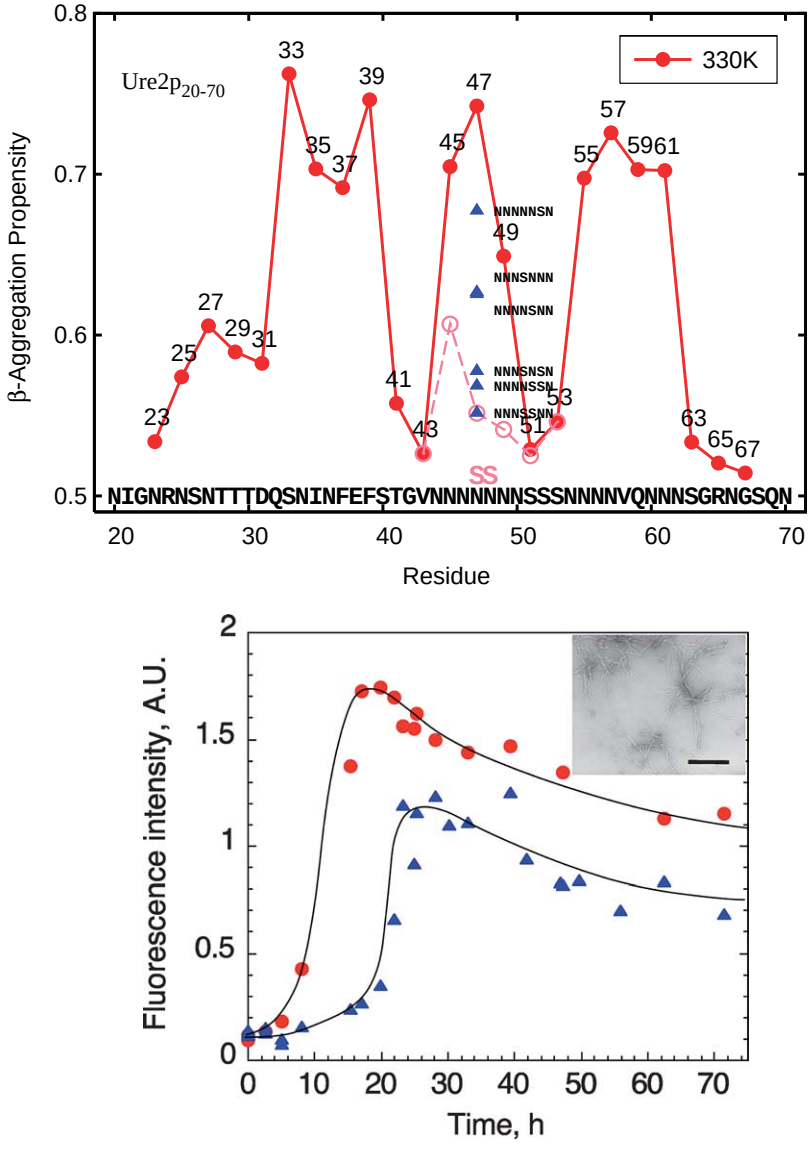

Figure 4. N-terminal domain of the prion protein Ure2: MD simulations and *in vitro* validation. Top: values of the β-aggregation propensity from 330 K constant temperature MD trajectories of trimeric seven-residue peptide systems are shown by filled circles. The segment identification number corresponds to the position of the central residue in full-length Ure2p (Table S4 in Supplementary Data). Blue triangles indicate β-aggregation propensities of single and double-point N-to-S mutants at positions 47, 48 and 49 of the stretch Ure2p$_{44–50}$ (segment 47). Open circles connected by a broken line highlight the mutational effect of the double mutation N4748S on the aggregation properties of Ure2p$_{1–94}$. Bottom: assembly kinetics of wild-type (Ure2p$_{1–94}$, red circles) and Ure2p$_{1–94}$ N4748S variant (blue triangles) monitored by ThT binding. The N4748S substitutions have a dramatic effect on the lag phase preceding assembly and validate the simulation results. Ure2p$_{1–94}$ fibrils are 4 nm wide and are shown in the electron micrograph (inset; the scale bar represent 100 nm).

aggregation process of the N-terminal domain as predicted by the implicit solvent simulations.

Concluding Discussion

A "divide-and-conquer" approach to investigate the aggregation properties of amyloid polypeptides is presented. The amino acid sequence is first decomposed in overlapping segments. Then, implicit solvent MD simulations of oligomeric systems are performed for each segment. The use of an implicit model of the solvent[36] allows for equilibrium sampling (total simulation time of hundreds of microseconds) starting from peptides well separated in space, i.e. without intermolecular contacts. To validate the structural stability of the ordered aggregates observed in the implicit solvent runs, 50 ns control simulations with explicit water are carried out for a subset of the segments.

The MD procedure is used here to determine the position dependence of the β-aggregation propensity along the polypeptide sequence, i.e. the β-aggregation profile. Despite higher computational demand with respect to analytical models recently developed to predict β-aggregation propensities,[31–33] the present method provides a structural interpretation of the β-aggregation profile, which is essential to rationalize the sequence dependence and predict mutational effects on amyloid aggregation. The use of segments to investigate the β-aggregation properties of a full-length sequence is justified, especially for parallel in-register aggregates. The dissected stretches are N-acetylated and C-amidated to reproduce their original context in the full-length sequence. Assuming in-register parallel arrangements, aggregation MD simulations of short stretches are a good approximation of the fibrillar environment and the observations made on the stretches can be extended to full-length polypeptides.

