ELSEVIER

Protein aggregation and amyloidosis: confusion of the kinds?

Frederic Rousseau[1], Joost Schymkowitz[1] and Luis Serrano[2]

Recent years have witnessed major advances in our understanding of the structural basis of protein aggregation on several fronts. Firstly, high-resolution structural information that remained elusive for many years was provided by a series of studies of amyloid fibers using NMR, X-ray crystallography and electron microscopy, thereby confirming earlier models based on lower resolution observations. Secondly, studies of the sequence determinants of protein aggregation culminated in the development of computer algorithms that predict aggregation-prone sequences with good accuracy, allowing the design of mutations that reduce aggregation. Thirdly, based on the first results from such predictions and on statistical analysis of naturally occurring aggregating sequences, a picture is emerging in which aggregation-prone sequences are capped by gatekeeper residues that oppose aggregation. In addition to their aggregation-opposing function, it seems that gatekeeper residues are also important in determining chaperone selectivity for strongly aggregating regions. Finally, recent computational and experimental work shows that preventing aggregation does not necessarily mean that amyloid formation is prevented and *vice versa*. Thus, although aggregation and amyloidosis correlate to a certain extent, they are different processes and should be treated as such.

Addresses
[1] SWITCH Laboratory, Flemish Interuniversity Institute for Biotechnology, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussel, Belgium
[2] European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany

Corresponding author: Serrano, Luis (serrano@embl-heidelberg.d)

Current Opinion in Structural Biology 2006, **16**:118–126

This review comes from a themed issue on
Folding and binding
Edited by Mikael Oliveberg and Eugene I Shakhnovich

Available online 24th January 2006

0959-440X/$ – see front matter

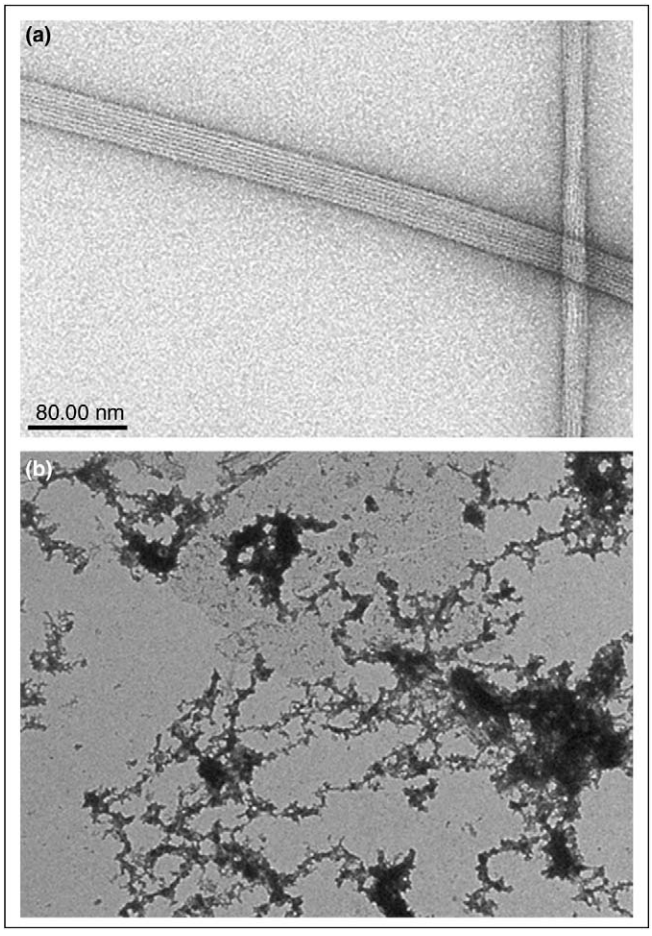
DOI 10.1016/j.sbi.2006.01.011

Introduction

Protein aggregation is essentially a self-association process in which many identical protein molecules form higher order conglomerates of low solubility that eventually precipitate. On the basis of their macroscopic morphology, they are generally classified as either ordered or disordered aggregates [1,2•]. Under physiological conditions, almost any protein can be induced at high concentration to form amorphous aggregates (Figure 1a); under the same conditions, a much smaller set of proteins form highly ordered β-rich amyloid fibers (Figure 1b) [1,3]. However, on a microscopic level, the differentiation between these two types of aggregates is more subtle. Amorphous aggregates are not just clusters of misfolded proteins that stick to each other through non-specific hydrophobic contacts. Rather, they are also often enriched in cross-β structure and their formation propensity correlates not only with hydrophobicity, but also with secondary structure propensity and charge, suggesting a specific mechanism of formation [4–6]. On the other hand, not all reported aggregates and fibers are enriched in β-structure, as both amorphous aggregates and fibers have been reported that retain native-like spectral properties and even enzymatic activity [7–11]. In these cases, aggregation is proposed to occur by other mechanisms of oligomerization, such as three-dimensional domain swapping [8,12–14]. The focus on protein aggregation is, for a large part, inspired by the observation that a range of human diseases are characterized by protein deposits composed of one or a very limited number of proteins [1]. β-Aggregation and amyloidosis often co-occur in these disease-associated protein aggregation processes and, when this is the case, the former is frequently observed as a precursor of the latter. It is now also generally accepted that a subgroup of these prefibrillar aggregates, not the mature fibers themselves, are associated with cytotoxicity [15,16]. The toxicity of prefibrillar aggregates is, among other factors, dependent on the size of these misfolded oligomers and seems to be determined by a universal mechanism, as toxic aggregates of very different proteins are recognized by a unique antibody [17,18]. The mechanism of toxicity of protein aggregates remains unclear, but accumulating evidence suggests that it is related to the interaction of protein aggregates and the cell membrane [19].

