Domain swapping and amyloid fibril conformation.

Patrick C.A. van der Wel

Department of Structural Biology, University of Pittsburgh School of Medicine,
3501 Fifth Ave, BST3 room 2044, Pittsburgh PA 15260
Keywords
Amyloid fibrils; Domain swapping; Protein aggregation;
Electron paramagnetic resonance; β₂-microglobulin; solid-state NMR;
magic angle spinning NMR

Abbreviations
β2m: β₂-microglobulin
EPR: electron paramagnetic resonance
GB1: immunoglobulin binding domain B1 of streptococcal protein G
IP: in-register parallel
MAS: magic angle spinning
Abstract

For several different proteins an apparent correlation has been observed between the propensity for dimerization by domain-swapping and the ability to aggregate into amyloid-like fibrils. Examples include the disease-related β2-microglobulin and transthyretin proteins. This has led to proposals that the amyloid-formation pathway may feature extensive domain swapping. One possible consequence of such an aggregation pathway is that the resulting fibrils would incorporate structural elements that resemble the domain-swapped forms of the protein, and thus reflect certain native-like structures or domain-interactions. In magic angle spinning solid-state NMR-based and other structural studies of such amyloid fibrils, it appears that multiple of these proteins form fibrils that are not native-like. Several fibrils instead have an in-register, parallel conformation, which is a common amyloid structural motif and is seen for instance in various prion fibrils. Such a lack of native structure in the fibrils suggests that the apparent connection between domain-swapping ability and amyloid-formation may be more subtle or complex than may be presumed at first glance.
We have recently reported\textsuperscript{1} on our magic-angle-spinning (MAS) solid-state NMR studies of the amyloid-like fibrils formed by a mutant form of the immunoglobulin binding domain B1 of streptococcal protein G (GB1). GB1 is a remarkably stable, small protein that has served as a model system for protein folding studies and has developed into a widely used test protein in both solution and solid-state NMR spectroscopy. One of its mutants has been found to form dimers in which a normally intra-molecular β-sheet, consisting of four β-strands β1-β4, becomes shared between the two proteins assembling into the dimer (see figure 1)\textsuperscript{2}. This involves a domain-swapping event with the domains each consisting of a β-hairpin\textsuperscript{3}. Interestingly, this same mutant was found to form amyloid-like fibrils, which the monomeric unmutated variant does not appear to do\textsuperscript{4}. Based on evaluations of related mutants, including the examination of specific disulfide bond formation\textsuperscript{4}, it appeared that there was a specific correlation between domain-swapping and amyloid-forming propensity.

Domain-swapping and amyloid formation

Such observations are of interest, as there have been indications that such a domain-swapping mechanism may represent a more general aspect of amyloid formation\textsuperscript{5, 6}. In particular, a domain-swapping based mechanism has been invoked for a number of amyloid-forming proteins - including some disease-related cases. The most prominent example has been the protein cystatin C, which causes cerebral amyloidosis\textsuperscript{7-11}. Other cases where similar mechanisms are invoked are in the aggregation of transthyretin\textsuperscript{12} and β\textsubscript{2}-microglobulin (β2m)\textsuperscript{13-15}, as well as various other amyloid-forming proteins\textsuperscript{2, 4, 16-19}.

Characterization of the molecular pathway leading to amyloid formation, and thus the identification of a potentially shared mechanism based on domain-swapping, is of
interest for various reasons. Although there is some debate over the (most) toxic species in the various amyloid-related diseases, it is clear that protein aggregation (i.e., amyloid fibril formation) is by definition the most prominent shared feature of all of these disorders. There are strong indications that there is a correlation between disease (onset) and the aggregation of the proteins involved, even if the specific type of aggregate that matters most may an issue of debate. Thus, if we want to understand how these diseases are triggered, we need to probe not just the nature of the mature fibrils, but also the pathway leading to fibril formation and the various steps along the way.

In various cases there are strong indications that the toxicity may be attributed to an oligomeric assembly of the protein involved, rather than the mature fibrils\textsuperscript{20}. These toxic oligomers are transient intermediates, and can thus be difficult to study directly. Nonetheless, there is a need to understand their structure and formation. Identifying and understanding a domain-swapping based pathway of aggregation would provide important clues about the likely or potential structures of on-pathway early aggregates. Already, studies of domain-swapping amyloidogenic protein has suggested that the domain-swapped oligomeric species could be toxic\textsuperscript{19}. In addition, a detailed structural analysis can yield a better appreciation of the role of native-like interactions or conformations, and how they may steer or initiate the aggregation process.

