# Emerging ideas on the molecular basis of protein and peptide aggregation

D Thirumalai*, DK Klimov and RI Dima

Several neurodegenerative diseases are associated with the unfolding and subsequent fibrillization of proteins. Although neither the assembly mechanism nor the atomic structures of the amyloid fibrils are known, recent experimental and computational studies suggest that a few general principles that govern protein aggregation may exist. Analysis of the results of several important recent studies has led to a set of tentative ideas concerning the oligomerization of proteins and peptides. General rules have been described that may be useful in predicting regions of known proteins (prions and transthyretin) that are susceptible to fluctuations, which give rise to structures that can aggregate by the nucleation-growth mechanism. Despite large variations in the sequence-dependent polymerization kinetics of several structurally unrelated proteins, there appear to be only a few plausible scenarios for protein and peptide aggregation.

**Addresses**
Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742, USA
*e-mail: thirum@glue.umd.edu

**Current Opinion in Structural Biology** 2003, **13**:XXX–XXX

This review comes from a themed issue on
Theory and simulation
Edited by Charles L Brooks III and David A Case

0959-440X/03/$ – see front matter


**Abbreviations**

| | |
|---|---|
| **ASA** | accessible surface area |
| **CD** | circular dichroism |
| **EM** | electron microscopy |
| **EPR** | electron paramagnetic resonance |
| **FTIR** | Fourier transform IR |
| **HB** | hydrogen bond |
| **MD** | molecular dynamics |
| **NCC** | nucleated conformational conversion |
| **OR²** | Oliveberg-Richardson-Richardson |
| **PDB** | Protein Data Bank |
| **PHF** | paired helical filament |
| **PrP^C** | cellular form of prion protein |
| **PrP^Sc** | pathogenic scrapie form of prion protein |
| **R²** | Richardson and Richardson |
| **SSE** | secondary structure element |
| **TTR** | transthyretin |

## Introduction

A large number of neurodegenerative diseases, including Alzheimer's disease [1,2] and the transmissible prion disorders [3,4], are associated with amyloid fibrils. Historically, amyloid referred to the extracellular deposits that were thought to be the cause of various diseases [2,5]. Many recent studies have found that mobile oligomers, which are precursors to fibril formation, may themselves be neurotoxic [2,6,7]. Several experiments have further shown that any generic protein, under suitable conditions, can form ordered aggregates, with morphologies that closely resemble the amyloid fibrils [8,9]. The finding that any protein can aggregate at high enough protein concentration and under suitable external conditions (pH, salt concentration, temperature) is not surprising. It is interesting, however, that even proteins and peptides that are not associated with known diseases form fibrils with the cross β-patterns that are characteristic of amyloid fibrils [9,10]. A more surprising finding is that the oligomers that form early in the aggregation process of even non-disease-related proteins may be cytotoxic [11••]. The formation of morphologically similar aggregates by a variety of proteins unrelated in sequence or structure suggests that certain general principles may govern fibrillization [10,12–14]. The vastness of sequence space and the heterogeneity of environmentally dependent intermolecular interactions make deciphering the principles of protein aggregation difficult.

The polymerization of proteins and peptides raises several questions of biophysical interest. First, what are the early events in the oligomerization process? In particular, what is the nature of the structural fluctuations that trigger the association of polypeptide chains? Second, it has been established, beginning with the classic studies of gelation of deoxyhemoglobin S [15], that polymerization occurs by a nucleation and growth process [16]. Nevertheless, several questions remain unanswered. What are the structural characteristics of the critical nuclei? Is the formation of distinct strains [17–19] reflected in the nature of the critical nuclei? Third, can the sequence and/or the structural characteristics of monomers provide insights into the sites that harbor amyloidogenic tendencies? Fourth, what are the principles that natural proteins use in preventing aggregation under physiological conditions? Fifth, can the variations in the fibrillization rates of naturally occurring mutants (in prions, transthyretins [TTRs] and Aβ peptides) be related to the biophysical characteristics of the monomers?

It is beyond the scope of this review to describe our current understanding of all the questions posed above. Topics related to the first two questions have been described in recent excellent reviews [10,12,13]. The past few years have witnessed considerable progress on several fronts in the field of protein aggregation. Here, we outline a few of these, with the emphasis on biophysical aspects of fibrillization. Some of the highlights are:

1. Solid-state NMR studies [20,21••,22,23•], imaging of Aβ oligomers using atomic force microscopy (AFM) [24] and cryo-EM [25,26] have been used to obtain insights into the structural organization of amyloid fibrils. The determination of Aβ fibril structures has led to computational strategies [27•] that distinguish between different models of fibril structures.

2. Several reports, especially in the context of Alzheimer's disease, have shown that the soluble oligomers themselves, rather than the protease-resistant plaque, may be the cause of neurotoxicity [2,6,7,28]. This finding has made it critical to understand the kinetics of aggregation of protofibrils, which are the precursors to the fibrils.

3. A detailed study of the fibrillization of Aβ peptides and their congeners has shown that the formation of fibrils with β-sheet architecture must involve the transient population of α-helical structures [29••]. Molecular dynamics (MD) simulations of the oligomerization of $A\beta_{16-22}$ peptides [30••] further suggest that, for this class of peptides, the formation of helical structures may be an obligatory intermediate step.

4. Exploration of the sequence and structural requirements needed to prevent fibrillization has given insights into the plausible regions in the cellular isoforms of prion proteins ($PrP^c$) and Aβ peptides that may be implicated in the transition to the fibrillar form [31••,32••]. These computational studies have led to testable predictions that are beginning to be confirmed in experiments.

5. Systematic studies of natural β-sheet proteins have led to the identification of the potential mechanisms that block aggregation [33••]. The translation of these observations into a simple computational rule allows us to predict regions that may be implicated in the production of intermediates that can grow into fibrils.

The purpose of this review is to formulate tentative ideas on the molecular origins of aggregation by synthesizing these important developments. A survey of seemingly unrelated studies suggests that a few qualitative principles about protein aggregation can be proposed. It is also clear that there are several outstanding issues that can only be addressed using a combination of experimental, theoretical and computational techniques. The review concludes with a description of a few of these outstanding problems.

## Conformational fluctuations of monomers provide a limited glimpse into fibrillization

It is known that aggregation kinetics depends on the sequence and the precise external conditions. Truncation of the two C-terminal residues of the Aβ peptide, whose sequence using single-letter code for amino acids is DAEFRHDSG[10]YEVHHQKLVF[20]FAEDVGSNKG[30]AI IGLMVGGV[40]VIA, results in substantial differences in the timescale of plaque formation for the $A\beta_{1-40}$ and $A\beta_{1-42}$ peptides [29••]. It has also been shown that E22Q ('Dutch') Aβ peptide has enhanced activity (as measured by peptide deposition rates) relative to the wild-type peptide for both the full-length (1–40) peptide and truncated (10–35) variants [34,35]. Similarly, aggregation times, under similar external conditions, vary greatly for wild-type TTR and its naturally occurring mutants [36••]. The difference in amyloidogenic characteristics is observed both in the rate of deposition of monomers onto existing fibrils and in the kinetics of oligomerization of peptides.

The observation of sequence-dependent deposition rates for Aβ peptides has been used to hypothesize that variations in the rates of amyloidogenesis may be explained by the propensity of different monomer sequences to form local structure [37,38]. To assess the validity of this hypothesis, Straub and co-workers [39,40,41••] have carried out a series of MD simulations of the wild-type and Dutch mutant of $A\beta_{10-35}$ peptide. Surprisingly, the analysis of multiple 1 ns MD simulations demonstrated that both peptides have very similar conformational properties. There is no appreciable difference in the β-structure propensities of the two peptides. The results of these studies imply that the structural characteristics of monomeric peptides may not be indicative of their amyloidogenic competence [39,40,41••]. Because, in general, the profound conformational changes are driven by interpeptide interactions, it is unlikely that the conformational dynamics of isolated peptides can fully explain variations in deposition rates.

We have proposed two alternative explanations for the change in the rate of amyloid formation between the wild type and E22Q mutant [41••]. Deletion of the charged residue (glutamic acid) is expected to compromise the solvation of E22Q peptide in water, which in turn leads to a reduction of the free energy barrier for fibril formation. It is also conceivable that the charged state of glutamic acid introduces destabilizing electrostatic interactions in the fibril itself. Therefore, the substitution E→Q may decrease the free energy barrier for forming assembly-competent structures.

The lack of correlation between the monomeric preferences of Aβ peptides and their observed propensities to form amyloid finds support in recent experimental studies. Wuthrich and co-workers [42•] investigated the conformational characteristics of $A\beta^{ox}_{1-40}$ and $A\beta^{ox}_{1-42}$ (*ox* means that methionine at position 35 occurs as sulfoxide) using solution NMR spectroscopy and found that there is close similarity between the solution structures of these peptides. The only discernible difference is found in the C-terminal region, starting with position 32. This is not surprising given that there are two additional (isoleucine and alanine) hydrophobic residues in the $A\beta^{ox}_{1-42}$ peptide. This finding is important, because $A\beta_{1-42}$ is known to fibrillize faster than $A\beta_{1-40}$ [43]. The study of Wuthrich and co-workers suggests that the two C-terminal

hydrophobic residues in $A\beta_{1-42}$ should critically affect the intermediate structures (oligomers) in fibril formation by lowering the free energy barrier for aggregation. It is also interesting that, similar to $A\beta_{10-35}$ [44], the structures of $A\beta^{ox}_{1-40}$ and $A\beta^{ox}_{1-42}$ peptides in aqueous solution have little long-range order or easily identifiable elements of secondary structure [42•]. The most rigid element of their structure seems to be the central hydrophobic cluster (17–21), which adopts similar conformations in both $A\beta^{ox}_{1-40}$ and $A\beta^{ox}_{1-42}$ peptides, as well as in $A\beta_{10-35}$ peptide. Thus, interpeptide interactions must be taken into account to understand the observed differences in the rate of amyloid formation between $A\beta_{1-40}$ and $A\beta_{1-42}$.