This article reviews recent advances in our understanding of structural and sequence determinants of protein aggregation and amyloidosis. Some proteins form fibers that are non-toxic and probably even functionally relevant, whereas other proteins form toxic aggregates without forming fibers [20,21]. Comparing the sequence determinants of both processes might therefore be of some relevance to understanding protein-misfolding-associated toxicity. Finally, as protein aggregation not only hampers protein folding but is also often toxic, the question arises as to whether proteins have evolved strategies to minimize aggregation. We review recent evidence of the existence of gatekeeper residues that fulfill such a role and how these residues could be linked to chaperone specificity.

Figure 1



Electron micrographs of the macromolecular morphology of protein aggregates and amyloid fibrils. **(a)** Amorphous aggregates (of the fusion proteins between the NusA protein and the Alzheimer's β-peptide) and **(b)** amyloid fibers (of the hexapeptide STVIIE [23]). Both aggregates have a high content of β-structure, but dramatically different morphologies when seen through an electron microscope.

The structure of amyloid fibrils

The cross-β model of the amyloid fiber structure was widely accepted, long before high-resolution structural data were available, as all lower resolution data from X-ray diffraction measurements, Fourier transform IR spectroscopy, circular dichroism and tinctorial assays (such as Congo red staining) indicated an enrichment in β-sheet structure [22••]. As the nucleation of amyloid fibers involves only a short segment of an amylogenic protein, fibers can be grown from peptides with the right amino acid sequence, a concept that was extensively demonstrated by Lopez-de la Paz *et al.* for the strongly amylogenic hexapeptide STVIIE [2•,23,24]. Peptide-based fibers offer the opportunity to uncouple questions regarding the structure of the amyloid backbone from the organization of the remainder of the protein.

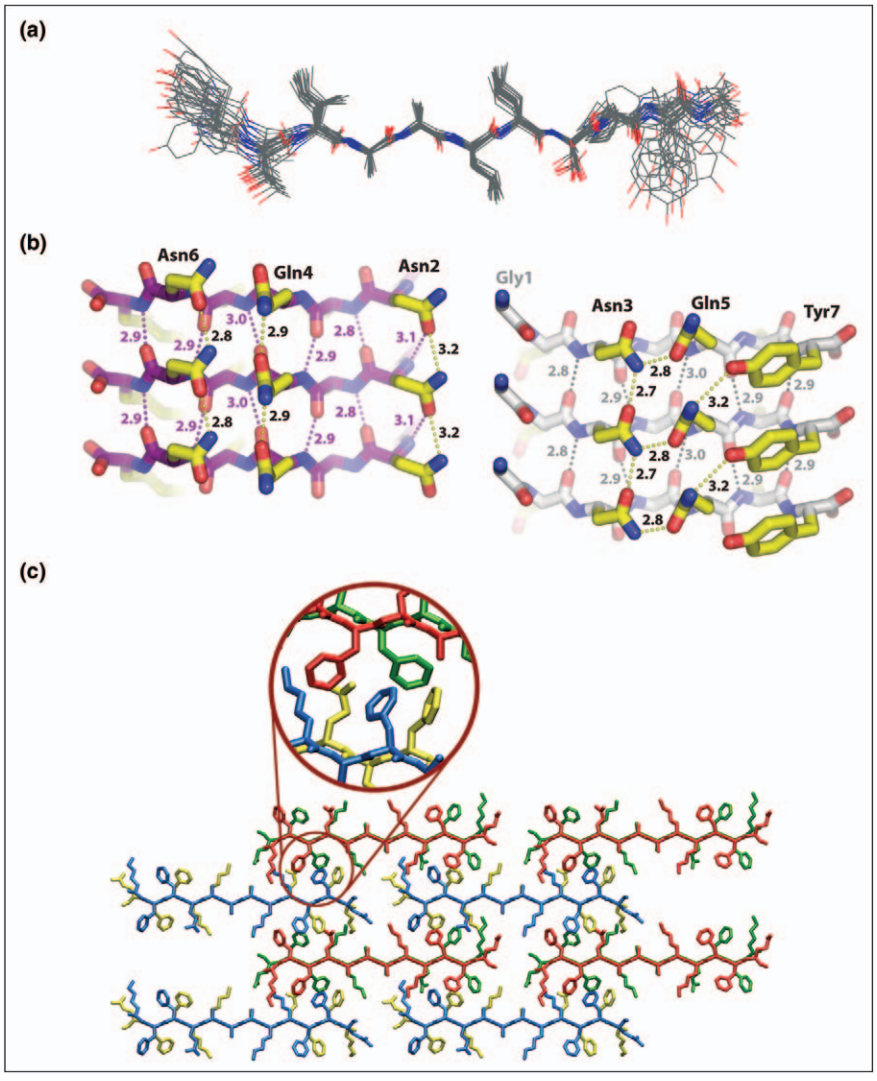
Recently, a series of high-resolution structural studies of amyloid fibers grown from small amylogenic peptides addressed the backbone structure of the fibril [22••,25••,26••]. In one of these pioneering studies, the Eisenberg laboratory used a peptide derived from the yeast prion protein Sup35 with the sequence GNNQQNY and a peptide with the related NNQQNY sequence to grow fibers that, with respect to all spectral and tinctorial properties and protease resistance, are highly similar to those grown from the full-length protein [26••] (Figure 2a). A similar approach was employed by Serpell and co-workers, this time using a designed 12-mer peptide with the sequence KFFEAAAKKFFE to form the fibers [25••] (Figure 2b). Finally, the high-resolution structure of a peptide fragment with the sequence YTIAALLSPYS from the amylogenic protein transthyretin was resolved in its fibrillar conformation using magic angle spinning NMR [22••] (Figure 2c). All three structures confirm that the peptides adopt an extended β-strand conformation in the amyloid fibrils. They also show that the fibrils achieve their exceptional stability and nearly crystalline order through near optimal values for both mainchain and sidechain dihedral angles, as well as through extensive packing and salt bridge formation between the peptides.