**Structural aspects**

It is by now well established that the central core of amyloid fibrils consists of long \(\beta\)-sheets that involve \(\beta\)-strands from numerous protein monomers. These intermolecular \(\beta\)-sheets can be arranged in different ways, but always feature extensive hydrogen bonding networks between the different peptide backbones. If we examine some of the
most commonly used case studies of domain-swapped dimers that are related to amyloid formation (e.g. Figure 2), then we see that in many of these dimers β-strands engage in antiparallel interactions with strands in other proteins, across the inter-domain (and inter-monomer) interface. Often these interactions in the domain-swapped dimer are intermolecular versions of normally intra-molecular β-sheets present in the monomeric native structure. This is for instance the case in the proposed structure for GB1 amyloid fibrils⁴, but also for many of the other proteins for which domain-swapping has been proposed to mediate their amyloid formation. Figure 2 shows a number of different dimeric structures for such proteins, based on solution NMR or X-ray crystallographic studies⁸, ¹³, ¹⁴, ²¹, ²².

Structurally it would seem reasonable to extrapolate the domain swapping beyond a dimeric assembly, as the observed β-sheet-based interactions could be accommodated in even longer β-sheets involving more proteins. This is shown schematically in Figure 1c. Such a model would predict that the amyloid core is formed from extensively inter-molecular β-sheets that contain the same native-like β-strands that are observed in the domain-swapped dimers, and would likely incorporate antiparallel β-strand interactions and β-hairpins (in analogy to the dimeric structures in Figure 2).

Clearly, this would present a seemingly elegant model for the formation of amyloid β-sheet structures, and explain the apparent correlations between domain swapping propensity and amyloid formation. Indeed, studies have reported various biochemical data that appear to be consistent with this general concept. In a number of different studies, selectively introduced disulfide bonds have been used to probe and control the interactions between domains and to observe the effect on amyloid formation as well as domain-swapping. Along with data suggesting the presence of native-like structures in
the fibrils or at least the early aggregates, these experiments appeared to support the formation of ‘domain-swapped fibrils’ that preserve the structural features typical of the dimeric forms and the native protein. Evidence of native-like structure has taken a variety of forms. Non-fibrillar toxic aggregates of p13suc1 were found to resemble the native form of the protein in terms of similar fluorescence and CD spectra, hydrogen/deuterium exchange characteristics, and protease sensitivity (monitored by SDS PAGE and mass spectrometry)\(^\text{19}\). In addition, there was a correlation between the amount of aggregation and the domain-swapping propensity of different mutants\(^\text{19}\). Activity-based evidence for domain-swapped mature fibrils was obtained in the elegant work by Sambashivan et al\(^\text{17}\), where fibrils were prepared from mixtures of individually inactive RNase A variants featuring inserts based on amyloidogenic polyQ or GNNQQQNY elements. The resulting fibrils showed enzymatic activity that required the swapping of domains from different mutants, thus revealing a native-like, but domain-swapped structure of functional domains, presumably with a ‘fibril backbone’ of $\beta$-sheets formed by the above-mentioned ‘inserts’\(^\text{17}\). Another approach for probing domain-swapped fibrils is through the incorporation of disulfide bonds within or between domains\(^\text{4, 10, 18, 23}\).

### Evidence for non-native mature fibrils

One crucial question in an aim to further validate this elegant and tempting mechanism is the structural characterization of the resulting fibrils on an atomic or molecular level. Two spectroscopic techniques that can provide such information are electron paramagnetic resonance (EPR) and solid-state NMR spectroscopy. Employing spin-labeled proteins or peptides, EPR can provide constraints on the fibril structure, indicating intra- but also (and importantly) inter-molecular proximities of specific residues\(^\text{24}\). Similarly, magic-angle-spinning solid state NMR provides structural data reflecting
the intramolecular and supramolecular assembly of amyloid fibrils. Gradually such studies are expanding our knowledge of common structural motifs in the fibrils. In addition, such methods have allowed the direct examination the structural similarities between fibrillar and crystalline aggregates of amyloid-forming peptide fragments. Analogously, these methods can provide the necessary data to detect the presence or absence of native-like structure in the mature fibrils formed by domain-swapping amyloidogenic proteins, in order to test the type of structural model as illustrated in Figure 1c.