Recently, two 10 ns trajectories generated by MD simulations of $PrP^c$ (Figure 1a) have been used to probe the initial events in the conformational transition to the aberrant aggregation-prone form [45•]. It is known that this transition can be driven by lowering the pH (i.e. under acidic conditions) [46,47]. At neutral pH, the ordered regions of $PrP^c$ remain stable during the simulation time. However, at low pH, substantial conformational fluctuations in residues 109–175, which include disordered N-terminal helix 1 and the two small β strands (Figure 1a), are observed. The authors conclude from examining several conformational snapshots that there might be a tendency for strand formation in helix 1. Moreover, the strands in $PrP^c$ have a tendency to lengthen. These simulations suggest that a glimpse into the early events of the fibrillization kinetics may be obtained using MD simulations over a range of external conditions. Bioinformatic analysis [31••,32••] and recent experiments [48••] suggest that parts of helices 2 and 3 may also be implicated in the transition from $PrP^c$ to the scrapie form ($PrP^{Sc}$ see below).

## Negative design: gatekeeper residues prevent aggregation

In the cell, a large fraction of proteins with varying architecture fold spontaneously by avoiding off-pathway processes that lead to aggregation. To prevent aberrant protein aggregation, nature employs molecular chaperones — nanomachines that actively assist the folding of proteins. A plausible link between the underexpression of molecular chaperones and the onset of certain classes of diseases suggests that these nanomachines may be utilized more widely than has been appreciated so far. However, it has been estimated that, in *Escherichia coli*, only about 5–10% of all proteins can afford to employ molecular chaperones to enable them to reach the folded state [49]. Thus, as envisioned by Anfinsen [50], most proteins must fold spontaneously and efficiently into the native state.

Anfinsen's hypothesis has led to the quest to understand how a polypeptide chain navigates the rough energy landscape to reach the native state. The past decade has seen numerous theoretical and experimental advances in our understanding of how a monomeric protein folds. However, from the perspective of aggregation, it is crucial to understand how proteins, under physiological conditions, avoid aggregation. A suggestion is that spontaneously folding proteins may have utilized negative design in generating sequences that not only can reach the final desired structure efficiently but also can avoid unproductive pathways [33••]. An important question that arises in the context of aggregation is: are there residues (gatekeepers in the terminology of Otzen and Oliveberg [51]) that implement the 'negative design' principle? Otzen *et al.* [52••] have suggested, based on re-engineering the β strand in ribosomal S6 protein from *Thermus thermophilus* (Figure 2) to have a sequence composition similar to that of Aβ peptide, that the gatekeepers, which preserve the structural integrity of the wild-type protein, are charged residues. They modified the β2 strand in S6 by replacing charged residues with hydrophobic residues. The resulting S6-Alz mutant, in which the six charged residues are replaced by hydrophobic species, forms a tetramer. Based on this finding, they proposed that the charged gatekeeper residues, which are not implicated in monomer folding, block aggregation by an electrostatic mechanism. The formation of the interfaces needed for oligomerization is prevented in wild-type S6 by electrostatic repulsion, but is promoted in S6-Alz by favorable interactions between hydrophobic residues.

A systematic bioinformatic approach has recently been used to identify potential gatekeeper residues in β strands [33••]. Motivated in part by the question posed above and by the finding that the majority of *de novo* designed all-β-sheet proteins tend to oligomerize, Richardson and Richardson ($R^2$) [33••] have proposed a set of rules for identifying aggregation-blocking mechanisms in β-sheet proteins. They note that, unless the edge strands utilize 'negative design', edge-to-edge aggregation can easily occur in all-β-sheet proteins. To understand how natural proteins avoid this unproductive route, they carried out an analysis of the architecture and sequence of the edge strands of β-sheet proteins. Their study reveals that there are two global 'blocking' strategies that nature utilizes to prevent edge-to-edge aggregation.

### Minimization and/or protection of dangling hydrogen bonds

One of the principles that emerges from the $R^2$ arguments [33••] is that, in the folded states of naturally occurring proteins, the number of dangling hydrogen bonds (HBs) is minimized. Conversely, the presence of a large number of dangling HBs promotes intermolecular association. The universal interaction that stabilizes β-sheet proteins is the formation of HBs. Proteins with β-barrel architecture have very few unsatisfied HBs. As a result, there are literally no edges in their structures. In β helices, which have been

suggested to be the nearly 'universal' structure of amyloid fibrils [53••], the edges are protected by large loops. Other β-sheet architectures, such as β propellers and single β-sheet proteins, use a combination of β bulges and charges to avoid aggregation.

A corollary of the $R^2$ findings is that low-stability β strands with a large number of unsatisfied HBs may be susceptible to aggregation. In $PrP^c$, it is likely that frustrated helices 2 and 3 could, upon conformational change, have strand conformation (see below). The percentages of unsatisfied HBs are 14, 14 and 9 in mouse $PrP^c$, Syrian hamster $PrP^c$ and $h1PrP^c$, respectively [32••]. These are larger than the average fraction (6% [54]) of residues in normal proteins that have unsatisfied buried HB donors/acceptors. The extended structure of helices 2 and 3 in $PrP^c$, together with the large number of unsatisfied HBs, makes this region susceptible to edge-to-edge aggregation.

**Inward-pointing charged residues block aggregation**

In the context of amyloid fibrils, β-sandwich and single β-sheet proteins are of particular interest. This is because disease-related proteins usually polymerize (see, however, [55••]) upon fibril formation into β-sandwich structures [2]. Edge-to-edge aggregation in naturally occurring β-sandwich proteins is prevented by placing an 'inward-pointing' charged residue on the hydrophobic side of a β strand [33••]. For a pair of β strands, a charged sidechain is 'inward pointing' if its $C_\alpha$–$C_\beta$ vector points towards the other strand backbone. Placement of just one such residue in the edge strand results in a minimal change in the stability of a protein, but prevents aggregation. The placement of a charged residue prevents aggregation either because of interstrand electrostatic repulsion, as envisioned by Otzen *et al.* [52••], or by the need to expose the charged residues to solvent. In the latter case, the distance between the β strands would be large enough to prevent the formation of HBs. Thus, any charged residue (+ or –) can be inward pointing provided the sidechain is long enough. We will refer to the principle underlying this blocking strategy as the $OR^2$ (Oliveberg-Richardson-Richardson) rule.

The other strategy, which is not as relevant to aggregation, involves creating a local β bulge, which effectively disrupts HBs between β strands [33••]. There is no sequence conservation at gatekeeper positions as might be deemed necessary for monomeric folding. The irregularities found in edge strands due to the placement of 'unusual' residues are purely for the purpose of negative design [33••].

There are a few experimental amyloidogenesis studies that illustrate the $OR^2$ criterion for preventing aggregation.

*S6 and variants*

Otzen *et al.* [52••] probed fibril formation in three 14-mer peptides corresponding to residues 36–49 in the β2 strand of S6. The wild-type RVEKVEELGLRRLA peptide, which has a net positive charge, has seven charged residues. Both wild-type peptide and the double mutant E41A/E42A are soluble. Otzen *et al.* noted that E41A/E42A forms amorphous (gel-like) aggregates at high peptide concentration. This observation points to the need for exploring phase diagrams of proteins with the protein concentration and other external conditions as appropriate variables [56••].

At a relatively low protein concentration, the mutant peptide S6-Alz (RVEKVAILGLMVLA) forms insoluble fibrils with morphology similar to that of Aβ aggregates. The S6-Alz peptide, which has β-sheet structure in water, has no charged residues in the middle. Because the aggregation-blocking mechanism is disabled, the $OR^2$ rule implies that S6-Alz would form β sandwiches stabilized by interpeptide interactions between hydrophobic sidechains. The middle of S6-Alz has the membrane protein motif HHHHGHHHHH (H stands for hydrophobic residue), which occurs with negligible probability in globular proteins.

The $OR^2$ rule can also be used in interpreting the tetramerization of the S6-Alz mutant. At high concentration, S6-Alz forms tetramers, in which edge strand β2 serves as an interface. Residues 38–44 in strand β2 of one of the molecules form an antiparallel β sheet with the same residues from an another molecule. Similarly, residues 47–50 of β2 form an intermolecular antiparallel β sheet with the β strand (residues 89–92) from another molecule. Aggregation of S6-Alz into tetramers becomes possible due to double mutations E41A/E42A and R46M/R47M, which remove charged residues from strand β2. This experimental result may be rationalized in light of the $OR^2$ rule. The negatively charged sidechains of glutamic acid residues, which appear in tandem at positions 41 and 42, are placed on both sides of β2 (Figure 2). Their sidechains are exposed to solvent (the relative, that is, with respect to a Gly-X-Gly construct, accessible surface areas [ASAs] are 0.46 and 0.44, respectively). The positively charged sidechains of the two arginine residues, which occur in tandem at positions 46 and 47, are also placed on opposite sides of the β strand (Figure 2). Furthermore, a twist in the β strand is observed next to R46 and R47. According to the $R^2$ rule [33••], these are the typical mechanisms that prevent edge-to-edge aggregation in single β-sheet proteins, such as S6.

The blocking method found in wild-type S6 is by no means unique. Similar aggregation-preventing mechanisms are observed in several other single β-sheet proteins. For example, the edge β-strand 4 of profilin (PDB code 1pne) contains two sequential charged residues, R74 and D75, whose ASAs are 0.59 and 0.26, respectively. In addition, several noticeable twists are observed in this edge β strand, in particular, near the positively charged K69. The same mechanisms seem to be operative in preventing aggregation in the edge β-strand 5 of chain A of monellin (PDB code 1mol). Two sequential charged residues (R82 and K83, with ASAs

of 0.32 and 0.61, respectively) are found at the beginning of this β strand. Thus, the mechanisms blocking aggregation in wild-type S6 fall in the categories described by $R^2$ [33••]. As observed by Otzen *et al.* [52••], mutating these naturally evolved structural gatekeepers in S6 should lead to tetramerization.