Up to now, the details of the amyloid structure and the extent to which it is uniquely defined are unclear. Initially, structural models with antiparallel β-sheets were favored.[56,57] However, solid state NMR measurements revealed that amyloid fibrils formed by $A\beta_{10–35}$[58] and by full-length $A\beta_{40}$[59,60] contain parallel β-strands exactly in register. Moreover, electron paramagnetic resonance (EPR) studies[26] and very recent vibrational dipolar coupling measurements[61] on fibrils formed by spin-labeled and isotope-labeled $A\beta_{40}$, respectively, have provided further evidence for the parallel in-register arrangement. Also, spin labeling experiments on $hIAPP_{1–37}$ fibrils have indicated an in-register parallel organization of the β-strands.[27] The compelling experimental[17,62–65] and computational[16,66] evidence that side-chains strongly influence amyloid aggregation suggests that a parallel organization of the strands in the fibrils should be, in general, preferred to antiparallel. By construction, in-register parallel arrangements favor the interactions of hydrophobic/aromatic side-chains. Hence, a preference for parallel aggregates is expected for polypeptide sequences with few charged residues. Short stretches with charged groups at the termini might prefer the antiparallel arrangement.[34]

The aggregation properties of the Alzheimer's amyloid-β peptide have been investigated by applying a two-residue shift: 18 seven-residue and 16 11-residue peptide segments were defined along the $A\beta_{42}$ sequence. Although the stretches of the two sets are rather diverse in both amino acid composition and length, the resulting amyloidogenicity profiles highlight the same region (from Val12 to Asp22) as the major hot-spot. Interestingly, this central zone is also the most prone to bind the radiolabeled full-length $A\beta_{40}$ peptide, as quantified by densitometry.[49] Since the autoradiography experiments were carried out with short (ten-residue) overlapping stretches, the comparison between *in vitro* and *in silico* results is appropriate and the former validates the latter. Furthermore, the central region includes the $K_{16}LVFF_{20}$ pentapeptide that was shown to be essential for amyloid fibril formation.[49] Although with a lower tendency, the C-terminal segment (residues 31–37) is also found to be aggregation-prone. In agreement with this *in silico* result, ThT fluorescence assays have shown that residues 30–35 (AIIGLM) promote the self-assembly by accelerating the aggregation process.[67] The enhanced β-aggregation propensity detected at the N terminus by the implicit solvent runs is rather surprising and in disagreement with solid state NMR[23] and EPR measurements.[26] It is likely that the approximations inherent to the solvation model and, in particular, the neutralization of formal charges, are too crude to correctly reproduce the behavior of polypeptide segments with many charged side-chains. Explicit solvent simulations started from parallel β-sheet conformations of segments located at the N terminus unveiled their marginal structural stability, in agreement with experiments. The structural details emerging from the simulations are consistent with the model for $A\beta_{40}$ fibrils derived from solid state NMR spectroscopy.[23] In the NMR model, the amyloid-β peptide bends to generate double-layered sheets that can pack in a parallel arrangement. Interestingly, the segments 12–24 and 31–37 correspond to the β-strands of the NMR model.

Thanks to the atomic detail provided by the MD simulations, the β-aggregation profile could be structurally characterized. Secondary structure analysis of the MD trajectories unveiled the presence of four turn-like sites along the amyloid-β sequence: S_8G_9, $G_{25}S_{26}$, $G_{29}A_{30}$, and $G_{38}V_{39}$. Interestingly, the location of the first three turns had been suggested by solution NMR,[68] solid state NMR[23] and proline scanning mutagenesis,[30] respectively. Although the four sites with turn propensity could have been detected by algorithms for secondary structure prediction, the consequences of such propensity within the context of an oligomeric system can be determined only by the

MD-simulation approach. The identified turn-like segments correspond to large drops in β-strand propensity and are located at the borders of aggregation-prone regions (Figure 3, top). Hence, their specific position on the sequence determines the location and width of the aggregation hot-spots, which are supposed to drive amyloid fibril formation and to have an influence on the fibrillar conformation of $A\beta_{42}$. In this regard, it is worth noting that although the structural model proposed by Tycko and co-workers[23] has a single turn located at residues 25–26, the original solid state NMR chemical shifts are fully compatible with the presence of a second turn located at residues 29–30, as suggested by the MD simulation results. An alternative structural model including a second turn and a different intramolecular register between the strands cannot be excluded because of the lack of experimental data in that region of the sequence and the usage of a minimization protocol by Petkova *et al.*[23] unable to investigate the whole conformational space compatible with experimental constraints.

The simulation-based approach has been further tested on the human amylin polypeptide ($hIAPP_{1–37}$) and the N-terminal domain of the yeast prion protein Ure2 ($Ure2p_{20–70}$). Unlike $A\beta_{42}$, these two polypeptide sequences contain very few charged side-chains (i.e. 2/37 and 4/51 in $hIAPP_{1–37}$ and $Ure2p_{20–70}$, respectively). Moreover, these charges are separated along the sequence (Lys1 and Arg11 in $hIAPP_{1–37}$; Arg24, Asp31, Glu38, and Arg65 in $Ure2p_{20–70}$) so that each blocked heptapeptide contains a maximum of one charge. Hence, explicit water runs were not deemed necessary. Two aggregation-prone regions have been identified along the $hIAPP_{1–37}$ sequence: a major hot-spot from Gln10 to Asn22, and a minor one from Ser28 to Val32. In contrast with experimental data obtained without blocking groups at the peptide termini,[69] the present simulation analysis indicates that the blocked NNFGAIL segment does not show high β-aggregation propensity. Interestingly, recent explicit water simulations of an octameric system of blocked NFGAIL peptides have also reported low aggregation tendency.[70] In fact, despite the usage of *ad hoc* conformational restraints, i.e. the main chains were completely restrained to ideal β-sheet conformations, only 8% of the sampled octamers were well ordered. The apparent disagreement between the simulation results (this work and work done by Wu *et al.*[70]) and the experiments suggests that short peptide stretches may show different aggregation properties if unblocked. Hence, it is more appropriate to consider blocked peptides to infer the aggregation properties of a polypeptide sequence from its segments.

The structural characterization of the β-aggregation profile of $hIAPP_{1–37}$ unveils the presence of two specific turn-like segments ($G_{24}A_{25}$ and $G_{33}S_{34}$). Similarly to what was found for $A\beta_{42}$, these two sites determine the overall shape of the aggregation profile, i.e., the location and extension of the hot-spots. It is worth noting that the $A\beta_{42}$ and $hIAPP_{1–37}$ aggregation profiles are strikingly similar (Figures 1 and 2). Both share a major aggregation hot-spot in the middle of the sequence, a less aggregation-prone region in the hydrophobic tail and a turn-like segment between them. Despite the rather low level of sequence identity (~21%) and similarity (~36%), the β-aggregation profiles suggest that $A\beta_{42}$ and $hIAPP_{1–37}$ might have similar fibrillar structures. In accord with these considerations, it has been found that $A\beta_{42}$ fibrils can act as efficient seeds for $hIAPP_{1–37}$ aggregation,[71] thus implying that at least under certain conditions $hIAPP_{1–37}$ can adopt a structure similar to that of $A\beta_{42}$ in the fibrillar form.