A possible extension to the single-strand amyloid backbone model is the incorporation of hairpin structures, in which each protein molecule contributes two strand regions, connected by a loop region [27••,28–31]. The NMR-derived three-dimensional structure of fibrils of the Alzheimer's β 1–42 peptide recently revealed such a structure, with residues 18–42 from each peptide adopting a β-turn-β motif that packs into two intermolecular parallel β-sheets that constitute the fibril [32••]. Several other plausible extensions and modulations of the cross-β theme have been proposed, including a four-sheeted model derived from X-ray diffraction data on fibrils of the B1 domain of protein G [33]. However, one of the few models to accommodate complete native structures in amyloid fibers is the zipper-spine model, proposed by Eisenberg *et al.* in 2001 based on two distinct domain-swapped structures of the RNase A molecule [13,14,34]. In this model, a cross-β spine is decorated with three-dimensional domain-swapped units that retain native-like structure and activity. The native-like fibers described by Melki and co-workers [7] could be consistent with such a model; recently, convincing evidence was presented by the Eisenberg laboratory using the RNase A model system, in which the hinge loop that mediates three-dimensional domain swapping was replaced by the amylogenic peptide from Sup35 discussed earlier [26••].

Sequence determinants of protein amyloid fibril formation

Crystal structures of amyloid fibrils generate a framework for understanding the sequence elements that predispose

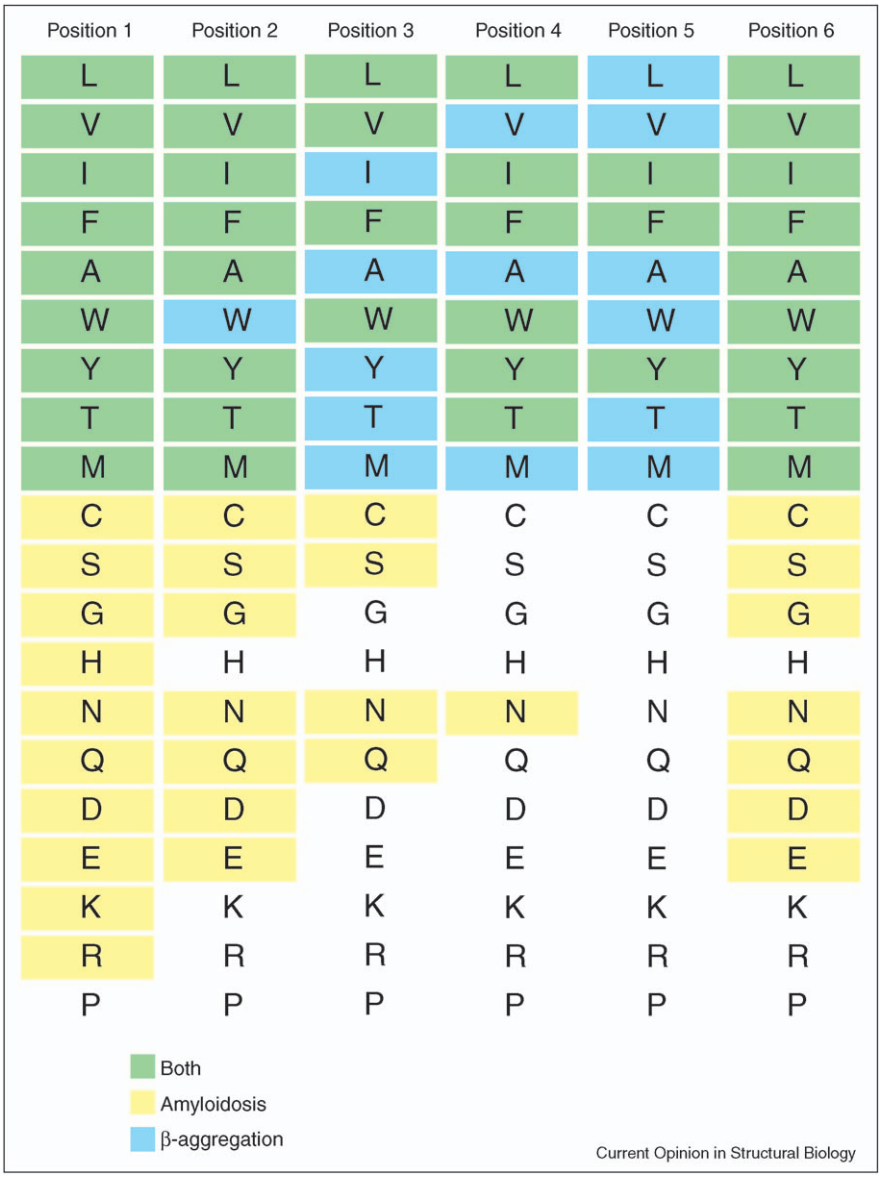
Figure 2



High-resolution structures of peptides in amyloid fibrils. **(a)** Structure derived from magic angle spinning NMR of fibrils grown from a peptide fragment of transthyretin [22••]. **(b)** Crystallographic model of a fibril comprising a peptide sequence derived from the yeast prion Sup35 [26••]. **(c)** High-resolution model from crystallography and electron microscopy of fibrils of a synthetic peptide sequence [25••].

a sequence to form fibers. Based on an antiparallel β-sheet consisting of six identical hexapeptides, sequences generated by the design algorithm PERLA were experimentally verified for amyloid fiber formation [23]. Not all sequences found by the design algorithm to be compatible with the antiparallel β-sheet backbone structure formed amyloid fibrils *in vitro*. Additional criteria, such as peptide charge, need to be taken into account to accurately predict amylogenesis. In a second study, Lopez de la Paz *et al.* [2•] performed an exhaustive experimental mutational scan of the most amylogenic peptide identified in the original study [23], which has the sequence STVIIE. Interestingly, only about 40% of the single point mutations of this strongly amylogenic peptide form amyloid fibrils, whereas the other peptides either stay in solution or form amorphous aggregates that are also rich in cross-β structure [2•]. Because this mutational study of single point mutations was performed on a single sequence only, it runs the risk of containing context-dependent biases, as well as covering insufficient sequence space. Nevertheless, the sequence pattern derived from this work captures a surprisingly large number of known amylogenic regions in a range of very different disease-associated and other amylogenic proteins [2•,28], including Alzheimer's β-peptide [35], human prion protein [27••] and yeast prion protein Sup35 [36]. Furthermore, peptide fragments from several disease-related proteins that were shown to form amyloid fibrils similar to those formed by the native protein correspond to regions predicted by the same sequence