A direct test of the $OR^2$] rule for preventing aggregation was provided by Wang and Hecht [57•]. Combinatorially *de novo* designed β-sheet proteins built of seven-residue β strands with an alternating hydrophobic/polar (PHPHPH) pattern form fibrils with amyloid-like characteristics. The $OR^2$ rule would suggest that, if the middle hydrophobic residue in the edge strand is replaced with lysine (i.e. PHPKPHP), then the protein would be soluble. If such proteins form fibrils, electrostatic repulsion or/and burial of uncompensated charge would render the fibrils unstable. In accord with the $OR^2$ rule, Wang and Hecht showed that the redesigned proteins with lysine in the middle of an otherwise alternating hydrophobic/polar edge strand sequence do not aggregate and form monomeric β-sheet structures.

*Fibrillization of transthyretin*

When TTR is subject to denaturation stress, conformational fluctuations in the monomer produce a state that can form amyloid fibrils [13]. The aberrant aggregation of TTR is associated with spontaneous and familial diseases in humans. By following the electron paramagnetic resonance (EPR) spectra before and after fibril formation, Serag *et al.* [58••] have established the arrangement of the strands in the amyloid fibrils. TTR, predominantly a β-sheet protein, forms a tetramer by burying hydrophobic strand H (Figure 3) at the interface between the four identical monomeric units. Kelly and co-workers [36••] have demonstrated that fluctuations (induced in denaturing environments) populate a partially unfolded intermediate that is susceptible to fibrillization by a nucleation and growth process. Recent studies from the Yeates laboratory [58••] suggest that, in this state, the F, F′, B and B′ β strands become exposed. The resulting structure can assemble and propagate by head-to-head and tail-to-tail arrangements, giving rise to the polymeric construct $(BEFF'E'B')_n$. In this proposed arrangement, the native-like interface contacts between the F and F′ β strands are preserved (Figure 3).

It was noted by Yeates and co-workers [58••] that the proposed architecture of TTR fibrils is consistent with the Richardson studies. Here, we describe an analysis of the structural characteristics of the TTR dimer (PDB code 2pab), which provides additional support to the proposed architecture of wild-type TTR fibrils (Figure 3). If the dimer is dissected into its constituents, the highly hydrophobic strand H, with the largest ASA (in the monomer state), is the edge strand. The $R^2$ observation would suggest that this vulnerable strand will form a β sheet with other strands, which explains why TTR is a tetramer in the natural state. A similar analysis of the dimer suggests that the ASA of strand H is greater than that of strand B, making the former more susceptible to conformational fluctuations. Furthermore, examination of the dimer structure indicates that HH′ interactions constitute the most stable region in the monomer interface, which is unlikely to dissolve given that the FF′ interactions are retained in the fibril [58••]. The resulting amyloid fibrils would form an additional $(AGHH'G'A')_n$ construct, a possibility that was not ruled out by Yeates and co-workers [58••].

We have calculated, using the protocol described elsewhere [59], the energies required to expose strands B and H. The energy loss in forming the misfolded structure that enables the formation of the BB′ interface is considerably smaller than that associated with the disruption of the HH′ interface. Exposure of strand B requires the removal of edge strands C and D (Figure 3). On the other hand, exposure of strand H requires breaking the entire interface (HH′ and FF′), which is stabilized by several sidechain contacts and HBs. The bulk of the interfacial energy gain in the wild-type TTR arises from the strong interactions between H and H'. As a result, it is unlikely that the partially folded structure involves conformational changes in the interfacial region. The current computations show that the use of the $R^2$ observation, together with the stability arguments, helps us understand the architecture of TTR amyloid fibrils. Because of the presence of the consecutive like charges (arginine and lysine) towards the end of strand B in an otherwise hydrophobic environment, it is easy to predict that B and B′ should be arranged in an antiparallel fashion [58••].

Another line of evidence that implicates strand B is the observation that many disease-causing mutations are clustered in this region. Therefore, this region of the protein may be intrinsically susceptible to fluctuations under suitable denaturation stress. It also follows from our analysis that mutations that destabilize the interface might lead to fibrils with a different architecture. The converse of this has been demonstrated by Kelly and co-workers [36••]. They showed that the mutation T119A, which stabilizes the tetramers, essentially prevents fibrillization.

## 'Frustrated' secondary structural elements may be harbingers of a tendency to polymerize

The ease of aggregation and the morphology of the aggregates depend not only on protein concentration but also on other external conditions, such as temperature, pH and salt concentration. Although most proteins can, under suitable conditions, aggregate, the observation that several disease-causing proteins form amyloid fibrils under physiologically relevant conditions raises the question: is aggregation or the need to avoid unproductive pathways encoded in the primary sequence itself? It is clear that sequences that contain a patch of hydrophobic residues are prone to forming

aggregates [60]. However, it is known that contiguous patches (three or more hydrophobic residues) occur with low probability in globular proteins [61]. For example, sequences with five hydrophobic residues (LVFFA in Aβ peptide) in a row are not well represented. Similarly, it is unusual to find hydrophobic residues concentrated in a specific region of helices, such as in helix 2 of $PrP^c$ [32••]. *De novo* design of α helices or β strands based on periodic binary patterned (sequences formed from hydrophobic and polar residues only) sequences often forms insoluble oligomers [60]. The morphology of these oligomers apparently has the characteristics of amyloid fibrils. These examples suggest that sequence alone in some cases might reveal the tendency towards aggregation of proteins.

It is natural to wonder if secondary structure elements (SSEs) bear signatures that could reveal amyloidogenic tendencies. The incompatibility of the nature of an SSE in the context of the entire protein may give insights into regions of the protein that may be susceptible to conformational fluctuations. Two studies have proposed that the extent of 'frustration' in SSEs may be a harbinger of amyloid fibril formation [31••,32••]. Because reliable secondary structure prediction requires knowing the context-dependent propensities and multiple sequence alignments (as used in PHD, a profile network from Heidelberg [62]), it is likely that assessment of the extent of frustration in SSEs, rather than analysis of sequence patterns, is a better predictor of fibril formation. Frustration in SSEs is defined as the incompatibility of the predicted (from PHD, for example) secondary structure and the experimentally determined structure [31••]. For example, if a secondary structure is predicted with high confidence to be in a β strand and if that segment is found (by NMR or X-ray crystallography) to be in a helix, then the structure is 'frustrated' (or discordant or mismatched). The α/β discordance, which can be correlated with amyloid formation, can be assessed using the score $S_{\alpha/\beta} = \frac{1}{L}\sum_{i=1}^{L}(R_i - 5)$, where $R_i$ is the reliability score predicted by PHD at position $i$ of the query sequence, 5 is the mean score and $L$ is the sequence length. The bounds on $S_{\alpha/\beta}$ are $0 \le S_{\alpha/\beta} \le 4$, with maximal frustration corresponding to $S_{\alpha/\beta} = 4$. Similarly, the measure $S_{\beta/\alpha}$ gives the extent of frustration in a region that is predicted to be helical and is found experimentally to be a strand. Using $S_{\alpha/\beta}$ and other structural characteristics, one can make predictions of the plausible regions that are most susceptible to large conformational fluctuations.

## $PrP^c$ and Dpl

Using the above concept of SSE frustration, the 23-residue sequence QNNFVHDCVNITIKQHTVTTTTK in mouse $PrP^c$ (Figure 1a), with a score of 1.83, was assessed to be frustrated or discordant [32••]. Other measures of quantifying the structure showed that the maximal frustration is localized in the second half (C-terminal of helix 2) [32••]. The validity of this prediction finds support in the analysis of mutants of the PRNP gene associated with inherited transmissible spongiform encephalopathies (familial Creutzfeldt–Jakob disease [CJD] and fatal familial insomnia [FFI]). According to SWISS-PROT [63], seven disease-causing point mutations (D178N, V180I, T183A, H187R, T188R, T188K and T188A) are localized in helix 2. (We have used the sequence numbering for mouse $PrP^c$). A naive use of propensities to form helices (similar to those of Chou and Fasman [64]) would suggest that, with the exception of D178N, all other point mutations should lead to better helix formation. However, the $S_{\alpha/\beta}$ scores for the mutants are 1.94, 1.80, 1.30, 1.80, 1.54, 1.94 and 1.94 for D178N, V180I, T183A, H187R, T188K, T188R and T188A respectively. Thus, in all these mutants, helix 2 is frustrated, making it susceptible to the conformational fluctuations that have to occur before fibrillization. The differences in $S_{\alpha/\beta}$, which can be correlated with local stability, suggest that stability alone may not be a good indicator of the kinetics of amyloid formation.

As stated earlier, there are many unsatisfied HBs in $PrP^c$. Several of these mismatches are found in helices 2 and 3 (Figure 1a). If these regions become exposed upon $PrP^c \rightarrow PrP^{c*}$ transition, then minimization of the dangling HBs can be accomplished by polymerization of $PrP^{c*}$. Measures of frustration and other structural characteristics suggest that even segments of the rigid and ordered part of $PrP^c$ may play a key role in the production of $PrP^{c*}$. When the theoretical studies (which showed that regions of helices 2 and 3 could be involved in the $PrP^c \rightarrow PrP^{c*}$ transition) appeared, there was no direct experimental support. Subsequently, using $^{15}N$-$^{1}H$ two-dimensional NMR measurements as a function of pressure, Kuwata *et al.* [48••] have concluded that, in $PrP^{c*}$, helices 2 and 3 are disordered. The disordered metastable intermediates may be precursors in the templated assembly that converts $PrP^c$ to $PrP^{c*}$. This study shows, in accord with the theoretical predictions [31••,32••], that the core of $PrP^c$ is involved in producing the assembly-competent $PrP^{c*}$. Although the mechanism leading to $PrP^{Sc}$ is still unknown, it is worth emphasizing that the concept of SSE frustration in the wild-type proteins may be a useful indicator of the regions that harbor amyloidogenic tendencies.