Taken together, the MD-simulation results of $A\beta_{42}$ and $hIAPP_{1–37}$ provide further evidence that alternating β-strands and turn (or bend) segments might be a general feature of amyloid polypeptides, as suggested by Kajava *et al.*[72,73] Assemblies with completely elongated peptide backbones are likely to be less favorable than partially folded arrangements of β-strands because of entropic effects as well as the tighter packing and minimal solvent exposure of hydrophobic residues in the latter.[23] In this view, the identified turn-like segments are expected to be very sensible to mutation and therefore optimal targets for reducing amyloid propensity.

When removed from its natural environment, $Ure2p_{1–94}$ assembles into 4 nm wide fibrils of amyloid nature (Figure 4). Three aggregation-prone regions have been identified along the sequence of $Ure2p_{20–70}$: residues 33–39, 45–49, and 55–61. Again, the presence of aggregation hot-spots intercalated with non-aggregation-prone segments has been observed. The simulation-based approach has been successfully applied to guide site-directed mutagenesis for reducing the amyloidogenic tendency of $Ure2p_{1–94}$. The double-point mutation N4748S designed *in silico* to reduce aggregation propensity has been verified experimentally (Figure 4).

In conclusion, the MD simulation approach yields the amyloidogenicity profile along a polypeptide sequence and the secondary structure propensity of its overlapping segments in the context of an ordered aggregate. The combination of both types of information is very helpful for the understanding of the sequence and structure determinants of amyloid fibril formation. The computational strategy may be ultimately used to guide the rational design of synthetic peptidic and non-peptidic molecules that hinder or prevent amyloid aggregation.

Materials and Methods

Implicit solvent simulations of aggregation

To simulate peptide aggregation, a strategy similar to that described by Gsponer *et al.*[16] was followed. Implicit

solvent simulations of aggregation were performed with the program CHARMM.[74] The oligomeric peptide systems were modeled by explicitly considering all heavy atoms and the hydrogen atoms bound to nitrogen or oxygen atoms (PARAM19 potential function[74,75]). The remaining hydrogen atoms are considered as part of the carbon atoms to which they are covalently bound (extended atom approximation). An implicit model based on the solvent-accessible surface was used to describe the main effects of the aqueous solvent on the solute.[36] The CHARMM PARAM19 default cutoffs for long range interactions were used, i.e., a shift function[74] was employed with a cutoff at 7.5 Å for both the electrostatic and van der Waals terms. The model is not biased toward any particular secondary structure type. In fact, the same force field and implicit solvent model have been recently used in MD simulations of aggregation,[16,48] folding of structured peptides (α-helices and β-sheets) ranging in size from 15 to 31 residues,[76–78] and small proteins of about 60 residues.[79,80] Moreover, the in-register parallel packing of three GNNQQNY peptides predicted by this model[16] has been recently validated by the X-ray microcrystal structure of the cross-β spine:[18] the β-strand alignment, the stacking interactions of the tyrosine rings and the hydrogen bonds between amide groups are essentially identical (compare Figure 2 of Gsponer *et al.*[16] with Figure 2(e) of Nelson *et al.*[18]). The same force field, solvation model and simulation protocol have been applied to polypeptide segments experimentally known not to form amyloid structures, i.e. the nonapeptide SQNGNQQRG (corresponding to Sup35 residues 17–25 with the Gln/Arg mutation at position 24, which showed solubility *in vivo* and *in vitro*[81]) and the alanine heptapeptide.[82] No ordered β-aggregate was observed in these control simulations,[16,48] which is particularly remarkable for the SQNGNQQRG sequence given its similarity to the amyloidogenic GNNQQNY peptide.

In the case of $A\beta_{42}$, the following constant temperature MD runs were performed: (i) 18 simulations of three seven-residue peptide copies at 330 K to monitor β-aggregation propensity along the sequence and highlight possible aggregation hot-spots; (ii) 18 simulations of three seven-residue peptide copies at 310 K to investigate the temperature effect on the aggregation properties of the overlapping stretches; (iii) 18 simulations of six seven-residue peptide copies at 330 K to investigate the effect of the system size, i.e. the number of molecules in the simulation box, on both β-aggregation propensity and structural properties of the aggregates; (iv) 13 310 K and four 330 K runs of three seven-residue peptide copies to predict mutational effects on the aggregation properties of $A\beta_{42}$; and (v) 16 simulations with three 11-residue peptide copies at 330 K to study the dependency of β-aggregation propensity on the length of the overlapping segments. To guarantee the correct sampling of peptide conformational space in physiological conditions,[48] the self-assembly of long segments (11-residue) was investigated by replica exchange molecular dynamics (REMD) simulations. In an REMD run, different copies of the system ("replicas") are simulated at the same time but at different temperature values. Each replica evolves independently by MD and every t_{swap} states *i,j* with neighbor temperatures are swapped (by velocity rescaling) with a probability $w_{ij}=\exp(-\Delta)$,[83] where $\Delta\equiv(\beta_i-\beta_j)(E_j-E_i)$, $\beta=1/kT$ and E is the effective energy (potential and solvation energy). During the simulation, each replica visits all temperatures of the set and realizes a free random walk in temperature space. High-temperature simulation segments facilitate the crossing of the energy barriers while low temperature ones explore energy minima in detail. Thus, the temperature swapping determines a random walk in energy space, which improves sampling efficiency. In this study, ten replicas were used with temperatures (in K): 294, 306, 318, 330, 343, 356, 369, 383, 397, 413. By using fixed value of $\Delta t_{swap}=10{,}000$ MD steps (20 ps),[48] temperature values were adjusted by trial and error until the acceptance ratios of exchange between neighbor temperatures converged to values between 40% and 50%.

