AmyPro: a database of proteins with validated amyloidogenic regions

Mihaly Varadi[1,*], Greet De Baets[2], Wim F. Vranken[3,4,5], Peter Tompa[3,5,6] and Rita Pancsa[2,*]

[1]Protein Data Bank in Europe, European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Hinxton, Cambridge CB10 1SD, UK, [2]MRC Laboratory of Molecular Biology, Cambridge Biomedical Campus, Cambridge CB2 0QH, UK, [3]Structural Biology Brussels, Vrije Universiteit Brussel (VUB), Brussels, 1050, Belgium, [4]Interuniversity Institute of Bioinformatics in Brussels (IB)[2], ULB-VUB, Brussels, 1050, Belgium, [5]VIB Center for Structural Biology, Vrije Universiteit Brussel (VUB), Brussels, 1050, Belgium and [6]Institute of Enzymology, Research Centre for Natural Sciences of the HAS, Budapest, 1117, Hungary

ABSTRACT

Soluble functional proteins may transform into insoluble amyloid fibrils that deposit in a variety of tissues. Amyloid formation is a hallmark of age-related degenerative disorders. Perhaps surprisingly, amyloid fibrils can also be beneficial and are frequently exploited for diverse functional roles in organisms. Here we introduce AmyPro, an open-access database providing a comprehensive, carefully curated collection of validated amyloid fibril-forming proteins from all kingdoms of life classified into broad functional categories (http://amypro.net). In particular, AmyPro provides the boundaries of experimentally validated amyloidogenic sequence regions, short descriptions of the functional relevance of the proteins and their amyloid state, a list of the experimental techniques applied to study the amyloid state, important structural/functional/variation/mutation data transferred from UniProt, a list of relevant PDB structures categorized according to protein states, database cross-references and literature references. AmyPro greatly improves on similar currently available resources by incorporating both prions and functional amyloids in addition to pathogenic amyloids, and allows users to screen their sequences against the entire collection of validated amyloidogenic sequence fragments. By enabling further elucidation of the sequential determinants of amyloid fibril formation, we hope AmyPro will enhance the development of new methods for the precise prediction of amyloidogenic regions within proteins.

INTRODUCTION

Proteins are crucial macromolecules mediating most cellular functions and the majority of them have well-defined folded structures to serve these roles. If they fail to adopt or to remain in their native structural states and misfold, they are prone to aggregate and can form amyloid fibrils that are associated with a number of degenerative human disorders such as type II diabetes, Huntington's and Alzheimer's diseases and systemic amyloidosis ([1–4]). Amyloids are non-covalent fibrillar structures of extended, inter-molecularly hydrogen bonded β-sheets that laterally self-assemble to yield insoluble, twisted fibrils typically of 10 nm in diameter, in which they are oriented perpendicular to the fibril axis. Due to this arrangement, the amyloid fold is often referred to as a cross β-sheet structure ([5,6]). Multiple recent lines of evidence suggest that the amyloid state can also be beneficial for different functional purposes ([7,8]), such as bacterial and fungal biofilm formation ([9–12]), antimicrobial activity ([13,14]), the storage of peptide hormones ([15]) and the formation of the *zona pellucida* for the protection of mammalian and fish oocytes ([16,17]), among others. Besides these functional benefits, another reason why aggregation and amyloid fibril formation are maintained during evolution is that the regions involved are often part of the hydrophobic core of these proteins, and are thus required for them to fold correctly ([18,19]). Although amyloids play a key role in numerous important human diseases, the molecular mechanisms, the effects of mutations and changes in cellular milieu, the kinetics and the potential functional-evolutionary driving forces of amyloid fibril formation are only sparsely understood.

The propensity for amyloid formation is encoded in the protein sequence, but a comprehensive, curated and regularly updated database that stores known amyloid precursor proteins and their amyloidogenic regions is still lacking. The AMYPdb database ([20]) which aimed to fulfil this

*To whom correspondence should be addressed. Tel: +44 1223 267 823; Fax: +44 1223 213 556; Email: ritapancsa@gmail.com
Correspondence may also be addressed to Mihaly Varadi. Tel: +44 1223 494 278; Fax: +44 1223 494 468; Email: mvaradi@ebi.ac.uk

role, contains only 33 amyloid precursor proteins forming mainly pathogenic amyloids in human, and was not updated since 2008. Other databases store information on the amyloid fibril-forming capacities of short peptides, for instance WaltzDB (21), while AmyLoad contains only a few amyloidogenic protein fragments (22). The recently published Curated Protein Aggregation Database (CPAD, (23)) is an integrated resource of aggregation data, but only stores the 33 amyloid precursor proteins that are also contained by the out-of-date AMYPdb, without any protein structure or amyloid classification information.