The gene encoding the Doppel protein (Dpl), termed *Prnd* [65], is a paralog of the prion protein gene, *Prnp*, to which it has about 25% identity. Normally, Dpl is not expressed in the central nervous system, but it is up-regulated in mice with knockout *Prnp* gene. In such cases, overexpression of Dpl causes ataxia with Purkinje cell degeneration [65], which in turn can be cured by the introduction of one copy of the wild-type PrP mouse gene [66]. NMR studies of the three-dimensional structure of mouse Dpl [67] (Figure 1b) showed that it is structurally similar to $PrP^c$. However,

$PrP^c$ and Dpl produce diseases of the central nervous system using very different mechanisms: $PrP^c$ causes disease only after conversion to the $PrP^{Sc}$ form, whereas simple overexpression of Dpl, with no requirement for the scrapie form, causes ataxia. The markedly different disease mechanisms of PrP and Dpl would suggest, in light of the findings for $PrP^c$, that mouse Dpl (PDB code 1i17) would not be frustrated. Indeed, prediction of secondary structure by PHD [62] for mouse Dpl correlates well with the experimentally derived structure. The only difference between the predicted and derived structures of Dpl is found in the first β-strand region, which is predicted to be helical by PHD. However, corresponding $S_{\beta/\alpha}$ is –3.0, indicating that this α-helix prediction is unreliable as this sequence has low complexity. Also, analysis of mouse Dpl with the WHAT CHECK program [68] reveals that, on average, there are only eight unsatisfied buried HB donors/acceptors, representing 7.4% of all residues in mouse Dpl. This is comparable with the average value of 6% found in normal proteins, but it is markedly smaller than the 14% seen in mouse PrP (PDB code 1ag2). This analysis rationalizes the lack of observed scrapie formation in Dpl.

## Structures of amyloid fibrils

To understand the assembly mechanisms of amyloid fibrils, it is necessary to determine the fibril structures in atomic detail. Noncrystallinity and insolubility of the amyloid fibrils have made it difficult to obtain high-resolution fibril structures. Nevertheless, in the past few years, a variety of experimental and computational techniques have been used to provide a glimpse into the detailed architecture of fibrils in a variety of systems. Here, we focus on three such systems.

### Human prion protein dimer

In an important paper, Knaus *et al.* [69••] announced a 2 Å crystal structure of the dimeric form of the human prion protein (residues 90–231). The structure suggests that dimerization occurs by a domain-swap mechanism, in which helix 3 from one monomer packs against helix 2 from another. In fact, Eisenberg and co-workers [70] have suggested that a domain-swapping mechanism may be a general route for amyloid fibril formation. The electron density map seems to suggest structural fluctuations in residues 189–198, which coincide with the maximally frustrated region predicted theoretically [31••,32••]. The dimer interface is stabilized by residues that are in helix 2 in the monomeric NMR structure. The header of the PDB file of the monomeric structure of human $PrP^c$ indicates that helix 2 ends at residue 194 and helix 3 begins at 200. The domain-swapped dimer structure shows that residues 190–198 exist largely in a β-strand conformation. It appears that the α→β transition minimizes frustration. One implication of the dimer structure is that oligomerization occurs by domain swapping, which, in $PrP^c$, may also involve the disulfide bond between the cysteine residues at 179 and 214. The role of the disulfide bond in $PrP^{Sc}$ formation remains controversial.

### Structural characteristics of Aβ fibrils

Recent solid-state NMR studies have provided, for the first time, direct measurements of interatomic distances between labeled residues in Aβ amyloid fibrils [22]. These studies have suggested that the arrangements of the strands in the fibrils depend on the length of the peptide. The parallel in-register organization of peptides in β sheets was proposed for both $A\beta_{10-35}$ [21••] and $A\beta_{1-40}$ [23•,71••] fibrils. Such an organization raises the question: how is the destabilizing electrostatic repulsion due to close placement of like charges in parallel registry accommodated in $A\beta_{1-40}$ fibrils? To answer this question, Tycko and co-workers [71••] proposed a novel structural model for $A\beta_{1-40}$ peptide organization into fibrils. The measurement of correlations between $^{13}C$ and $^{15}N$ chemical shifts to probe the conformations (in terms of ϕ and ψ angles) of individual residues in $A\beta_{1-40}$ peptides showed that residues 12–24 and 30–40 adopt β-strand structure. Residues 25–29 form a bend. On the basis of these and previous findings [22,23•], Tycko and co-workers proposed that there are two β strands in $A\beta_{1-40}$, which form an in-register parallel β sheet. The β sheet HBs run parallel to the fibril axis. The formation of an intrapeptide salt bridge between charged residues D23 and K28 lends stability to the interstrand interactions. Taking into account the measurements of mass-per-length by scanning transmission electron spectroscopy, they proposed that individual $A\beta_{1-40}$ peptides are juxtaposed to form a dimer, which serves as an elementary building block of parallel β sheets. Their structural model is based on the premise, which finds support in MD simulations of $A\beta_{16-22}$ peptides [30••], that fibril structures form by maximizing favorable hydrophobic and electrostatic (salt bridge) interactions.

The structural model for $A\beta_{1-40}$ amyloids proposed by Tycko and co-workers was independently predicted by Ma and Nussinov [27•]. Using MD simulations, they probed the stabilities of various structural arrangements of $A\beta_{10-35}$ peptides. Although the simulation results cannot be conclusive because of the short duration and the lack of equilibration of the initial structures, they provide valuable insights into the fibrillar architecture. Similar to the model of Tycko and co-workers, a turn (at positions 24–27) is proposed, which is reinforced by the intrapeptide salt bridge D23–K28.

In contrast to $A\beta_{1-40}$, a different structural organization is envisioned for $A\beta_{10-35}$ fibrils. To maximize hydrophobic interactions, two $A\beta_{10-35}$ peptides in turn conformations are docked end-to-end, locking unmatched residues 10–16. A hydrophobic core is centered near position L34. Thus, both solid-state NMR measurements and MD simulations suggest that the parallel in-registry structure of Aβ peptides with a turn in the middle appears to be the most stable arrangement for long Aβ peptides. The proposed structure for $A\beta_{10-35}$ fibrils [27•] is at variance

with the structure suggested by Lynn and co-workers [21••,72]. In contrast to $A\beta_{1-40}$ [71••] and $A\beta_{10-35}$ [21••], the 16–22 and 34–42 fragments have been shown to form antiparallel β sheets [20,73]. We found that $A\beta_{16-22}$ forms an in-registry antiparallel organization, which is favored by interpeptide K16–D22 salt bridges and hydrophobic interactions between aromatic residues [30••].

Experiments so far have not probed the structural organization of short peptides in fibrils beyond single β sheets. Several MD simulations have been performed to examine the three-dimensional structures of amyloid fibrils [27•,74]. The interesting study of Ma and Nussinov [27•] found that $A\beta_{16-22}$ forms the most stable fibril structure, which is arranged in antiparallel in-registry β sheets that propagate parallel to each other. This structural organization provides close contacts between oppositely charged lysine and glutamic acid residues, and also establishes optimal (parallel) registry between phenylalanine residues in neighboring β sheets. Simulations of 24-mer fibril blocks revealed a significant twist angle of 15° per peptide.

### Structure of aggregates of tau protein

Paired helical filaments (PHFs), which are primarily aggregates of the microtubule-associated tau protein, also accumulate in neurons of patients with Alzheimer's disease. The prevailing view is that the insoluble filaments are composed of β sheets, giving credence to the notion that the formation of such structures is a universal characteristic of all disease-causing proteins. The soluble tau protein, in the monomeric form, is known to be unstructured [75]. Using far-UV circular dichroism (CD) and Fourier transform infrared (FTIR) spectroscopy, Sadqi *et al.* [55••] showed that the PHF, contrary to popular belief, is predominantly helical. In this case, there must be a structural transition in the major protein component of PHF, namely tau, from a random coil to α helix [55••,76•].

## Multiple routes to fibril formation

Although significant progress has been made in the determination of low-resolution structures of fibrils, relatively little is known, at the molecular level, about the cascade of events that leads to aggregation. Several experimental studies suggest that, generically, fibril formation exhibits all the characteristics of a nucleation-growth process [16]. The kinetics of fibril formation has a lag phase provided the protein concentration exceeds a critical value. The lag phase disappears if a seed or preformed nucleus is present in the supersaturated solution. The seeded growth of fibrils, which closely resembles the templated assembly envisioned by Griffith to explain self-replication of proteins [77], has been explicitly verified in simple lattice models [56••,78].

One of the most popular beliefs is that fibrillization requires partial unfolding of the native state or partial folding of the unfolded state [10,12–14]. Both events, which are likely to involve crossing free energy barriers (Figure 4), produce the assembly-competent structure **N***. The **N*** state in TTR, which has a higher free energy than the native state **N**, is formed upon the unraveling of edge strands C and D, thus exposing strand B [13,58••]. One can also envision a scenario in which **N*** has a lower free energy than **N**, thus making the monomeric native state metastable (Figure 4). We conjecture that amyloidogenic proteins, in which nearly complete transformation of the structure takes place upon fibrillization, may follow the second scenario. Both of these possibilities follow from an energy landscape perspective of aggregation [79••]. In both cases (Figure 4), fibrillization kinetics should be determined by an 'unfolding' free energy barrier separating the **N** and **N*** states. Recent studies of the fibrillization of $PrP^c$ and TTR provide experimental support for this concept [36••]. The perspective sketched in Figure 4 also suggests that the free energy of stability of **N** may not be a good indicator of the rates of fibrillization.

In scenario I, the amyloidogenic state **N*** is formed by denaturation stress. The production of **N*** in scenario II can occur by two distinct routes. If **N** is metastable, as is apparently the case for $PrP^c$ [80•], then conformational fluctuations can lead to **N***. Alternatively, formation of **N*** can also be triggered by intermolecular interactions (this possibility presumably applies for Aβ peptides). In the latter case, **N*** can only form when the protein concentration exceeds a threshold value.

To better understand the kinetics of fibrillization, it is necessary to characterize the early events and pathways leading to the formation of the critical nucleus. In terms of the two scenarios outlined above, the structures of **N***, the ensemble of transition state structures and the conformations of the critical nuclei must be known to fully understand the assembly kinetics. A significant step in this direction has been taken by Teplow and co-workers [29••], who have followed the growth of fibrils for 18 peptides, including $A\beta_{1-40}$ and $A\beta_{1-42}$. In all cases, the formation of amyloids by $A\beta_{1-40}$ and $A\beta_{1-42}$ is preceded by the formation of the intermediate oligomeric state with high α-helical content. This is remarkable given that both the monomers and fibrils have little or no α-helical content. Therefore, the transient accumulation of α-helical structure represents an obligatory (on-pathway) intermediate state, which coincides with the onset of oligomerization.