In a similar fashion, to determine the β-aggregation profile and identify the aggregation hot-spots along both $hIAPP_{1–37}$ and $Ure2p_{20–70}$, 16 and 23 simulations of three seven-residue peptide copies at 330 K were carried out, respectively. Finally, ten simulations of three seven-residue peptide copies were performed at 330 K to investigate mutational effects on $Ure2p_{1–94}$ aggregation properties.

All implicit solvent simulations were performed starting from random conformations, positions, and orientations of the peptide copies. In the initial random positions there was no intermolecular contact, i.e. the peptides were separated in space. Each system was simulated in a cubic box whose side was adjusted to yield a sample concentration of 5 mg/ml. Langevin dynamics with a friction value of 0.15 ps^{-1} was used. This friction coefficient is much smaller than that of water (43 ps^{-1} at 330 K) to allow for sufficient sampling within the microsecond time-scale of the simulation. The small friction does not influence the thermodynamic properties of the system. The SHAKE algorithm[84] was used to fix the length of the covalent bonds involving hydrogen atoms, which allows an integration time-step of 2 fs. Furthermore, the non-bonded interactions were updated every ten dynamics steps and coordinate frames were saved every 20 ps for a total of 5×10^4 conformations/μs. On a 2.1 GHz Athlon processor, a 1 μs run requires approximately 10.4 days, 25.2 days and 22.3 days for three, seven-residue peptides, six, seven-residue peptides, and three 11-residue peptides, respectively. Simulations were run on a Beowulf cluster for a total simulation time of 0.35 ms, 0.02 ms and 0.05 ms for $A\beta_{42}$, $hIAPP_{1–37}$ and $Ure2p_{20–70}$, respectively.

Explicit solvent simulations started from ordered aggregates

For a subset of four $A\beta_{42}$ seven-residue stretches ($A\beta_{1–7}$, $A\beta_{5–11}$, $A\beta_{13–19}$ and $A\beta_{29–35}$), explicit solvent MD simulations were carried out starting from in-register parallel β-sheet conformations. The starting structure of each run was selected among the implicit solvent conformations with a fraction of parallel contacts larger than 0.85 (these contacts were defined following the procedure described by Gsponer *et al.*[16]). Explicit solvent simulations were performed with the program NAMD2[85] using the CHARMM all-hydrogen force field (PARAM22 potential function[86]) along with the TIP3P model for water molecules.[87] Non-polar hydrogen atoms were added by CHARMM (HBUILD module) and their position was optimized *in vacuo*. The resulting structure was solvated in a water box of appropriate dimensions so that the distance between periodic images was not smaller than 25 Å. Chloride and sodium ions were added to neutralize the systems, yielding a salt concentration of 150 mM. Long-range

electrostatic forces were accounted for by using the particle mesh Ewald summation method[88] with real space cutoff distance of 10 Å and a grid width of 0.93 Å. The simulations were run at constant temperature (310 K) and pressure (1 atm) by applying the Berendsen thermostat[89] with a coupling decay time of 1 ps and the hybrid Nose–Hoover Langevin pressure control.[90] The SHAKE algorithm[84] was used to allow for an integration time-step of 2 fs. Solvent molecules and counterions were equilibrated at 310 K while holding the peptide system rigid for 1 ns. Two 0.5 ns equilibration cycles were then performed applying a harmonic constraint to all peptide atoms with force constants of 1.0 and 0.1 kcal/mol Å^2, respectively. Upon releasing the constraints, 50 ns production runs were performed for each peptide system.

Order parameters and β-aggregation propensity

The nematic order parameter $\bar{P}_2$ was considered to monitor the aggregation process as described by Cecchini *et al*.[48] This order parameter is widely used for studying the properties of anisotropic fluids such as liquid crystals[91–93] and is defined as:

$$\bar{P}_2 = \frac{1}{N}\sum_{i=1}^{N}\frac{3}{2}(\hat{\mathbf{z}}_i \cdot \hat{\mathbf{d}})^2 - \frac{1}{2} \quad (1)$$

where $\hat{\mathbf{d}}$ (the director) is a unit vector defining the preferred direction of alignment, $\hat{\mathbf{z}}_i$ is a suitably defined molecular vector, and N is the number of molecules in the simulation box, i.e., three or six peptides in this study. The director is defined as the eigenvector of the ordering matrix[94] that corresponds to the largest positive eigenvalue. Here, the molecular vectors $\hat{\mathbf{z}}_i$ were defined as unit vectors linking the peptide's termini (from the N to the C terminus). The nematic $\bar{P}_2$ describes the orientational order of the system and discriminates between ordered and disordered conformations.

As it has been recently shown,[48] the average over the canonical ensemble of $\bar{P}_2$ is descriptive of the thermodynamic stability of the ordered state of oligomeric peptide systems. This scalar value, which ranges from 0 (complete disorder) to 1 (perfect order), can then be used to estimate and compare the amyloidogenic propensity of different peptide sequences. Here, the value of $\bar{P}_2$ averaged over the MD trajectory, referred to as β-aggregation propensity, is used to determine the β-aggregation profiles.

Progress variables

Radius of gyration

The radius of gyration of the oligomeric system R_g was considered to monitor the "condensation" equilibrium along the simulation trajectories. Conformations of the system producing non-interacting peptides, namely conformations where all inter-peptide atomic distances are larger than the long-range interactions cutoffs (7.5 Å in this case), were used to determine R_g^C. In other words, R_g^C is the lowest value of the radius of gyration measured for the snapshots when the three peptides are far apart from each other. Conformations with one or more isolated peptides ($R_g > R_g^C$) describe the "uncondensed state" of the system.

Secondary structure

Strings of secondary structure (SS) were considered to monitor peptide conformational changes along the trajectory. For each oligomeric snapshot (Cartesian coordinates of the atomic nuclei) the SS of each chain was calculated.[95] The resulting strings of SS elements (one element per residue) were used to describe peptide conformations and monitor the aggregation process. The SS alphabet includes four possible letters: "H", "E", "T" and "–", which stand for α-helix, extended (β-strand), β-turn or bend, and random coil, respectively. Terminal caps as well as N and C-terminal residues are always assigned a "–". Albeit devoid of the atomic detail, SS strings are useful because they provide an intuitive description of the shape of the peptide backbone. However, these strings are suitable to monitor peptide conformational changes only at a coarse-grained level. In fact, in a perfect in-register arrangement a capped hendecapeptide would present the following SS string "–EEEEEEEEE–" independently of the polarity of the assembly, i.e. the number of parallel and antiparallel β-strands.