In order to offer a more comprehensive and progressive online resource, we have collected proteins that have been demonstrated to form either pathogenic or functional amyloid fibrils or prions, with special focus on their amyloidogenic sequence regions through manually curated in-depth literature mining. This functionally classified set of amyloid-forming proteins is made available in AmyPro and offers a significantly more detailed view than its peers, as it already contains 143 entries as of September 2017, out of which for 127 the amyloidogenic regions or prion domains are also defined. Our data collection pipeline enables continuous updates of the database through dedicated searches of the recent literature, in addition to inviting the scientific community to contribute to the project by submitting their published data via our user-friendly submission interface. AmyPro provides comprehensive information on the entry sequences, including the precursor proteins, experimentally validated amyloidogenic sequence regions, a short description of the functional relevance of their soluble and amyloid states, a list of the experimental techniques used to discover and investigate the amyloid state, structural/functional/variation/mutation annotations obtained from UniProt (24) and relevant, classified structures linking to the PDBe/RCSB/PDBj entry pages of the entries, with representative structures also displayed as static pictures for the fibrillar and/or soluble states with amyloidogenic regions highlighted in color. Additionally, AmyPro offers a sequence screening service allowing the users to screen their sequences of interest for matching fragments with the validated amyloidogenic regions stored in the database.

MATERIALS AND METHODS

Data collection and classification

The initial dataset incorporated the curated data of 33 amyloid precursor proteins from the Supplementary Data of Tsolis *et al*. (25). For many of these, e.g. calcitonin, apolipoprotein A-I, transthyretin and cystatin C, the annotated amyloidogenic regions were also updated based on recently published analyses (26–30). We then collected other proteins from the literature for which amyloid fibril formation was experimentally demonstrated, with the only exception of different variants of antibody light chains that are numerous and are already available in ALBase maintained by Boston University (http://albase.bumc.bu.edu/aldb/). The amyloidogenic sequence regions were confirmed for most of the proteins, although in some cases this information was missing or the proposed region was only defined based on predictions. Such predictions were deemed to be

of lower reliability, and thus for these cases the amyloidogenic regions were not defined on a residue level in AmyPro. Sixteen entries are for proteins that are known to be amyloidogenic, but their amyloid forming regions are not yet defined. It is also important to note that for amyloid-forming proteins and peptides that are <45 residues in length, for instance peptide hormones (15) and antimicrobial peptides (13,14), we did not necessarily require definition of an amyloidogenic region, but accepted the entire sequence as amyloidogenic. In the case of prions, we accepted full prion domains as amyloidogenic regions and also indicated shorter segments of the domain if those were explicitly defined as amyloidogenic core regions. AmyPro provides a flag to easily identify entries with prion domains in order to help users and method developers to distinguish them from better resolved amyloidogenic regions, since those display distinctly different length distributions and amino acid compositions (see the Stats page of AmyPro).

The database stores the protein sequences investigated in the experiments addressing amyloid formation, which often lack initiator methionines, signal peptides and propeptides, and may comprise only a single domain compared to the corresponding UniProt entries (24). UniProt residue boundaries and information on the corresponding precursor protein are indicated. In a few cases where the wild-type protein does not form amyloid, but a mutant form does (31–34), we included the mutated sequence into the database and clearly indicated the change compared to the corresponding UniProt entry sequence. The boundaries of amyloidogenic regions are provided with regard to both the displayed protein sequences and the corresponding UniProt entries. The collected amyloidogenic proteins were screened against the Protein Data Bank (35) to obtain available structural information on them. Similarly, they were mapped onto UniProt (24), from where the relevant sections of feature tables were retrieved, including e.g. variation data, providing detailed information on relevant disease mutations.