Because it is experimentally difficult to atomically map the events leading to fibrillization, we have carried out multiple long MD simulations to probe the oligomerization of $A\beta_{16-22}$ peptides [30••]. This peptide, which is disordered in the monomeric form, assembles into an antiparallel β structure through interpeptide interactions. Even in the oligomerization of these small peptides from the Aβ family, the assembly was preceded by the formation of an on-pathway α-helical intermediate. Based on our findings and the work by Teplow and co-workers, we postulated that the

formation of oligomers rich in α-helical structure may be a universal mechanism for Aβ peptides.

Formation of the on-pathway α-helical intermediate may be rationalized using the following arguments. The initial events involve the formation of 'disordered' oligomers, driven by hydrophobic interactions that reduce the effective volume available to each Aβ peptide. In the confined space, peptides adopt α-helical structure. Further structural changes are determined by the requirement of maximizing the number of favorable hydrophobic and electrostatic interactions. This can be achieved if Aβ peptides adopt ordered extended β-like conformations provided that oligomers contain a sufficiently large number of peptides.

There is some similarity between the aggregation mechanism postulated for Aβ peptides and the nucleated conformational conversion (NCC) model envisioned for the conversion of Sup35 to [$PSI^+$] in *Saccharomyces cerevisiae*. By studying the assembly kinetics of Sup35, Serio *et al.* [81••] proposed the NCC model, which combines parts of the templated assembly and nucleation-growth mechanisms. The hallmark of the NCC model is the formation of a critical-sized mobile oligomer, in which Sup35 adopts a conformation that may be distinct from its monomeric random coil or the conformation it adopts in the aggregated state. The formation of a critical nucleus, to which other Sup35 can assemble, involves a conformational change to the state that it adopts in self-propagating [$PSI^+$]. The α-helical intermediate seen in Aβ peptides may well correspond to the mobile oligomer that has the 'wrong' conformation to induce further assembly. Thus, as noted by Lindquist and co-workers [81••], NCC may serve as a unifying model for protein aggregation.

## Conclusions

The development of methods to envision the structure of amyloid fibrils has enabled us to obtain molecular insights into the assembly process itself. Computational and experimental studies are beginning to provide detailed information, at the residue level, about the regions in a given protein that harbor amyloidogenic tendencies. We have harnessed these developments to propose tentative ideas on the molecular basis of protein aggregation. These principles (or, more precisely, rules of thumb) may be useful in the interpretation and design of new experiments.

Despite the great progress that has been made in the past few years, several outstanding issues still require clarification. Are there common pathways involved in the self-assembly of fibrils? Because of the paucity of the structural description of the intermediates involved in the aggregation process, a definitive answer cannot be currently provided. The energy landscape perspective, summarized briefly in Figure 4, suggests that multiple scenarios for assembly exist. Although the generic nucleation-growth mechanism governs fibril formation, the details can vary considerably. A complete understanding will require experiments along the lines initiated by Teplow and co-workers [29••]. The microscopic basis for the formation of distinct strains in mammalian prions and in yeast prions remains a mystery. Are these merely associated with the heterogeneous seeds or are there unidentified mechanisms that lead to their formation? What are the factors that determine the variations in the fibrillization kinetics for the wild type and the mutants? A tentative proposal is that the kinetics of polymerization is determined by the rate of production of **N*** (Figure 4) [82], which in turn is controlled by barriers separating **N** and **N*** [32••,36••]. In this scenario, the stability of **N** plays a secondary role. The generality of this observation has not yet been established. Finally, how can one design better therapeutic agents based on enhanced knowledge of the assembly mechanism? Even in the case of sickle cell disease, viable therapies began to emerge only long after the biophysical aspects of gelation were understood [83].

## Update

Recently, Bitan *et al.* [86] showed that $A\beta_{1-40}$ and $A\beta_{1-42}$ oligomerize by distinct pathways. The oligomerization of this class of peptides follows scenario II in Figure 4. The distinct rates of fibril formation of $A\beta_{1-40}$ and $A\beta_{1-42}$ can be rationalized in terms of the variations in the free energy barrier heights separating U and N*. A quantitative assessment of this proposal will require temperature-dependent measurements of oligomerization rates.

The scenarios for fibrillization shown in Figure 4 imply that aggregation may be prevented by destabilizing N*. Hammarstrom *et al.* [87] have recently used this strategy to devise a way to prevent transthyretin amyloidosis by having inhibitors increase the kinetic barrier separating N and N*. Based on this study, they propose that using the small-molecule binding strategy is an effective way of treating a number of amyloid diseases. This study also highlights the use of biophysical methods in coming up with plausible therapies for this class of amyloid disease.


## Acknowledgements

We thank John Straub and Robert Tycko for several useful conversations. The comments by David Teplow are greatly appreciated. This work was supported in part by a grant from the National Institutes of Health (IR01 NS41356-01).


## References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest


1. Selkoe DJ: **Alzheimer's disease: genes, proteins, and therapy.** *Physiol Rev* 2001, **81**:741-766.
2. Hardy J, Selkoe DJ: **The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics.** *Science* 2002, **297**:353-356.
3. Prusiner S: **Prions.** *Proc Natl Acad Sci USA* 1998, **95**:12580-12585.
4. Scott MR, Robert W, Ironside J, Nguyen H-OB, Tremblay P, DeArmond SJ, Prusiner SB: **Compelling transgenetic evidence for transmission of bovine spongiform**

**encephalopathy prions to humans.** *Proc Natl Acad Sci USA* 1999, **96**:15137-15142.

5. Kisilevsky R: **Review: Amyloidogenesis - unquestioned answers and unanswered questions.** *J Struct Biol* 2000, **130**:99-108.

6. Lambert MP, Barlow AK, Chromy BA, Edwards C, Freed R, Liosatos M, Morgan TE, Rozovsky I, Trommer B, Viola KL: **Diffusible, nonfibrillar ligands derived from Aβ1-42 are potent central nervous system neurotoxins.** *Proc Natl Acad Sci USA* 1998, **95**:6448-6453.

7. Walsh DM, Hartley DM, Kusumoto Y, Fezoui Y, Condron MM, Lomakin A, Benedek GB, Selkoe DJ, Teplow DB: **Amyloid β-protein fibrillogenesis. Structure and biological activity of protofibrillar intermediates.** *J Biol Chem* 1999, **274**:25945-25952.

8. Dobson CM: **Protein misfolding, evolution and disease.** *Trends Biochem Sci* 1999, **24**:329-332.

9. Chiti F, Webster P, Taddei N, Clark A, Stefani M, Ramponi G, Dobson CM: **Designing conditions for *in vitro* formation of amyloid protofilaments and fibrils.** *Proc Natl Acad Sci USA* 1999, **96**:3590-3594.

10. Rochet JC, Lansbury PT: **Amyloid fibrillogenesis: themes and variations.** *Curr Opin Struct Biol* 2000, **10**:60-68.

••11. Bucciantini M, Giannoni E, Chiti F, Baroni F, Formigli L, Zurdo J, Taddei N, Ramponi G, Dobson CM, Stefani M: **Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases.** *Nature* 2002, **416**:507-511.

It is shown that species that form early in the aggregation of proteins (SH3 domain and bovine phosphatidyl-inositol 3′ kinase) that are not associated with known diseases may be inherently cytotoxic. A clear implication of this study is that natural proteins have evolved not only to fold to the desired folded state but also to avoid aggregation.

12. Kelly JW: **Alternative conformations of amyloidogenic proteins govern their behavior.** *Curr Opin Struct Biol* 1996, **6**:11-17.

13. Kelly JW: **The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways.** *Curr Opin Struct Biol* 1998, **8**:101-106.

14. Fink AL: **Protein aggregation: folding aggregates, inclusion bodies and amyloid.** *Fold Des* 1998, **3**:R9-R23.

15. Eaton WA, Hofrichter J: **Sickle cell hemoglobin polymerization.** *Adv Protein Chem* 1990, **40**:63-279.

16. Harper JD, Lansbury PT: **Models of amyloid seeding in Alzheimer's disease and scrapie: mechanistic truths and physiological consequences of time-dependent stability of amyloid proteins.** *Annu Rev Biochem* 1997, **66**:385-407.

17. Bessen RA, Kocisko DA, Raymond GJ, Nandan S, Lansbury PT, Caughey B: **Non-genetic propagation of strain-specific properties of scrapie prion protein.** *Nature* 1995, **375**:698-700.

18. Collinge J, Sidle KC, Meads J, Ironside J, Hill AF: **Molecular analysis of prion strain variation and the etiology of new variant CJD.** *Nature* 1996, **383**:685-690.

19. Safar J, Wille H, Itri V, Groth D, Serban H, Torchia M, Cohen FE, Prusiner SB: **Eight prion strains have $PrP^{Sc}$ molecules with different conformations.** *Nat Med* 1998, **4**:1157-1165.

20. Lansbury PT, Costa PR, Griffiths JM, Simon EJ, Auger M, Halverson K, Kocisko DA, Hendsch ZS, Ashburn TT, Spencer R *et al.*: **Structural model for the beta-amyloid fibril based on interstrand alignment of an antiparallel-sheet comprising a C-terminal peptide.** *Nat Struct Biol* 1995, **2**:990-998.

••21. Burkoth TS, Benzinger T, Urban V, Morgan DM, Gregory DM, Thiyagarajan P, Botto RE, Meredith SC, Lynn DG: **Structure of the β-amyloid(10-35) fibril.** *J Am Chem Soc* 2000, **122**:7883-7889.

The authors show that $A\beta_{10-35}$ strands are arranged in a parallel manner in the fibrils. This study also emphasizes cooperative motion of hydrogen bonds that stabilize the parallel arrangement of the $A\beta_{10-35}$ peptides.