Mutagenesis

The N-terminal domain of Ure2p (Ure2p$_{1–94}$) is highly insoluble and forms inclusion bodies in *E. coli*. In contrast, it is soluble when attached to the C-terminal domain of the protein. We therefore engineered a specific cleavage site between the two domains in order to generate soluble Ure2p$_{1–94}$ at the onset of the assembly reaction by cleavage with the specific protease Factor Xa as described by Bousset *et al.*[96] The variant Ure2p-N4748S expression vector was obtained by site-directed mutagenesis. Mutagenesis was achieved using the QuickChange site-directed mutagenesis kit (Stratagene Europe, Amsterdam, The Netherlands) and the primers 5′-CAGGTGTAAATAATAATAGTAGTAA-CAATAGCAGTAGTAATAAC-3′ and 5′-GTTATTAC-TACTGCTATTGTTACTACTATTATTATTTACACCTG-3′.

Protein purification, generation of soluble Ure2p$_{1–94}$ and assembly of Ure2p$_{1–94}$ into fibrils

Recombinant Ure2p-I91EGR94 and Ure2p-N4748S-I91EGR94 were over expressed as soluble proteins in *E. coli* and purified as described.[46] The Ure2p$_{1–94}$ fragment was generated as described by Bousset *et al.*[96] The assembly reactions were monitored using thioflavin T binding ,[97] using a Quantamaster QM 2000-4 spectrofluorimeter (Photon Technology International, Inc. NJ). Ure2p$_{1–94}$ fibrils were also examined following negative staining with 1% uranyl acetate on carbon-coated grids (200-mesh) in a Philips EM 410 electron microscope (Philips Inc., The Netherlands).

Acknowledgements

We thank R. Pellarin for introducing periodic boundary conditions in the SASA module in CHARMM (version 29) and S. Dubois for excellent technical assistance. We are grateful to E. Guarnera, F. Rao and G. G. Tartaglia for helpful discussions, and to Professor F. E. Cohen for suggesting the control runs with explicit solvent. We thank

A. Widmer (Novartis Pharma, Basel) for providing the molecular modeling program Wit!P, which was used for visual analysis of the trajectories. We are very grateful to Dr M. Seeber (University of Modena and Reggio Emilia) for his computer program for the efficient analysis of CHARMM trajectories and to Dr M. Schaefer (Sygenta, Basel, Switzerland) for providing the program used for the clustering with RMSD between all pairs of structures. The simulations were performed on the Matterhorn Beowulf cluster at the Computing Center of the University of Zurich. We thank C. Bolliger and Dr A. Godknecht for setting up the cluster and the Canton of Zurich for generous hardware support. This work was supported by grants from the Swiss National Competence Center in Neural Plasticity and Repair (NCCR) and the Swiss National Science Foundation to AC and the French Ministry of Education and Research through the GIS Prions grant to RM.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2006.01.009

References

1. Thomas, P., Qu, B. & Pedersen, P. (1995). Defective protein folding as a basis of human disease. *Trends Biochem. Sci.* **20**, 456–459.
2. Dobson, C. M. (2001). The structural basis of protein folding and its links with human disease. *Phil. Trans. Roy. Soc. Ser. B*, **356**, 133–145.
3. Horwich, A. (2002). Protein aggregation in disease: a role for folding intermediates forming specific multimeric interactions. *J. Clin. Invest.* **110**, 1221–1232.
4. Westermark, P., Benson, M., Buxbaum, J., Cohen, A., Frangione, B., Ikeda, S. *et al*. (2002). Amyloid fibril protein nomenclature, *Amyloid*, **9**, 197–200.
5. Dobson, C. M. (1999). Protein misfolding, evolution and disease. *Trends Biochem. Sci.* **24**, 329–332.
6. Perutz, M. F. (1999). Glutamine repeats and neurodegenerative diseases: molecular aspects. *Trends Biochem. Sci.* **24**, 58–63.