However, our MAS NMR studies of the mature fibrils of the GB1 model protein failed to reveal the anticipated native-like conformations. No evidence was found for the native \(\alpha\)-helix, nor for the intramolecular \(\beta\)-hairpins that are present in both the native monomeric conformation and in the domain-swapped dimers. Instead, the amyloid fibrils are highly rigid, predominantly \(\beta\)-sheet in structure, and adopt an in-register parallel (IP) conformation. An examination of the literature also revealed evidence for such IP assemblies for various of the proteins proposed to undergo domain-swapping as part of the amyloid-formation process, based on MAS NMR and EPR. For example, studies of fibrils of \(\beta\)2m by both EPR and MAS NMR indicated a non-native IP structure. Similar studies of the human prion protein, which has also been implicated in a domain-swapping associated aggregation pathway, have come to similar conclusions.

Some experimental studies of transthyretin fibrils have supported a non-native, possibly IP \(\beta\)-sheet assembly. On the other hand, a recent NMR study yields structural constraints for transthyretin fibrils that indicate a fibril core structure that is neither native-like nor parallel in-register, and suggests that this may reflect a polymorphic form occurs alongside a more native-like fibril conformation.
It is perhaps worth pointing out here that this parallel in-register conformation reflects a commonly recurring motif in amyloid fibril structure \(^{24, 37}\). This structural motif has recently been referred to as a \(\beta\)-arcade \(^{37}\), and is illustrated in Figure 3. This type of assembly has been found in many different amyloid fibrils, including for instance amyloid-\(\beta\) and various yeast prion proteins. It aligns the peptide backbones such that they can participate in the hydrogen bonding networks that are thought to provide much of the stability typical of most amyloid fibrils\(^{38}\). It also results in a natural alignment of the same residue position in different monomers, within the fibrils, thus permitting stabilizing interactions via side chains of many residues: e.g. by clustering hydrophobic residues, the generation of hydrogen-bonded Gln/Asn ladders and through favorable interactions among aromatic residues. Charged residues may necessitate compensating counter-charges nearby or may be predominantly present in loop regions that intersperse the \(\beta\)-strands that assemble the amyloid core.

One consequence of the characteristic same-same interactions in these IP assemblies is that there is a class of experiments that can specifically probe such interactions, as illustrated in Figure 3. For MAS NMR this includes a variety of methods, typically relying on the specific isotopic \((^{13}\text{C})\) labeling of individual residues or even atoms, and detecting the proximity (or lack thereof\(^{25}\)) of the same labels in the corresponding residues in neighboring monomers\(^{26}\). There are similar approaches using site-specific spin labels in EPR\(^{24}\). These types of interactions are much less likely to be seen in structures where \(\beta\)-strands are not ‘stacked’ in a parallel in-register manner (see Figure 3). Note that all the domain-swapped dimers shown in Figure 2 feature an antiparallel alignment of the \(\beta\)-strands that form the dimer-interface. Furthermore, based on some of the observed domain-swapping, one may expect the inclusion of intramolecular \(\beta\)-hairpins within the fibrillar \(\beta\)-sheets, which would also prevent the same close proximity of same-to-same
residues in different monomers (Fig. 3a). Aside from these spectroscopic approaches, similar conclusions may be drawn on the basis of e.g. intermolecular disulfide bond formation\textsuperscript{23}.

**Role for domain-swapping?**

Based on such studies of the associated fibrils, it appears that at the least for a number of domain-swapping proteins the mature fibrils do not feature the predicted (e.g. Figure 1) native-like structural elements, but rather adopt the seemingly common $\beta$-arcade amyloid motif with IP $\beta$-sheets\textsuperscript{37} (or other non-native structure\textsuperscript{35}. What then could be the role of domain-swapping in amyloid formation of these various proteins?

One intriguing possibility is that the domain-swapping-related structural rearrangements play critical roles in the early stages of the aggregation process. It has been noted that the early aggregation states do feature native-like structures for e.g. $\beta_2m$ and the GB1 amyloids\textsuperscript{4, 39}. It is quite possible for the initial aggregates and the mature fibrils to feature different conformations, as these reflect kinetically as well as structurally distinct events along the aggregation pathway. Structural studies of the different intermediate states along the pathway are challenging due to their transient nature. There are studies that suggest a role for transient $\alpha$-helical structure in early oligomeric aggregates, which is lost upon formation of fibrils, in particular for inherently disordered proteins or peptides\textsuperscript{40-42}. In contrast, MAS NMR studies of A$\beta$ oligomers seems to show a highly $\beta$-sheet rich assembly that is relatively similar to the secondary sheet elements in the mature fibrils\textsuperscript{43}.