Figure 3



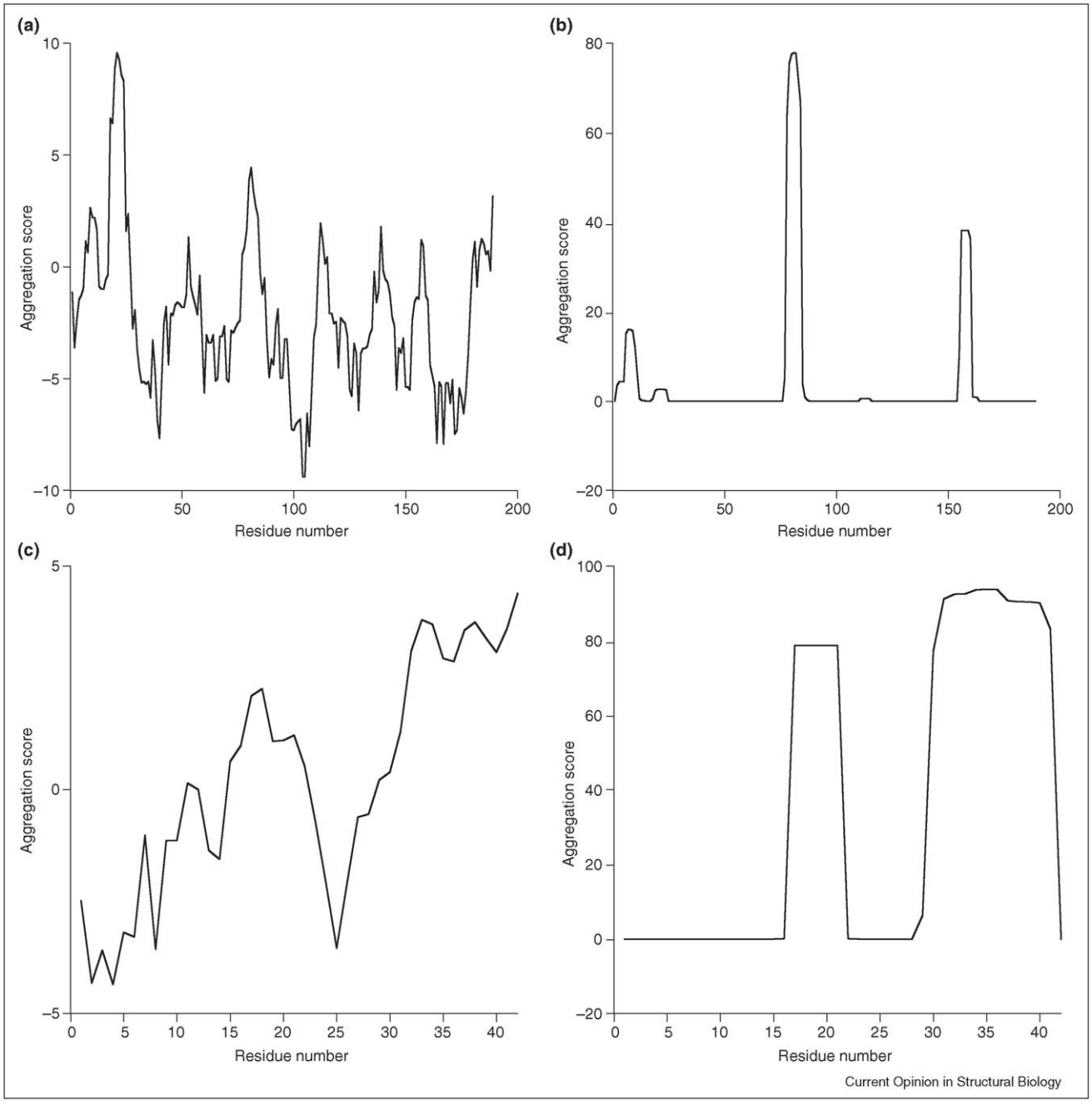
Overlap of sequence space for cross-β aggregation [37••] and amyloidosis [2•] of a hexapeptide. For the hexapeptide sequence, amino acid preferences are color coded: amino acids that are allowed at a given position in an amyloid fibril are shown in yellow, whereas amino acids allowed in amorphous cross-β aggregates are shown in blue. Residues that are compatible with both structures are shown in green.

pattern (L Serrano *et al.*, unpublished). This mutational study also highlights the thin line dividing sequences compatible with highly ordered cross-β amyloid structures and sequences that form amorphous cross-β aggregates (Figure 3).