22. Tycko R: **Solid-state NMR as a probe of amyloid fibril structure.** *Curr Opin Chem Biol* 2000, **4**:500-506.

•23. Antzutkin ON, Balbach JJ, Leapman RD, Rizzo NW, Reed J, Tycko R: **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* 2000, **97**:13045-13050.

One of the first reports to suggest that $A\beta_{1-40}$ peptides are arranged as parallel β sheets in amyloid fibrils.

24. Harper JD, Wong SS, Lieber CM, Lansbury PT: **Assembly of Aβ amyloid protofibrils: an *in vitro* model for a possible early event in Alzheimer's disease.** *Biochemistry* 1999, **38**:8972-8980.

25. Jimenez JL, Nettleton EJ, Bouchard M, Robinson CV, Dobson CM, Saibil HR: **The protofilament structure of insulin amyloid fibrils.** *Proc Natl Acad Sci USA* 2002, **99**:9196-9201.

26. Serpell JC, Smith JM: **Direct visualisation of the beta-sheet structure of synthetic Alzheimer's amyloid.** *J Mol Biol* 2000, **299**:225-231.

•27. Ma B, Nussinov R: **Stabilities and conformations of Alzheimer's β-amyloid peptide oligomers ($A\beta_{16-22}$, $A\beta_{16-35}$, and $A\beta_{10-35}$): sequence effects.** *Proc Natl Acad Sci USA* 2002, **99**:14126-14131.

MD simulations were used to assess the stabilities of Aβ peptides. By comparing the stabilities of different arrangements of peptides in the fibril, the authors have proposed a novel structural model for aggregates of Aβ peptides.

28. Walsh DM, Klyubin I, Fadeeva JV, Cullen WK, Anwyl R, Wolfe MS, Rowan MJ, Selkoe DJ: **Naturally secreted oligomers of amyloid beta protein potently inhibit hippocampal long-term potentiation *in vivo.*** *Nature* 2002, **416**:535-539.

••29. Kirkitadze MD, Condron MM, Teplow DB: **Identification and characterization of key kinetic intermediates in amyloid β-protein fibrillogenesis.** *J Mol Biol* 2001, **312**:1103-1119.

One of the first studies to describe the pathways in the assembly of Aβ peptides into fibrils. By studying the assembly kinetics of 18 different Aβ peptides, the authors suggest that fibril formation is preceded by the transient population of on-pathway α-helical intermediates.

••30. Klimov DK, Thirumalai D: **Dissecting the assembly of A $\beta_{16-22}$ amyloid peptides into antiparallel β-sheets.** *Structure* 2003, **11**:295-307.

Using a series of MD simulations, it is shown that $A\beta_{16-22}$ peptides form antiparallel β-sheet structures. Even in this system, α-helical intermediates are transiently populated. This led the authors to propose that fibril formation by Aβ peptides, which occurs by maximizing the number of salt bridges and hydrophobic interactions, must involve oligomers with high α-helical content.

••31. Kallberg Y, Gustafsson M, Persson B, Thyberg J, Johansson J: **Prediction of amyloid fibril-forming proteins.** *J Biol Chem* 2001, **276**:12945-12950.

This study shows that several amyloid-forming proteins harbor α helices in regions that are theoretically predicted to be β strands. Using this incompatibility, the authors predicted the specific regions of $PrP^c$ and Aβ peptides that may be involved in fibril formation.

••32. Dima RI, Thirumalai D: **Exploring the propensities of helices in $PrP^c$ to form β sheet using NMR structures and sequence alignments.** *Biophys J* 2002, **83**:1268-1280.

Using several sequence alignments and structural characteristics of $PrP^c$, including the concept of frustration of SSEs (see also

[31••]), the authors predict that helices 2 and 3 are implicated in the conformational transition from $PrP^c$ to $PrP^{c*}$.

••33. Richardson JS, Richardson DC: **Natural β-sheet proteins use negative design to avoid edge-to-edge aggregation.** *Proc Natl Acad Sci USA* 2002, **99**:2754-2759.

This novel study illustrates the general strategies that naturally occurring β sheets use to avoid aggregation. The principles that emerge from this paper are not only relevant to the prediction of regions in proteins that are susceptible to aggregation, but also are important in the *de novo* design of proteins with β-sheet architecture.

34. Esler WP, Stimson ER, Lachenmann MJ, Ghilardi JR, Lu Y, Vinters HV, Mantyh PW, Lee JP, Maggio JE: **Activation barriers to structural transition determine deposition rates of Alzheimer's disease.** *J Struct Biol* 2000, **130**:174-183.

35. Sian AK, Frears ER, El-Agnaf OA, Patel BP, Manca MF, Siligardi G, Hussain R, Austen BM: **Oligomerization of β-amyloid of the Alzheimer's and the dutch-cerebral-haemorrhage.** *Biochem J* 2000, **349**:299-308.

••36. Hammarstrom P, Jiang X, Hurshman AR, Powers ET, Kelly JW: **Sequence-dependent denaturation energetics: a major determinant in amyloid disease diversity.** *Proc Natl Acad Sci USA* 2002, **99**:16427-16432.

By studying the kinetics of polymerization of wild-type TTR and several mutants, the authors suggest that unfolding rates of the functional state of the normal protein are a key determinant of aggregation. A similar conclusion has been reached in [32••] in the context of $PrP^c$.

37. Watson DJ, Landers AD, Selkoe DJ: **Heparin-binding properties of the amyloidogenic peptides Aβ and amylin.** *J Biol Chem* 1997, **272**:31617-31624.

38. Miravalle L, Tokuda T, Chiarle R, Giaccone G, Bugiani O, Ragliavini F, Frangione B, Ghiso J: **Substitutions at codon 22 of Alzheimer's Aβ peptide induce diverse conformational changes and apoptotic effects in human cerebral endothelial cells.** *J Biol Chem* 2000, **275**:27110-27116.

39. Massi F, Peng JW, Lee JP, Straub JE: **Simulation study of the structure and dynamics of the Alzheimer's amyloid peptide congener in solution.** *Biophys J* 2001, **80**:31-44.

40. Massi F, Straub JE: **Probing the origins of increased activity of the E22Q Dutch mutant of the Alzheimer's β-amyloid peptide.** *Biophys J* 2001, **81**:697-709.

••41. Massi F, Klimov D, Thirumalai D, Straub JE: **Charged 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* 2002, **11**:1639-1647.

To test the hypothesis that differences in the tendency to form β strands may explain the variations in the deposition rates of wild-type Aβ and the E22Q mutant, the authors performed a series of MD simulations. They showed that the desolvation penalty, rather than intrinsic preferences for β structure, explains the differences in the fibrillization kinetics.

•42. Riek R, Guntert P, Dobeli H, Wipf B, Wuthrich K: **NMR studies in aqueous solution fail to identify significant conformational differences between the monomeric forms of two Alzheimer peptides with widely different plaque-competence, $A\beta^{ox}_{1-40}$ and $A\beta^{ox}_{1-42}$.** *Eur J Biochem* 2001, **268**:5930-5936.

The lack of differences in the solution NMR structures of $A\beta^{ox}_{1-40}$ and $A\beta^{ox}_{1-42}$ shows that monomer conformations alone are not indicative of the mechanism of polymerization.

43. Jarrett JT, Berger EP, Lansbury PT: **The carboxy terminus of β amyloid protein in critical for the seeding of amyloid formation: implications for the pathogenesis of Alzheimer's disease.** *Biochemistry* 1993, **32**:4693-4697.

44. Lee JP, Stimson ER, Ghilardi JR, Mantyh PW, Lu YA, Felix AM, Llanos W, Behbin A, Cummings M, Criekinge MV *et al.*: **$^1$H NMR of Aβ amyloid peptide congeners in water solution. Conformational changes correlate with plaque competence.** *Biochemistry* 1995, **34**:5191-5200.

•45. Alonso DOV, DeArmond SJ, Cohen FE, Daggett V: **Mapping the early steps in the pH-induced conformational conversion of the prion protein.** *Proc Natl Acad Sci USA* 2001, **98**:2985-2989.

MD simulations of $PrP^c$ at neutral and low pH were used to probe the early steps in the conformational transition. The study shows that it is necessary to simulate conformational fluctuations under a variety of external conditions to obtain insights into the nature of events in the transition from $PrP^c$ to $PrP^{c*}$.

46. Swietnicki W, Petersen R, Gambetti P, Surewicz WK: **pH-dependent stability and conformation of the recombinant human prion protein PrP(90-231).** *J Biol Chem* 1997, **272**:27517-27520.

47. Hornemann S, Glockshuber R: **A scrapie-like unfolding intermediate of the prion protein domain PrP(121-231) induced by acidic pH.** *Proc Natl Acad Sci USA* 1998, **95**:6010-6014.

••48. Kuwata K, Li H, Yamada H, Legname G, Prusiner SB, Akasaka K, James TL: **Locally disordered conformer of the hamster prion protein: a crucial intermediate to $PrP^{Sc}$?** *Biochemistry* 2002, **41**:12277-12283.

Using two-dimensional NMR measurements at various pressures on Syrian hamster $PrP^c$ (90–231), the authors propose that, in the amyloidogenic intermediate $PrP^{c*}$, regions of helices 2 and 3 are disordered. Their key finding that helices 2 and 3 are implicated in the $PrP^c \rightarrow PrP^{c*}$ transition is in accord with previous theoretical predictions [31••,32••].

49. Lorimer GH: **A quantitative assessment of the role of the chaperonin proteins in protein folding *in vivo.*** *FASEB J* 1996, **10**:5-9.

50. Anfinsen CB: **Principles that govern the folding of protein chains.** *Science* 1973, **181**:223-230.

51. Otzen DE, Oliveberg M: **Salt-induced detour through compact regions of the protein folding landscape.** *Proc Natl Acad Sci USA* 1999, **96**:11746-11751.

••52. Otzen DE, Kristensen O, Oliveberg M: **Designed protein tetramer zipped together with a hydrophobic Alzheimer homology: a structural clue to amyloid assembly.** *Proc Natl Acad Sci USA* 2000, **97**:9907-9912.