7. Blake, C. & Serpell, L. (1996). Synchrotron X-ray studies suggest that the core of the transthyretin amyloid fibril is a continuous β-sheet helix. *Structure*, **4**, 989–998.
8. Malinchik, S. B., Inouye, H., Szumowski, K. E. & Kirschner, D. A. (1998). Structural analysis of Alzheimer's β_{1-40} amyloid: protofilament assembly of tubular fibrils. *Biophys. J.* **74**, 537–545.
9. Sunde, M. & Blake, C. C. F. (1997). The structure of amyloid fibrils by electron microscopy and X-ray diffraction. *Advan. Protein Chem.* **50**, 123–159.
10. Guijarro, J., Sunde, M., Jones, J., Campbell, I. & Dobson, C. (1998). Amyloid fibril formation by an SH3 domain. *Proc. Natl Acad. Sci. USA.* **95**, 4224–4228.
11. Konno, T., Murata, K. & Nagayama, K. (1998). Amyloid-like aggregates of a plant protein: a case of a sweet-tasting protein, monellin. *FEBS Letters*, **95**, 4224–4228.
12. Fandrich, M., Fletcher, M. & Dobson, C. (2001). Amyloid fibrils from muscle myoglobin. *Nature*, **410**, 165–166.
13. Dobson, C. M. (2003). Protein folding and misfolding. *Nature*, **426**, 884–890.
14. Gazit, E. (2002). A possible role for π-stacking in the self-assembly of amyloid fibrils. *FASEB J.* **16**, 77–83.
15. Tartaglia, G. G., Pellarin, R., Cavalli, A. & Caflisch, A. (2004). The role of aromaticity, exposed surface and dipole moment in determining protein aggregation rates. *Protein Sci.* **13**, 1939–1941.
16. Gsponer, J., Habertür, U. & Caflisch, A. (2003). 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. USA*, **100**, 5154–5159.
17. Lopez de la Paz, M. & Serrano, L. (2004). Sequence determinants of amyloid fibril formation. *Proc. Natl Acad. Sci. USA.* **101**, 87–92.
18. Nelson, R., Sawaya, M., Balbirnie, M., Madsen, A., Riekel, C., Grothe, R. & Eisenberg, D. (2005). Structure of the cross-β spine of amyloid-like fibrils. *Nature*, **435**, 773–778.
19. Griffiths, J., Ashburn, T., Auger, M., Costa, P., Griffin, R. & Lansbury, P. (1995). Rotational resonance solid-state NMR elucidates a structural model of pancreatic amyloid. *J. Am. Chem. Soc.* **117**, 3539–3546.
20. Burkoth, T., Benzinger, T., Urban, V., Morgan, D., Gregory, D., Thiyagarajan, P. *et al*. (2000). Structure of the β-amyloid (10–35). *J. Am. Chem. Soc.* **122**, 7883–7889.
21. Benzinger, T., Gregory, D., Burkoth, T., Miller-Auer, H., Lynn, D., Botto, R. & Meredith, S. (2000). Two-dimensional structure of β-amyloid (10–35) fibrils. *Biochemistry*, **39**, 3491–3499.
22. Antzutkin, O., Leapman, R., Balbach, J. & Tycko, R. (2002). Supramolecular structural constraints on Alzheimer's β-amyloid fibrils from electron microscopy and solid-state nuclear magnetic resonance. *Biochemistry*, **41**, 15436–15450.
23. Petkova, A. T., Ishii, Y., Balbach, J. J., Antzutkin, O. N., Leapman, R. D., Delaglio, F. & Tycko, R. (2002). A structural model for Alzheimer's β-amyloid fibrils based on experimental constraints from solid state NMR. *Proc. Natl Acad. Sci. USA.* **99**, 16742–16747.
24. Jaroniec, C., MacPhee, C., Bajaj, V., McMahon, M., Dobson, C. & Griffin, R. (2004). High-resolution molecular structure of a peptide in an amyloid fibril determined by magic angle spinning NMR spectroscopy. *Proc. Natl Acad. Sci. USA.* **101**, 711–716.
25. Serag, A., Altenbach, C., Gingery, M., Hubbell, W. & Yeates, T. (2001). Identification of a subunit interface in transthyretin amyloid fibrils: evidence for self-assembly from oligomeric building blocks. *Biochemistry*, **40**, 9089–9096.
26. Torok, M., Milton, S., Kayed, R., Wu, P., McIntire, T., Glabe, C. & Langen, R. (2002). Structural and dynamic features of Alzheimer's Aβ peptide in amyloid fibrils studied by site-directed spin labeling. *J. Biol. Chem.* **277**, 40810–40815.
27. Jayasinghe, S. & Langen, R. (2004). Identifying structural features of fibrillar islet amyloid polypep-