Sequence determinants of protein aggregation

Mutational studies of the kinetics of aggregation of full-length proteins revealed simple correlations between aggregation and physico-chemical properties such as β-sheet propensity, hydrophobicity and charge [5]. This prompted the development of computer algorithms that identify aggregation-prone regions in the amino acid sequence of a protein [37••,38••] (Figure 4). Based on the correlations mentioned above, the Zyggregator algorithm of Dobson *et al.* [38••] identifies aggregation-prone sequences by comparing the aggregation propensity score of a given amino acid sequence with an average propensity calculated for a set of sequences of similar length. The statistical mechanics algorithm TANGO [37••], on the other hand, balances the same physico-chemical parameters, supplemented by the assumption that an amino acid is fully buried in the aggregated state: this

Figure 4



Example outputs of algorithms for the prediction of aggregation-prone sequences. **(a)** Zyggregator [38••] output for p21 Ras. **(b)** Tango [37••] output for p21 Ras. **(c)** Zyggregator [38••] output for Alzheimer's β-peptide. **(d)** Tango [37••] output for Alzheimer's β-peptide. Note that both algorithms predict the intrinsic aggregation tendency of protein sequences. Aggregation propensities can be further modulated by the stability of the native state.

means it becomes fully desolvated and entropically restricted. From an input sequence, TANGO generates an extensive sample of fragments for which competing structural propensities, such as helix or hairpin formation, are considered. All the fragments are then balanced in a global partition sum, which allows the identification of sequence regions that predominantly form aggregates. The TANGO algorithm has an accuracy of more than 90% for a set of 176 experimentally validated peptides. Importantly, both the Zyggregator algorithm and TANGO perform well for peptides and denatured proteins. For globular proteins, a partly folded molecule can

either refold to the native state or misfold into an aggregated state. As a result, both reactions are in competition and a precise understanding of the kinetics is essential to predict the final outcome in terms of folding or misfolding.

Similarities and differences between aggregation and amyloidosis

Comparing the sequence space of β-aggregation predicted by TANGO or Zyggregator with the sequence space of amyloidosis derived from experimental studies of the STVIIE amylogenic peptide reveals the similarities, but also interesting differences between both processes. Indeed, as both amyloid formation and amorphous cross-β aggregation require amino acid compositions that are compatible with a β-strand conformation, an overlap in sequence space is to be expected (Figure 3). However, the structure of amorphous cross-β aggregates is not clearly defined and seems to be characterized by a high degree of flexibility. On the other hand, the structure of amyloid fibers is quasi-crystalline. As a consequence, amino acid preferences will be much more position specific in an amyloid fiber than in amorphous cross-β aggregates. For the STVIIE sequence, for example, positions 3 and 4 are extremely selective, as only some amino acid types are compatible with a highly ordered amyloid structure. On the other hand, positions 1, 2 and 6 are much more tolerant, as almost any residue type allows amyloid formation [2•]. In contrast, almost any residue can be accommodated at any position of an hexapeptide for β-aggregation to occur, as long as the sequence as a whole has a good propensity to be in a β-extended conformation, and is sufficiently hydrophobic and/or neutral in charge [37••]. Amylogenic sequences are therefore more position specific, but also more tolerant to polar and charged residues than β-aggregating sequences. This will also have consequences on the kinetics of both processes. Due to its less stringent conformational requirements, β-aggregation is generally much faster than amyloidosis, although fast amyloidosis has also been observed [21]. As β-aggregates are often observed as precursors on the path to fiber formation, the stability of these precursor aggregates will strongly influence the kinetics of amyloidosis [39]. Stable β-aggregated amyloid precursors will therefore probably slow down amyloidosis. Polar amylogenic sequences, as observed in yeast prion proteins, will have a much lower β-aggregation propensity and will therefore be much more favorable for the kinetics of amyloidosis. In summary, amorphous cross-β aggregation and amyloidosis can occur in common, and the stability and kinetics of both processes will be determined by the extent to which the structural requirements of both processes are fulfilled. At one end of the spectrum, generally highly hydrophobic sequences have a strong tendency to form amorphous cross-β aggregates, but do not form fibers due to steric constraints. At the other end of the spectrum, generally more polar sequences will form amyloid fibers via precursor aggregates that are only marginally stable. Between the two, a whole spectrum of behaviors will probably be observed.

Evolutionary pressure against aggregation

Aggregating sequences are very common in globular proteins, and occur with about the same frequency in α, β, α+β and α/β proteins (SCOP classification [40]) [41]. These short aggregation-prone stretches are sufficient to induce aggregation of a protein. For example, transplanting a short six amino acid stretch from the amylogenic Src homology 3 (SH3) domain of phosphatidyl-inositol-3′-kinase provokes fiber formation in the otherwise non-amylogenic α-spectrin SH3 domain [42]. Similar results were obtained when grafting non-SH3 amylogenic sequences [43] and other proteins [44]. It can be considered that aggregation-sensitive protein sequences are the price to be paid for the existence of globular protein structures: as tertiary sidechain interactions mainly occur in the hydrophobic core, protein stretches spanning this region generally have a propensity to aggregate. Accordingly, intrinsically disordered proteins that lack tertiary structure are much less hydrophobic and thus have a much lower aggregation propensity [41]. However, for native globular proteins, aggregation is generally not an issue, as aggregation-prone protein stretches are generally sequestered by the protein structure and thereby protected from self-association [45]. On the other hand, during protein translation and folding, or in the case of cellular stress or destabilizing mutations, partially unfolded states are much more likely to self-associate and induce aggregation and amyloidosis [46–50].

The toxicity of protein aggregates is believed to be caused by the partitioning of these aggregates into the lipid phase of biological membranes [19,51,52]. Obviously, in a cellular context, this and other toxic effects associated with aggregation need to be tightly controlled and minimized where possible. Using the TANGO algorithm, we performed a study of the aggregation of 28 full proteomes [53]. From this study, it appeared that there is a clear evolutionary pressure to minimize the aggregation propensity of these sequences. First, although many proteins possess aggregation-prone sequences, most are of low aggregation propensity, with only a minority of high aggregation propensity. Second, as aggregation cannot be completely eliminated, because of structural constraints from the native state, aggregation is further contained by placing not only charged residues but also prolines and glycines at the flanks of aggregating sequence segments. These effectively act as gatekeeper residues, opposing aggregation and thereby promoting the native folding reaction [54•]. The existence of gatekeeper residues was proposed several years ago in a folding study of the S6 protein [54•,55–57], and they have since been observed in naturally occurring amylo-

genic and amorphous aggregating regions of several proteins [2•,58,59]. It has also been reported on several occasions that the introduction of charged residues, prolines or glycines in aggregation-prone sequences reduces aggregation [59,60]. The aggregation-opposing properties of proline and glycine originate primarily from their structure-breaking properties. Identically charged residues are also very effective at opposing aggregation, because of the huge repulsive force generated upon self-assembly. Interestingly, the statistical analysis performed using TANGO showed that arginine and lysine are preferred over glutamate and aspartate at the flanks of strongly aggregating sequences [53]. The reason for this preference might be that, in addition to charge, arginine and lysine also have much larger conformational entropy [61], making it very costly to immobilize them in densely packed aggregates. As most chaperones have a preference for hydrophobic sequences and positive charges [62–64], and disfavor negative charges, it is entirely possible that chaperones exploited the properties of gatekeeper residues to develop specificity for strongly aggregating hydrophobic sequences.