This is the first study that used mutational analysis to investigate the factors that prevent wild-type S6 from aggregation. The experiments were used to introduce a key idea – natural proteins use charged gatekeeper residues (ones that are not involved in directing the folding process) to prevent intermolecular association.

••53. Perutz MF, Finch JT, Berriman J, Lesk A: **Amyloid fibers are water-filled nanotubes.** *Proc Natl Acad Sci USA* 2002, **99**:5591-5595.

Using available experimental data and novel modeling methods, this study suggests that amyloid fibrils may have a near 'universal' structure.

54. McDonald I, Thornton JM: **Satisfying hydrogen bonding potential in proteins.** *J Mol Biol* 1994, **238**:777-793.

••55. Sadqi M, Hernandez F, Pan UM, Perez M, Schaeberle MD, Avila J, Munoz V: **Alpha-helix structure in Alzheimer's disease aggregates of tau-protein.** *Biochemistry* 2002, **41**:7150-7155.

This study challenges the commonly held belief that all amyloid fibrils comprise β-sheet structures. Using CD and FTIR spectroscopy, the authors suggest that, in the PHFs, the tau protein adopts an α-helical conformation.

••56. Dima RI, Thirumalai D: **Exploring protein aggregation and self-propagation using lattice models: phase diagram and kinetics.** *Protein Sci* 2002, **11**:1036-1049.
The authors use lattice models to describe various scenarios for protein aggregation. They use a toy model to verify the templated aggregation model for prion propagation [77].

•57. Wang W, Hecht MH: **Rationally designed mutations convert *de novo* amyloid-like fibrils into monomeric β-sheet proteins.** *Proc Natl Acad Sci USA* 2002, **99**:2760-2765.
The authors show that inserting lysine in the middle of an alternating sequence of hydrophobic and hydrophilic residues results in soluble β-sheet proteins. This study provides direct experimental support to the ideas developed in [33••].

••58. Serag AA, Altenbach C, Gingery M, Hubbell WL, Yeates TO: **Arrangement of subunits and ordering of β-strands in an amyloid sheet.** *Nat Struct Biol* 2002, **9**:734-739.
Using EPR, the authors propose a model for TTR fibrils. The rationale for this model follows from a combination of ideas presented in [33••,59].

59. Klimov DK, Thirumalai D: **Native topology determines force-induced unfolding pathways in globular proteins.** *Proc Natl Acad Sci USA* 2000, **97**:7254-7259.

60. West MW, Wang W, Patterson J, Mancias JD, Beasley JR, Hecht MH: ***De novo* amyloid proteins from designed combinatorial libraries.** *Proc Natl Acad Sci USA* 1999, **96**:11211-11216.

61. Schwartz R, Istrail S, King J: **Frequencies of amino acid string in globular protein sequences indicate suppression of blocks of consecutive hydrophobic residues.** *Protein Sci* 2001, **10**:1023-1031.

62. Rost B, Sander C: **Prediction of protein secondary structure at better than 70% accuracy.** *J Mol Biol* 1993, **232**:584-599.

63. Bairoch A, Apweiler R: **The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000.** *Nucleic Acid Res* 2000, **28**:275-284.

64. Chou PY, Fasman GD: **Empirical predictions of protein conformation.** *Annu Rev Biochem* 1978, **47**:251-276.

65. Moore RC, Lee IY, Silverman GL, Harrison PM, Strome R, Heinrich C, Karunaratne A, Pasternak SH, Chishti MA, Liang Y *et al.*: **Ataxia in prion protein (PrP)-deficient mice is associated with upregulation of the novel PrP-like protein Doppel.** *J Mol Biol* 1999, **292**:797-817.

66. Nishida N, Tremblay P, Sugimoto T, Shigematsu K, Shirabe S, Petromilli C, Erpel SP, Nakaoke R, Atarashi R, Houtani T *et al.*: **A mouse prion protein transgene rescues mice deficient for the prion protein gene from Purkinje cell degeneration and demyelination.** *Lab Invest* 1999, **79**:689-697.

67. Mo H, Moore RC, Cohen FE, Westaway D, Prusiner SB, Wright PE, Dyson HJ: **Two different neurodegenerative diseases caused by proteins with similar structures.** *Proc Natl Acad Sci USA* 2001, **98**:2352-2357.

68. Hooft RWW, Vriend G, Sander C: **Errors in protein structures.** *Nature* 1996, **381**:272-273.

••69. Knaus KJ, Morillas M, Swietnicki W, Malone M, Surewicz WK, Yee VC: **Crystal structure of the human prion protein reveals a mechanism for oligomerization.** *Nat Struct Biol* 2001, **8**:770-774.
The first crystal structure of a dimer of human prion protein is reported at 2 Å resolution. This structure, which may be a building block for $PrP^{Sc}$, shows that the dimer forms by a domain-swap mechanism.

70. Liu Y, Gotte G, Libonati M, Eisenberg D: **A domain-swapped RNase A dimer with implications for amyloid formation.** *Nat Struct Biol* 2002, **8**:211-214.

••71. Petkova AT, Ishii Y, Balbach JJ, Antzutkin ON, Leapman RD, Delaglio F, Tycko R: **A structural model for Alzheimer's β-amyloid fibrils based on experimental constraints from solid state NMR.** *Biophys J* 2003, in press.
A novel structural model for the $A\beta_{1-40}$ amyloid fibril is proposed using solid-state NMR constraints. The authors showed that the number of salt bridges and hydrophobic interactions can be maximized by introducing a bend in each peptide (see also [27•]).

72. Benzinger T, Gregory DM, Burkoth TS, Miller-Auer H, Lynn DG, Botto RE, Meredith SC: **Propagating structure of Alzheimer's β-amyloid$_{10-35}$ is parallel β-sheet with residues in exact order.** *Proc Natl Acad Sci USA* 1998, **95**:13407-13412.

73. Balbach JJ, Ishii Y, Antzutkin ON, Leapman RD, Rizzo NW, Dyda F, Reed J, Tycko R: **Amyloid fibril formation by A$\beta_{16-22}$, a seven-residue fragment of the Alzheimer's β-amyloid peptide, and structural characterization by solid state NMR.** *Biochemistry* 2000, **39**:13748-13759.

74. Ma B, Nussinov R: **Molecular dynamics simulations of alanine rich β-sheet oligomers: insight into amyloid formation.** *Protein Sci* 2002, **11**:2335-2350.

75. Schweer O, Schonbrunn-Hanebeck E, Marx A, Mandelkow E: **Structural studies of tau protein and Alzheimer paired helical filaments show no evidence for beta-structure.** *J Biol Chem* 1994, **269**:24290-24297.

•76. Goux WJ: **The conformations of filamentous and soluble tau associated with Alzheimer paired helical filaments.** *Biochemistry* 2002, **41**:13798-13806.
Extensive investigation of the structure of the PHFs further confirms the conclusions reached in [55••].

77. Griffith JS: **Self-replication and scrapie.** *Nature* 1967, **215**:1043-1044.

78. Harrison PM, Chan HS, Prusiner SB, Cohen FE: **Conformational propagation with prion-like characteristics in a simple model of protein folding.** *Protein Sci* 2001, **10**:819-835.

••79. Massi F, Straub JE: **Energy landscape theory for Alzheimer's amyloid β-peptide fibril elongation.** *Proteins* 2001, **42**:217-229.
The authors have proposed several schemes for the kinetics of fibril formation by Aβ peptides. By using the analogy with the energy landscape perspective of protein folding, many new predictions are made for the assembly of amyloid fibrils.

•80. Baskakov IV, Legname G, Prusiner SB, Cohen FE: **Folding of prion protein to its native α-helical conformation is under kinetic control.** *J Biol Chem* 2001, **276**:19687-19690.
It is shown that $PrP^c$ may be metastable, which suggests that the rarity of prion disorders is associated with a large barrier separating $PrP^c$ and $PrP^{c*}$. This study also implies that the kinetics of $PrP^c$ formation follows scenario II (Figure 4).

••81. Serio TR, Cashikar AG, Kowal AS, Sawicki GJ, Moslehi JJ, Serpell L, Arnsdorf MF, Lindquist SL: **Nucleated conformational conversion and the replication of conformational information by a prion determinant.** *Science* 2000, **289**:1317-1321.
Based on the kinetics of $[PSI^+]$ production from Sup35 in *S. cerevisiae*, the authors proposed a novel model that incorporate elements of the templated assembly and nucleation-growth mechanisms. It is likely that recent studies on Aβ aggregation provide additional examples of the NCC model.

82. Ramirez-Alvarado M, Merkel JS, Regan L: **A systematic exploration of the influence of the protein stability on amyloid fibril formation *in vitro*.** *Proc Natl Acad Sci USA* 2000, **97**:8979-8984.

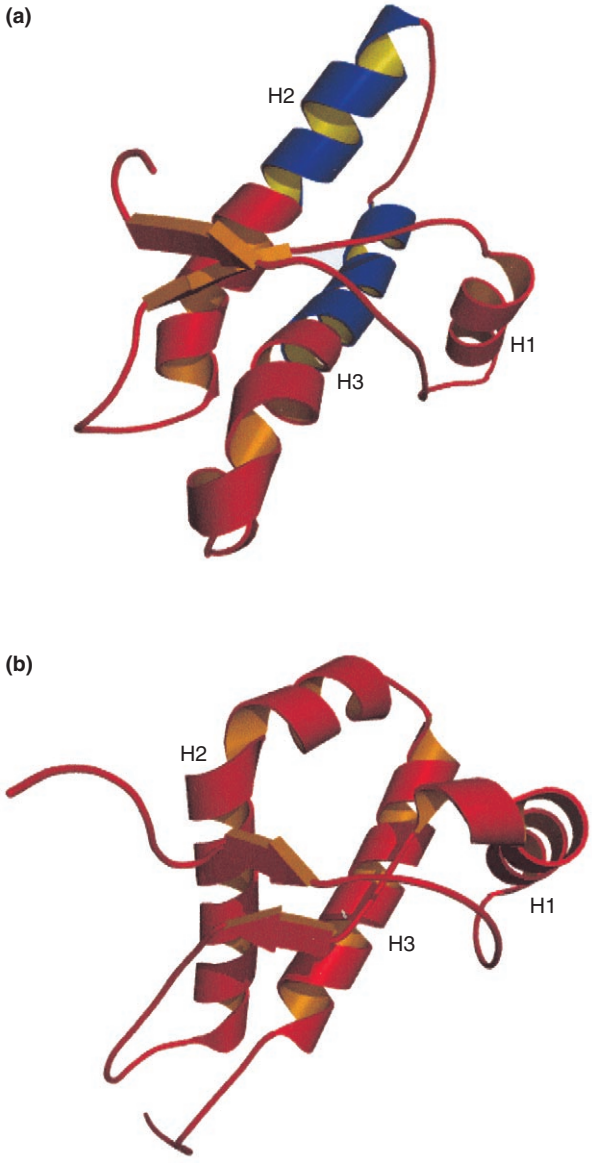
83. Eaton WA, Hofrichter J: **The biophysics of sickle cell hydroxyurea therapy.** *Science* 1995, **268**:1142-1143.