tide using site-directed spin labeling. *J. Biol. Chem.* **279**, 48420–48425.
28. Jimenez, J., Guijarro, J., Orlova, E., Zurdo, J., Dobson, C., Sunde, M. & Saibil, H. (1999). Cryo-electron microscopy structure of an SH3 amyloid fibril and model of the molecular packing. *EMBO J.* **18**, 815–821.
29. Kishimoto, A., Hasegawa, K., Suzuki, H., Taguchi, H., Namba, K. & Yoshida, M. (2004). β-helix is a likely core structure of yeast prion Sup35 amyloid fiber-scryo-electron microscopy structure of an SH3 amyloid. *Biochim. Biophys. Acta*, **315**, 739–745.
30. Williams, A. D., Portelius, E., Kheterpal, I., Guo, J., Cook, K. D., Xu, Y. & Wetzel, R. (2004). Mapping Aβ amyloid fibril secondary structure using scanning proline mutagenesis. *J. Mol. Biol.* **335**, 833–842.
31. Dubay, K. F., Pawar, A. P., Chiti, F., Zurdo, J., Dobson, C. M. & Vendruscolo, M. (2004). Prediction of the absolute aggregation rates of amyloidogenic polypeptide chains. *J. Mol. Biol.* **341**, 1317–1326.
32. Fernandez-Escamilla, A. M., Rousseau, F., Schymkowitz, J. & Serrano, L. (2004). Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nature Biotechnol.* **22**, 1302–1306.
33. Tartaglia, G. G., Pellarin, R., Cavalli, A. & Caflisch, A. (2005). Prediction of aggregation rate and aggregation-prone segments in polypeptide chains. *Protein Sci.* **14**, 2723–2734.
34. Lopez de la Paz, M., de Mori, G., Serrano, L. & Colombo, G. (2005). Sequence dependence of amyloid fibril formation: insights from molecular dynamics simulations. *J. Mol. Biol.* **349**, 583–596.
35. Buchete, N., Tycko, R. & Hummer, G. (2005). Molecular dynamics simulations of Alzheimer's β-amyloid protofilaments. *J. Mol. Biol.* **353**, 804–821.
36. Ferrara, P., Apostolakis, J. & Caflisch, A. (2002). Evaluation of a fast implicit solvent model for molecular dynamics simulations. *Proteins: Struct. Funct. Genet.* **46**, 24–33.
37. Selkoe, D. J. (1999). Translating cell biology into therapeutic advances in Alzheimer's disease. *Nature*, **399**, A23–A31.
38. Citron, M. (2004). β-Secretase inhibition for the treatment of Alzheimer's disease: promise and challenge. *Trends Pharmacol. Sci.* **25**, 92–97.
39. Westermark, P. & Wilander, E. (1978). Influence of amyloid deposits on islet volume in maturity onset diabetes-mellitus. *Diabetologia*, **15**, 417–421.
40. King, H., Aubert, R. & Herman, W. (1998). Global burden of diabetes, 1995–2025–prevalence, numerical estimates, and projections. *Diabetes Care*, **21**, 1414–1431.
41. Makin, S. & Serpell, L. C. (2004). Structural characterization of islet amyloid polypeptide fibrils. *J. Mol. Biol.* **335**, 1279–1288.
42. Wickner, R. (1994). [URE3] as an altered Ure2 protein–evidence for a prion analog in *Saccharomyces cerevisiae*. *Science*, **264**, 566–569.
43. Courchesne, W. & Magasanik, B. (1988). Regulation of nitrogen assimilation in *Saccharomyces cerevisiae*–roles of the ure2 and gln3 genes. *J. Bacteriol.* **170**, 708–713.
44. Lacroute, F. (1971). Non-mendelian mutation allowing ureidosuccinic acid uptake in yeast. *J. Bacteriol.* **106**, 519–522.
45. Masison, D., Maddelein, M. & Wickner, R. (1997). The prion model for [URE3] of yeast: spontaneous generation and requirements for propagation. Prion-inducing domain of yeast Ure2p and protease. *Proc. Natl Acad. Sci. USA.* **94**, 12503–12508.
46. Thual, C., Komar, A., Bousset, L., Fernandez-Bellot, E., Cullin, C. & Melki, R. (1999). Structural characterization of *Saccharomyces cerevisiae* prion-like protein Ure2. *J. Biol. Chem.* **274**, 13666–13674.
47. Taylor, K., Cheng, N., Williams, R., Steven, A. & Wickner, R. (1999). Prion domain initiation of amyloid formation *in vitro* from native Ure2p. *Science*, **283**, 1339–1343.
48. Cecchini, M., Rao, F., Seeber, M. & Caflisch, A. (2004). Replica exchange molecular dynamics simulations of amyloid peptide aggregation. *J. Chem. Phys.* **121**, 10748–10756.
49. Tjernberg, L. O., Näslund, J., Lindqvist, F., Johansson, J., Karlstrom, A. R., Thyberg, J. *et al*. (1996). Arrest of β-amyloid fibril formation by a pentapeptide ligand. *J. Biol. Chem.* **271**, 8545–8548.
50. Riek, R., Güntert, P., Döbeli, H., Wipf, B. & Wüthrich, K. (2001). NMR studies in aqueous solution fail to identify significant conformational differences between the monomeric forms of two Alzheimer peptides with widely different plaque-competence, $A\beta(1–40)^{ox}$ and $A\beta(1–42)^{ox}$. *Eur. J. Biochem.* **268**, 5930–5936.
51. Wurth, C., Guimard, N. & Hecht, M. (2002). Mutations that reduce aggregation of the Alzheimer's $A\beta_{42}$ peptide: an unbiased search for the sequence determinants of Aβ amyloidogenesis. *J. Mol. Biol.* **319**, 1279–1290.
52. Nilsberth, C., Westlind-Danielsson, A., Eckman, C., Condron, M., Axelman, K., Forsell, C. *et al*. (2001). The "Arctic" APP mutation (E693G) causes Alzheimer's disease by enhanced Aβ protofibril formation. *Nature Neurosci.* **4**, 887–893.
53. Päiviö, A., Jarvet, J., Gräslund, A., Lannfelt, L. & Westlind-Danielsson, A. (2004). Unique physicochemical profile of β-amyloid peptide variant $A\beta_{1–40}$ E22G protofibrils: conceivable neuropathogen in arctic mutant carriers. *J. Mol. Biol.* **339**, 145–159.
54. Massi, F., Klimov, D., Thirumalai, D. & Straub, J. E. (2002). Charge states rather than propensity for β-structure determine enhanced fibrillogenesis in wild-type Alzheimer's β-amyloid peptide compared to E22Q Dutch mutant. *Protein Sci.* **11**, 1639–1647.
55. Mazor, Y., Gilead, S., Benhar, I. & Gazit, E. (2002). Identification and characterization of a novel molecular-recognition and self-assembly domain within the islet amyloid polypeptide. *J. Mol. Biol.* **322**, 1013–1024.
56. Tjernberg, L. O., Callaway, D. J. E., Tjernberg, A., Hahne, S., Lilliehöök, C., Terenius, L. *et al*. (1999). A molecular model of Alzheimer amyloid β-peptide fibril formation. *J. Biol. Chem.* **274**, 12619–12625.
57. Li, L., Darden, T., Bartolotti, L., Kominos, D. & Pedersen, L. (1999). An atomic model for the pleated β-sheet structure of Aβ amyloid protofilaments. *Biophys. J.* **76**, 2871–2878.
58. Benzinger, T., Gregory, D., Burkoth, T., Miller-Auer, H., Lynn, D., Botto, R. & Meredith, S. (1998). Propagating structure of Alzheimer's β-amyloid (10–35) is parallel β-sheet with residues in exact register. *Proc. Natl Acad. Sci. USA*, **95**, 13407–13412.
59. Antzutkin, J., Balbach, O. N., Leapman, R., Rizzo, N., Reed, J. & Tycko, R. (2000). Multiple quantum solid-state NMR indicates a parallel, not antiparallel, organization of β-sheets in Alzheimer's β-amyloid fibrils. *Proc. Natl. Acad. Sci. USA.* **97**, 13045–13050.
60. Balbach, J., Petkova, A., Oyler, N., Antzutkin, O., Gordon, D., Meredith, S. & Tycko, R. (2002). Supramolecular structure in full-length Alzheimer's