Conclusions

Recent years saw the elucidation of high-resolution amyloid fibrils. The detailed structure of amorphous aggregates remains beyond the limits of current experimental imaging techniques. However, for aggregation and amyloidosis, the accumulated structural and mechanistic knowledge allowed the discovery of sequence determinants for both processes. It is clear that, although both β-aggregation and amyloidosis are related phenomena, they still display significant differences in terms of both structural and biophysical properties. Comparing the sequence preferences of both processes should help to rationalize the mechanisms of toxicity of protein aggregation and amyloidosis. Furthermore, evidence is now accumulating that protein aggregation is a universal flaw in the folding landscape of globular proteins and that evolution enforces the use of gatekeeper residues to minimize misfolding as much as possible. Finally, the interesting possibility emerges that gatekeeper properties and chaperone specificity are directly linked.

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. Dobson CM: **Principles of protein folding, misfolding and aggregation**. *Semin Cell Dev Biol* 2004, **15**:3-16.

2. • Lopez de la Paz M, Serrano L: **Sequence determinants of amyloid fibril formation**. *Proc Natl Acad Sci USA* 2004, **101**:87-92.

This extensive mutational study of the hexapeptide STVIIE using electron microscopy reveals the sequence determinants of amyloid formation. Although the study is based on a single amylogenic sequence, the resulting sequence profile successfully identifies a range of known amylogenic regions in proteins associated with human disease.

3. Dobson CM: **Protein folding and disease: a view from the first Horizon Symposium**. *Nat Rev Drug Discov* 2003, **2**:154-160.

4. Chiti F, Calamai M, Taddei N, Stefani M, Ramponi G, Dobson CM: **Studies of the aggregation of mutant proteins *in vitro* provide insights into the genetics of amyloid diseases**. *Proc Natl Acad Sci USA* 2002, **99**:16419-16426.

5. Chiti F, Stefani M, Taddei N, Ramponi G, Dobson CM: **Rationalization of the effects of mutations on peptide and protein aggregation rates**. *Nature* 2003, **424**:805-808.

6. Chiti F, Taddei N, Baroni F, Capanni C, Stefani M, Ramponi G, Dobson CM: **Kinetic partitioning of protein folding and aggregation**. *Nat Struct Biol* 2002, **9**:137-143.

7. Bousset L, Thomson NH, Radford SE, Melki R: **The yeast prion Ure2p retains its native alpha-helical conformation upon assembly into protein fibrils *in vitro***. *EMBO J* 2002, **21**:2903-2911.

8. Rousseau F, Schymkowitz JWH, Wilkinson HR, Itzhaki LS: **Three-dimensional domain swapping in p13suc1 occurs in the unfolded state and is controlled by conserved proline residues**. *Proc Natl Acad Sci USA* 2001, **98**:5596-5601.

9. Soldi G, Bemporad F, Torrassa S, Relini A, Ramazzotti M, Taddei N, Chiti F: **Amyloid formation of a protein in the absence of unfolding and destabilisation of the native state**. *Biophys J* 2005, **89**:4234-4244.

10. Marcon G, Plakoutsi G, Canale C, Relini A, Taddei N, Dobson CM, Ramponi G, Chiti F: **Amyloid formation from HypF-N under conditions in which the protein is initially in its native state**. *J Mol Biol* 2005, **347**:323-335.

11. Ganesh C, Zaidi FN, Udgaonkar JB, Varadarajan R: **Reversible formation of on-pathway macroscopic aggregates during the folding of maltose binding protein**. *Protein Sci* 2001, **10**:1635-1644.

12. Silow M, Oliveberg M: **Transient aggregates in protein folding are easily mistaken for folding intermediates**. *Proc Natl Acad Sci USA* 1997, **94**:6084-6086.

13. Liu Y, Eisenberg D: **3D domain swapping: as domains continue to swap**. *Protein Sci* 2002, **11**:1285-1299.

14. Liu Y, Gotte G, Libonati M, Eisenberg D: **Structures of the two 3D domain-swapped RNase A trimers**. *Protein Sci* 2002, **11**:371-380.

15. 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.

16. Bucciantini M, Calloni G, Chiti F, Formigli L, Nosi D, Dobson CM, Stefani M: **Pre-fibrillar amyloid protein aggregates share common features of cytotoxicity**. *J Biol Chem* 2004, **279**:31374-31382.

17. Kayed R, Head E, Thompson JL, McIntire TM, Milton SC, Cotman CW, Glabe CG: **Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis**. *Science* 2003, **300**:486-489.

18. O'Nuallain B, Wetzel R: **Conformational Abs recognizing a generic amyloid fibril epitope**. *Proc Natl Acad Sci USA* 2002, **99**:1485-1490.

19. Anderluh G, Gutierrez-Aguirre I, Rabzelj S, Ceru S, Kopitar-Jerala N, Macek P, Turk V, Zerovnik E: **Interaction of human stefin B in the prefibrillar oligomeric form with membranes. Correlation with cellular toxicity**. *FEBS J* 2005, **272**:3042-3051.