One thing that differs for the domain-swapping proteins compared to these examples is that they do have an ordered native conformation. This implies that a partial destabilization of the ordered native conformation has to be achieved in order for
amyloid formation to take place. Similarly, a partial unfolding must accompany the domain-swapping into the dimeric assemblies (Figure 4b). Thus, the domain-swapping propensity may play a role at the step of destabilizing the native state, and possibly in the subsequent first steps of the aggregation process.

If these initial steps result in a somewhat native-like oligomeric species that indeed shares structural features with the domain-swapped dimers (ref. figure 4), this could account for the experimental data indicating native structure at this stage of aggregation. Note that the toxic, non-fibrillar aggregates of p13suc1 (see above) seem to match such a description\textsuperscript{19}. Native-like early aggregates may well lead to ‘domain-swapped’ fibrils\textsuperscript{16}, in a way that is consistent with a templated elongation process (Figure 4c)\textsuperscript{44}. However, as delineated above, much of the structural fibril data appear thus consistent with the generation of structurally different amyloid fibrils (Figure 4d) that are parallel, in-register and thus reflect a loss of native structure – as we reported for the GB1 amyloid fibrils\textsuperscript{1}. A significant structural rearrangement between the oligomeric and fibrillar states is not necessarily unexpected\textsuperscript{37, 41}, although unambiguous evidence for this is still lacking. Given the potential involvement of this mechanism in a number of different amyloid-related disorders and the interest in the oligomer structure due to the potential toxic role of these species, further studies to examine this issue are clearly warranted. As with the characterization of transient oligomeric species in other amyloid diseases, such studies will pose some practical challenges but should be of great value to elucidate the aggregation process and possibly the nature of the toxic species.
Acknowledgments

The author acknowledges the members of the Van der Wel group, useful discussions with Ronald Wetzel, and the financial support by start-up funds from the University of Pittsburgh.
References


34. Blake C, Serpell L. Synchrotron X-ray studies suggest that the core of the transthyretin amyloid fibril is a continuous beta-sheet helix. Structure 1996; 4:989-98.


**Figure Legends**

Figure 1 – (a) Structure of monomeric GB1 (PDB ID 2JSV)\(^4\), along with schematic cartoon showing the β-hairpin ‘domains’. (b) domain-swapped dimer of mutant GB1 (PDB 1Q10)\(^2\), with a color-coded cartoon illustrating the domain swapped β-hairpin ‘domains’ (c) Highly schematic representation of a hypothetical assembly pattern for amyloid fibrils, based on extensive, intermolecular domain swapping\(^4\).

Figure 2 – Structures of domain-swapped dimers associated with amyloid formation. (a) bovine pancreatic RNase A (PDB 1A2W)\(^{21}\), (b) p13suc1 tetramer (1SCE)\(^{22}\), (c) human cystatin C (1T1J)\(^8\), (d) antibody bound tetramer of β2m (2X89)\(^{13}\), and (e) β2m dimer (3LOW)\(^{14}\). Individual monomers are color-coded. In each of these assemblies the interface is largely mediated by hydrogen bonding of antiparallel β-strands from two different monomers.

Figure 3 – Domain-swapped interfaces are commonly anti-parallel (a) and sometimes involved intra-sheet β-hairpins (b). This contrasts to the in-register parallel conformation (c), as experimentally observed for several of the actual fibrils. These measurements often include the detected of the same-to-same contacts that characterize the IP structure, as illustrated in (d) based on an IP peptide amyloid conformation\(^{28,46}\). Black spheres illustrate the varying positioning and proximity of identical labeled sites in different monomers.
Figure 4 – Schematic illustration of domain-swapping mechanisms in amyloid formation.
(a) The formation of domain-swapped dimers follows the flexing and partial unfolding of different domains relative to each other. Partial unfolding is also a pre-requisite for amyloid formation and may yield native-like oligomers with structural features that relate to ‘run-away domain-swapping’ \(^5,18\). (b) This may lead to the formation of amyloid fibrils that have native-like structure. However, various experimentally characterized amyloid fibrils do not show native-like structure (see text). (c) In those cases, a somewhat native-like oligomeric species may have undergone significant structural rearrangements, to yield a more canonical amyloid fibril lacking native structure, featuring for instance in-register parallel \(\beta\)-sheets.