β-amyloid fibrils: evidence for a parallel β-sheet organization from solid-state nuclear magnetic resonance. *Biophys. J.* **83**, 1205–1216.
61. Paul, C. & Axelsen, P. (2005). β sheet structure in amyloid-β fibrils and vibrational dipolar coupling. *J. Am. Chem. Soc.* **127**, 5754–5755.
62. West, M., Wang, W., Patterson, J., Mancias, J., Beasley, J. & Hecht, M. (1999). *De novo* amyloid proteins from designed combinatorial libraries. *Proc. Natl Acad. Sci. USA*, **96**, 11211–11216.
63. Lopez de la Paz, M., Goldie, K., Zurdo, J., Lacroix, E., Dobson, C. M., Hoenger, A. & Serrano, L. (2002). *De novo* designed peptide-based amyloid fibrils. *Proc. Natl Acad. Sci. USA*, **99**, 16052–16057.
64. Hammarstrom, P., Jiang, X., Hurshman, A., Powers, E. & Kelly, J. (2002). Sequence-dependent denaturation energetics: a major determinant in amyloid disease diversity. *Proc. Natl Acad. Sci. USA*, **99**, 16427–16432.
65. Chiti, F., Stefani, M., Taddei, N., Ramponi, G. & Dobson, C. (2003). Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature*, **424**, 805–808.
66. Klimov, D. & Thirumalai, D. (2003). Dissecting the assembly of $A\beta_{16-22}$ amyloid peptides into antiparallel β sheets. *Structure*, **11**, 295–307.
67. Liu, R., McAllister, C., Lyubchenko, Y. & Sierks, M. R. (2003). Residues 17–20 and 30–35 of β-amyloid play crucial roles in aggregation. *J. Neurosci. Res.* **75**, 162–171.
68. Hou, L., Shao, H., Zhang, Y., Hua, L., Menon, N. K., Neuhaus, E. *et al*. (2004). Solution NMR studies of the Aβ(1–40) and Aβ(1–42) peptides established that Met35 oxidation state affects the mechanism of amyloid formation. *J. Am. Chem. Soc.* **126**, 1992–2005.
69. Tenidis, K., Waldner, M., Bernhagen, J., Fischle, W., Bergmann, M., Weber, M., Kapurniotu, A. *et al*. (2000). Identification of a penta- and hexapeptide of islet amyloid polypeptide (IAPP) with amyloidogenic and cytotoxic properties. *J. Mol. Biol.* **295**, 1055–1071.
70. Wu, C., Lei, H. & Duan, Y. (2005). The role of phe in the formation of well-ordered oligomers of amyloidogenic hexapeptide (NFGAIL) observed in molecular dynamics simulations with explicit solvent. *Biophys. J.* **88**, 2897–2906.
71. O'Nuallain, B., Williams, A., Westermark, P. & Wetzel, R. (2004). Seeding specificity in amyloid growth induced by heterologous fibrils. *J. Biol. Chem.* **279**, 17490–17499.
72. Kajava, A., Baxa, U., Wickner, R. & Steven, A. (2004). A model for Ure2p prion filaments and other amyloids: the parallel superpleated β-structure. *Proc. Natl Acad. Sci. USA*, **101**, 7885–7890.
73. Kajava, A., Ueli, A. & Steven, A. (2005). The parallel superpleated beta-structure as a model for amyloid fibrils of human amylin. *J. Mol. Biol.* **348**, 247–252.
74. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983). CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187–217.
75. Neria, E., Fischer, S. & Karplus, M. (1996). Simulation of activation free energies in molecular systems. *J. Chem. Phys.* **105**, 1902–1921.
76. Hiltpold, A., Ferrara, P., Gsponer, J. & Caflisch, A. (2000). Free energy surface of the helical peptide $Y(MEARA)^6$. *J. Phys. Chem. B*, **104**, 10080–10086.
77. Ferrara, P. & Caflisch, A. (2000). Folding simulations of a three-stranded antiparallel β-sheet peptide. *Proc. Natl Acad. Sci. USA*, **97**, 10780–10785.
78. Ferrara, P. & Caflisch, A. (2001). Native topology or specific interactions: what is more important for peptide folding? *J. Mol. Biol.* **306**, 837–850.
79. Gsponer, J. & Caflisch, A. (2001). Role of native topology investigated by multiple unfolding simulations of four SH3 domains. *J. Mol. Biol.* **309**, 285–298.
80. Gsponer, J. & Caflisch, A. (2002). Molecular dynamics simulations of protein folding from the transition state. *Proc. Natl Acad. Sci. USA*, **99**, 6719–6724.
81. Balbirnie, M., Grothe, R. & Eisenberg, D. (2001). An amyloid-forming peptide from the yeast prion Sup35 reveals a dehydrated β-sheet structure for amyloid. *Proc. Natl Acad. Sci. USA*, **98**, 2375–2380.
82. Perutz, M. F., Pope, B. J., Owen, D., Wanker, E. E. & Scherzinger, E. (2002). Aggregation of proteins with expanded glutamine and alanine repeats of the glutamine-rich and asparagine-rich domains of Sup35 and of the amyloid β-peptide of amyloid plaques. *Proc. Natl Acad. Sci. USA*, **99**, 5596–5600.
83. Sugita, Y. & Okamoto, Y. (1999). Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Letters*, **314**, 141–151.
84. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. (1977). Numerical integration of the Cartesian equation of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comp. Phys.* **23**, 327–341.
85. Kalè, L., Skeel, R., Bhandarkar, M., Brunner, R., Gursoy, A., Krawetz, N. *et al*. (1999). Namd2: greater scalability for parallel molecular dynamics. *J. Comput. Phys.* **151**, 283–312.
86. MacKerell, A., Jr, Bashford, D., Bellott, M., Dunbrack, R., Jr, Evanseck, J., Field, M. *et al*. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*, **102**, 3586–3616.
87. Jorgensen, W. L., Chandrasekhar, J., Madura, J., Impey, R. W. & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935.
88. Darden, T., York, D. & Pedersen, L. (1993). Particle mesh Ewald–an $N\log(N)$ method for ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092.
89. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690.
90. Hoover, W. (1985). Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A*, **31**, 1695–1697.
91. Chandrasekhar, S. (1992). *Liquid Crystals*, Cambridge University Press, Cambridge, England.
92. de Gennes, P. G. & Prost, J. (1993). *The Physics of Liquid Crystals* 2nd edit., Oxford University Press, Oxford.
93. Zannoni, C. (2001). Molecular design and computer simulations of novel mesophases. *J. Mater. Chem.* **11**, 2637–2646.
94. Allen, M. P. & Tildesley, D. J. (1987). *Computer Simulation of Liquids*, Oxford Science Publications, Oxford, UK.
95. Andersen, C. A. F., Palmer, A. G., Brunak, S. & Rost, B. (2002). Continuum secondary structure captures protein flexibility. *Structure*, **10**, 174–184.

96. Bousset, L., Redeker, V., Decottignies, P., Dubois, S., LeMarechal, P. & Melki, R. (2004). Structural characterization of the fibrillar form of the yeast *Saccharomyces cerevisiae* prion Ure2p. *Biochemistry*, **43**, 5022–5032.

97. McParland, V., Kad, N., Kalverda, A., Brown, A., Kirwin-Jones, P., Hunter, M. *et al*. (2000). Partially unfolded states of β_2-microglobulin and amyloid formation *in vitro*. *Biochemistry*, **39**, 8735–8746.

Edited by F. E. Cohen

(Received 11 July 2005; received in revised form 21 November 2005; accepted 4 January 2006)
Available online 26 January 2006