20. Si K, Lindquist S, Kandel ER: **A neuronal isoform of the aplysia CPEB has prion-like properties**. *Cell* 2003, **115**:879-891.

21. Fowler DM, Koulov AV, Alory-Jost C, Marks MS, Balch WE, Kelly JW: **Functional amyloid formation within mammalian tissue**. *PLoS Biol* 2005, **4**:e6.

22. •• Jaroniec CP, MacPhee CE, Bajaj VS, McMahon MT, Dobson CM, Griffin RG: **High-resolution molecular structure of a peptide in an amyloid fibril determined by magic angle spinning NMR spectroscopy**. *Proc Natl Acad Sci USA* 2004, **101**:711-716.

Three groups [22••,25••,26••] obtained high-resolution structural information on peptide-derived amyloid fibers within a few months of each other, each confirming the established model of the cross-β spine of the amyloid fiber and providing atomic detail for the interactions. Taken together, the structures give insights into sidechain packing in the fiber structure.

23. Lopez De La Paz M, Goldie K, Zurdo J, Lacroix E, Dobson CM, Hoenger A, Serrano L: ***De novo* designed peptide-based amyloid fibrils**. *Proc Natl Acad Sci USA* 2002, **99**:16052-16057.

24. Lopez de la Paz M, Lacroix E, Ramirez-Alvarado M, Serrano L: **Computer-aided design of beta-sheet peptides**. *J Mol Biol* 2001, **312**:229-246.

25. •• Makin OS, Atkins E, Sikorski P, Johansson J, Serpell LC: **Molecular basis for amyloid fibril formation and stability**. *Proc Natl Acad Sci USA* 2005, **102**:315-320.
See annotation to [22••].

26. •• Nelson R, Sawaya MR, Balbirnie M, Madsen AO, Riekel C, Grothe R, Eisenberg D: **Structure of the cross-beta spine of amyloid-like fibrils**. *Nature* 2005, **435**:773-778.
See annotation to [22••].

27. •• Krishnan R, Lindquist SL: **Structural insights into a yeast prion illuminate nucleation and strain diversity**. *Nature* 2005, **435**:765-772.
This extensive study of the structure/function relationship of a 250 amino acid fragment of the yeast prion protein Sup35 identifies two 15 amino acid stretches that together form the key interactions that stabilize the prion form of the protein.

28. Ritter C, Maddelein ML, Siemer AB, Luhrs T, Ernst M, Meier BH, Saupe SJ, Riek R: **Correlation of structural elements and infectivity of the HET-s prion**. *Nature* 2005, **435**:844-848.

29. Daidone I, Simona F, Roccatano D, Broglia RA, Tiana G, Colombo G, Di Nola A: **Beta-hairpin conformation of fibrillogenic peptides: structure and alpha-beta transition mechanism revealed by molecular dynamics simulations**. *Proteins* 2004, **57**:198-204.

30. Aravinda S, Harini VV, Shamala N, Das C, Balaram P: **Structure and assembly of designed beta-hairpin peptides in crystals as models for beta-sheet aggregation**. *Biochemistry* 2004, **43**:1832-1846.

31. Torok M, Milton S, Kayed R, Wu P, McIntire T, Glabe CG, Langen R: **Structural and dynamic features of Alzheimer's Abeta peptide in amyloid fibrils studied by site-directed spin labeling**. *J Biol Chem* 2002, **277**:40810-40815.

32. •• Luehrs T, Ritter C, Adrian M, Riek-Loher D, Bohrmann B, Doebeli H, Schubert D, Riek R: **3D structure of Alzheimer's amyloid-beta(1-42) fibrils**. *Proc Natl Acad Sci USA* 2005, **102**:17342-17347.
The structure of Alzheimer's β-peptide in the fibril conformation was solved using a combination of NMR constraints derived from quenched hydrogen/deuterium exchange, pairwise mutagenesis studies and solid-state NMR. The structure shows that the peptides are stacked like hairpins to form the fibril; residues 18–26 and 31–42 adopt the β-strand conformation, and form intermolecular β-sheets with identical molecules.

33. Wang J, Gulich S, Bradford C, Ramirez-Alvarado M, Regan L: **A twisted four-sheeted model for an amyloid fibril**. *Structure (Camb)* 2005, **13**:1279-1288.

34. Ogihara NL, Ghirlanda G, Bryson JW, Gingery M, DeGrado WF, Eisenberg D: **Design of three-dimensional domain-swapped dimers and fibrous oligomers**. *Proc Natl Acad Sci USA* 2001, **98**:1404-1409.

35. Wurth C, Guimard NK, Hecht MH: **Mutations that reduce aggregation of the Alzheimer's Abeta42 peptide: an unbiased search for the sequence determinants of Abeta amyloidogenesis**. *J Mol Biol* 2002, **319**:1279-1290.

36. Balbirnie M, Grothe R, Eisenberg DS: **An amyloid-forming peptide from the yeast prion Sup35 reveals a dehydrated beta-sheet structure for amyloid**. *Proc Natl Acad Sci USA* 2001, **98**:2375-2380.

37. •• Fernandez-Escamilla AM, Rousseau F, Schymkowitz J, Serrano L: **Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins**. *Nat Biotechnol* 2004, **22**:1302-1306.
This paper introduces the TANGO algorithm for the prediction of the protein aggregation propensity of peptides and denatured proteins. This statistical mechanics algorithm balances several physico-chemical parameters, such as charge, β-sheet propensity and hydrophobicity, as well as competition from other structural states (e.g. α helix), to reach a prediction.

38. •• Pawar AP, Dubay KF, Zurdo J, Chiti F, Vendruscolo M, Dobson CM: **Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases**. *J Mol Biol* 2005, **350**:379-392.
The authors introduce the Zyggregator method for predicting protein aggregation. It is based on the physico-chemical profiling of amino acid sequence parameters, such as hydrophobicity, β-sheet propensity and charge.