84. Kraulis PJ: **MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures.** *J Appl Crystallogr* 1991, **24**:946-950.

85. Koradi R, Billeter M, Wuthrich K: **Molmol: a program for display and analysis of macromolecular structures.** *J Mol Graph* 1996, **14**:51-55.

86. Bitan G, Kirkitadze MA, Lomakin A, Vollers SS, Benedeck GB, Teplow DB: **Amyloid β-protein (Aβ) assembly: Aβ40 and Aβ42 oligomerize through distinct pathways.** *Proc Natl Acad Sci USA* 2003, **100**:330-335.

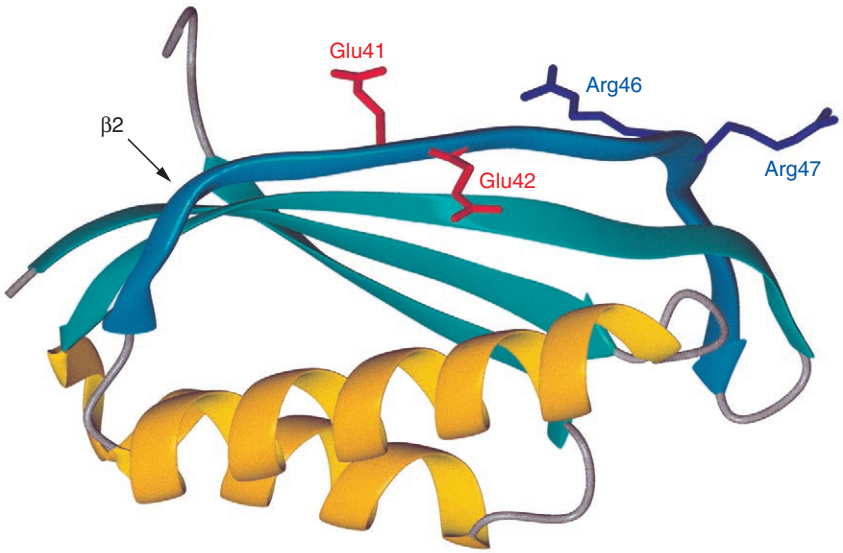
87. Hammarstrom P, Wiseman RL, Powers ET, Kelly JW: **Prevention of transthyretin amyloid disease by changing protein misfolding energetics.** *Science* 2003, **299**:713-716.




Current Opinion in Structural Biology

**Figure 1**

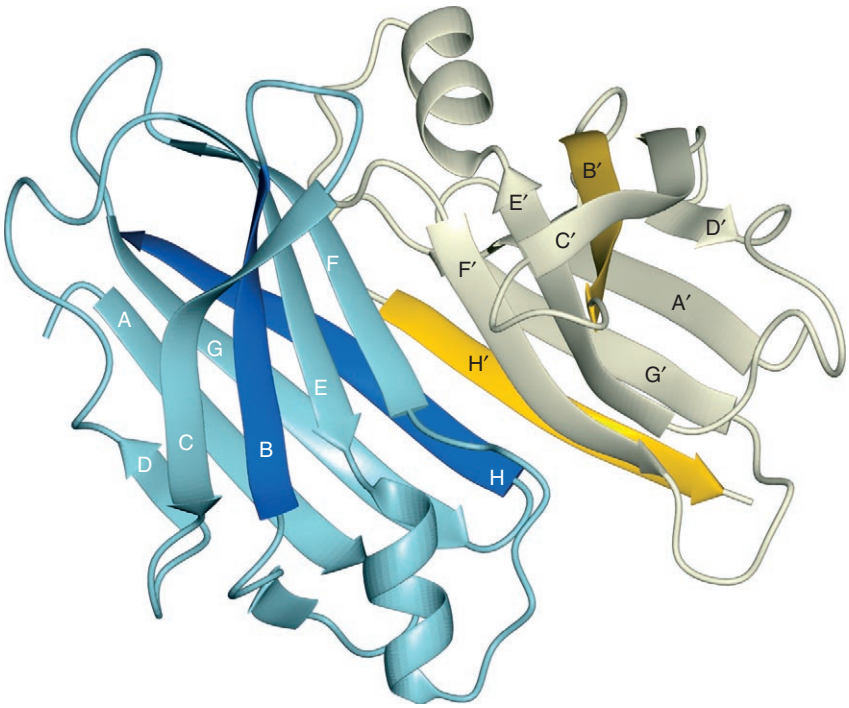
**(a)** NMR structure of mouse $PrP^{c}$ (PDB code 1ag2). Various measures of frustration (see text) between the sequence and its native three-dimensional structure show that at least the second half of helix 2 (H2) and part of the first half of helix 3 (H3) (colored in blue) are frustrated in the helical state. These regions, together with the disordered N-terminal segments, are implicated in the transition from $PrP^{c}$ to the assembly-competent structure $PrP^{c}$*. **(b)** Solution structure of mouse Dpl (PDB code 1i17). Despite the similarity of the two structures, no frustrated region is found in Dpl. This may explain the absence of the scrapie form in Doppel. The figures were produced with the program Molscript [84].




Current Opinion in Structural Biology

**Figure 2**

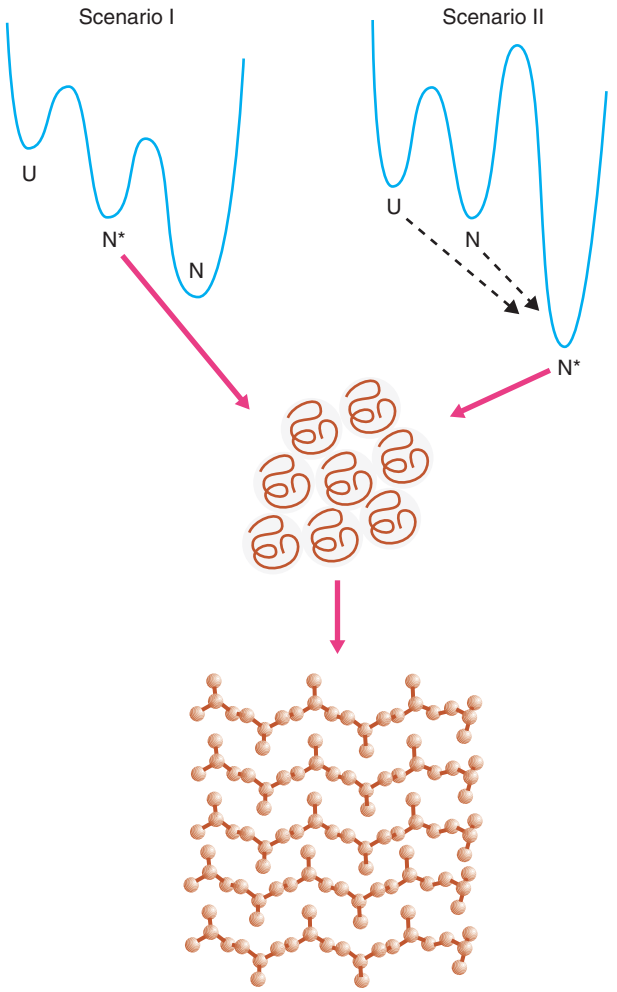
Ribbon diagram of the native structure of S6 (PDB code 1lou), which contains a single β sheet and two α helices. According to the OR[2] rule, β2, which is the edge strand (shown in blue), is protected against fibrillization by a combination of two mechanisms. The first one, based on electrostatic considerations, is enabled by the presence of two pairs of consecutive like charges (E41/E42 and R46/R47). Protection against aggregation in the second mechanism is afforded by the presence of a sharp twist and bend near the second pair of charged residues. Deletion of the pairs of charged residues by the double mutations E41A/E42A and R46M/R47M creates the mutant S6-Alz, which is prone to tetramerization. The figure was produced using MolMol [85].




Current Opinion in Structural Biology

**Figure 3**

The native dimeric assembly of TTR protein, which consists of β sheets CBEFF′E′B′C′ and DAGHH′G′A′D′. Experiments by Yeates and co-workers [58••] suggest that the first β sheet in the dimer turns into the elementary building block of the fibril by unfolding strands C and D (C′ and D′), and exposing the amyloidogenic strands B and B′. We also argue that the second β sheet (except D,D′) is likely to be preserved in fibrils because of the strong interactions within the HH′ interface. Protection of the strand B by short and twisted edge strands C and D is achieved by one of the aggregation-blocking mechanisms envisioned by R[2]. The figure was produced using MolMol [85].




Current Opinion in Structural Biology

**Figure 4**

Schematic diagram of the two plausible scenarios of fibrillization based on the free energy landscape perspective. According to scenario I, the assembly-competent state **N*** is metastable with respect to the monomeric native state **N** and is formed through partial unfolding. In scenario II, **N*** is formed upon structural conversion either of the native state **N** (as in prions) or directly from the unfolded state **U** (as in Aβ amyloid peptides). In both cases, proteins (or peptides) in **N*** states must coalescence into larger oligomers capable of growth into fibrils.