39. Nguyen HD, Hall CK: **Kinetics of fibril formation by polyalanine peptides**. *J Biol Chem* 2005, **280**:9074-9082.

40. Lo Conte L, Ailey B, Hubbard TJ, Brenner SE, Murzin AG, Chothia C: **SCOP: a structural classification of proteins database**. *Nucleic Acids Res* 2000, **28**:257-259.

41. Linding R, Schymkowitz J, Rousseau F, Diella F, Serrano L: **A comparative study of the relationship between protein structure and beta-aggregation in globular and intrinsically disordered proteins**. *J Mol Biol* 2004, **342**:345-353.

42. Ventura S, Zurdo J, Narayanan S, Parreno M, Mangues R, Reif B, Chiti F, Giannoni E, Dobson CM, Aviles FX *et al.*: **Short amino acid stretches can mediate amyloid formation in globular proteins: the Src homology 3 (SH3) case**. *Proc Natl Acad Sci USA* 2004, **101**:7258-7263.

43. Esteras-Chopo A, Serrano L, Lopez de la Paz M: **The amyloid stretch hypothesis: recruiting proteins toward the dark side**. *Proc Natl Acad Sci USA* 2005, **102**:16672-16677.

44. Otzen DE, Miron S, Akke M, Oliveberg M: **Transient aggregation and stable dimerization induced by introducing an Alzheimer sequence into a water-soluble protein**. *Biochemistry* 2004, **43**:12964-12978.

45. Clark LA: **Protein aggregation determinants from a simplified model: cooperative folders resist aggregation**. *Protein Sci* 2005, **14**:653-662.

46. Johnson RJ, Christodoulou J, Dumoulin M, Caddy GL, Alcocer MJ, Murtagh GJ, Kumita JR, Larsson G, Robinson CV, Archer DB *et al.*: **Rationalising lysozyme amyloidosis: insights from the structure and solution dynamics of T70N lysozyme**. *J Mol Biol* 2005, **352**:823-836.

47. Wiseman RL, Green NS, Kelly JW: **Kinetic stabilization of an oligomeric protein under physiological conditions demonstrated by a lack of subunit exchange: implications for transthyretin amyloidosis**. *Biochemistry* 2005, **44**:9265-9274.

48. Foss TR, Kelker MS, Wiseman RL, Wilson IA, Kelly JW: **Kinetic stabilization of the native state by protein engineering: implications for inhibition of transthyretin amyloidogenesis**. *J Mol Biol* 2005, **347**:841-854.

49. Hurshman AR, White JT, Powers ET, Kelly JW: **Transthyretin aggregation under partially denaturing conditions is a downhill polymerization**. *Biochemistry* 2004, **43**:7365-7381.

50. Vernaglia BA, Huang J, Clark ED: **Guanidine hydrochloride can induce amyloid fibril formation from hen egg-white lysozyme**. *Biomacromolecules* 2004, **5**:1362-1370.

51. Jao CC, Der-Sarkissian A, Chen J, Langen R: **Structure of membrane-bound alpha-synuclein studied by site-directed spin labeling**. *Proc Natl Acad Sci USA* 2004, **101**:8331-8336.

52. Molander-Melin M, Blennow K, Bogdanovic N, Dellheden B, Mansson JE, Fredman P: **Structural membrane alterations in Alzheimer brains found to be associated with regional disease development; increased density of gangliosides GM1 and GM2 and loss of cholesterol in detergent-resistant membrane domains**. *J Neurochem* 2005, **92**:171-182.

53. Rousseau F, Serrano L, Schymkowitz JWH: **How evolutionary pressure against protein aggregation shaped chaperone specificity**. *J Mol Biol* 2006, **355**:1037-1047.

54. Pedersen JS, Christensen G, Otzen DE: **Modulation of S6 fibrillation by unfolding rates and gatekeeper residues**. *J Mol Biol* 2004, **341**:575-588.
An experimental study of the role of gatekeeper residues in the folding and misfolding of the S6 protein. This work is an experimental demonstration of the importance of gatekeeper residues to prevent protein misfolding.

55. Mogensen JE, Ipsen H, Holm J, Otzen DE: **Elimination of a misfolded folding intermediate by a single point mutation**. *Biochemistry* 2004, **43**:3357-3367.

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

57. 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.

58. Parrini C, Taddei N, Ramazzotti M, Degl'Innocenti D, Ramponi G, Dobson CM, Chiti F: **Glycine residues appear to be evolutionarily conserved for their ability to inhibit aggregation**. *Structure (Camb)* 2005, **13**:1143-1151.

59. Fowler SB, Poon S, Muff R, Chiti F, Dobson CM, Zurdo J: **Rational design of aggregation-resistant bioactive peptides: reengineering human calcitonin**. *Proc Natl Acad Sci USA* 2005, **102**:10105-10110.

60. Calloni G, Zoffoli S, Stefani M, Dobson CM, Chiti F: **Investigating the effects of mutations on protein aggregation in the cell**. *J Biol Chem* 2005, **280**:10607-10613.

61. Abagyan R, Totrov M: **Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins**. *J Mol Biol* 1994, **235**:983-1002.

62. Rudiger S, Germeroth L, Schneider-Mergener J, Bukau B: **Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries**. *EMBO J* 1997, **16**:1501-1507.

63. Patzelt H, Rudiger S, Brehmer D, Kramer G, Vorderwulbecke S, Schaffitzel E, Waitz A, Hesterkamp T, Dong L, Schneider-Mergener J *et al.*: **Binding specificity of *Escherichia coli* trigger factor**. *Proc Natl Acad Sci USA* 2001, **98**:14244-14249.

64. Knoblauch NT, Rudiger S, Schonfeld HJ, Driessen AJ, Schneider-Mergener J, Bukau B: **Substrate specificity of the SecB chaperone**. *J Biol Chem* 1999, **274**:34219-34225.