The collected proteins were classified into five broad categories based on the proposed functions of their amyloid states in the literature: (i) pathogenic amyloids that are associated with known human diseases, (ii) functional amyloids that confer well-described functional benefits to the organism, (iii) functional prions that are self-propagating amyloid forms of proteins that confer functional benefits to the organism, (iv) amyloids whose functional relevance is not known and (v) biologically not relevant amyloids that were observed under conditions whose biological relevance is questionable. Besides the classification, a short description is provided for each entry that explains the function of the protein and the proposed role of its amyloid state.

Implementation

AmyPro is implemented as an AngularJS web application fueled by the underlying data stored in multi-level JSON. The advantages of this technology are that it scales well with increasing data traffic and it enables responsive and dynamic behavior on multiple media, such as personal computers and hand-held devices.

Each data record contains an entry ID, cross-references to other resources (PDBe (35), UniProt (24), PubMed),

source organism, protein name and sequence, alternative names for both the investigated and the parent protein (if relevant) and the recommended amyloid name (if relevant), mutations (if relevant), functional classification, a list of the experimental techniques applied to study the amyloid state and the actual amyloidogenic sequence region(s). Where available, determined structures grouped according to relevant protein states are also provided as direct links to the corresponding wwPDB pages.

Data can be accessed via either the online user interface, direct URLs serving up FASTA or raw TXT formats and a RESTful API serving JSON, for example http://amypro. net/data/entries/AP00017.txt directly provides the data of entry AP00017 in TXT format, and http://amypro.net/data/ entries/AP00007.json displays entry AP00007 in JSON.

RESULTS

The AmyPro database (http://amypro.net) greatly improves accessibility to amyloid data both in a qualitative and quantitative sense, by providing 143 carefully curated, experimentally validated amyloid fibril-forming proteins (as of September 2017), including 127 with defined amyloidogenic regions totaling 5819 residues (2587 in prion domains and 3232 in annotated shorter amyloidogenic fragments). It is important to emphasize that AmyPro only contains aggregating proteins that form amyloids/prions, but it is not intended to host proteins implicated in phase separation or in forming amorphous aggregates. Data collection will continue in order to maintain a regularly updated online resource, with the goal of periodic releases twice every year. To this end, we also kindly invite the scientific community to contribute by sending data using our online submission interface (http://www.amypro.net/#/submit). AmyPro currently includes 57 pathogenic and 54 functional amyloids, 17 functional prions and 15 cases where the role of the amyloid state is not known or has only been observed under non-native conditions and thus was considered as biologically not relevant. Apolipoprotein A-I is an illustrative example (http://www.amypro.net/#/entries/AP00003) of the increased quality of our database: the amyloidogenic region was previously loosely defined to residues 1-93 (25), but is now redefined to a 14-residue stretch following a recent publication (30).

User interface and website features

The online user interface of AmyPro offers convenient ways to access the amyloidogenic data. The database can be browsed and searched by clicking on the respective button on the top navigation bar. While browsing, the displayed entry list can be filtered by category (pathogenic, functional prion, functional amyloid, unknown, biologically not relevant) and by typing in filtering terms, such as a species- or protein name. Similarly, searching can be performed by typing in a search term, and then the results may be further refined by adding species name or category. Examples of search terms are AmyPro or UniProt IDs, protein/gene names or species.

Both browsing and searching results are displayed as lists of relevant entries arranged as a table. Clicking on any row

will direct the user to the dedicated entry page. Clicking on the plus signs on the left hand side allows the user to select multiple entries, and then perform a batch download in JSON format by pressing the 'Download selected' button.

Searching can also be performed using an input sequence on the search screen. Pasting a sequence and pressing 'Match' will run a screening process and return all those entries where at least six consecutive residues match any segment within the user input. Again, clicking on any row in the results table will lead to the entry page.

In addition to accessing specific entries, the complete dataset is also available for download in three different data formats. Pressing the 'Download' button on the top navigation menu displays a drop-down menu from which the data format can be selected. AmyPro is currently available for download in JSON, FASTA and raw TXT formats.

We invite the scientific community to submit their own published data by filling in our user-friendly online submission form. Clicking on the 'Submit' button on the top navigation bar directs the user to this form.

Finally, we provide database statistics (sequence length distributions and amino acid compositions) on the 'Stats' page, and an online documentation of all the functionalities of AmyPro on the 'Help' page, both accesses using the top navigation menu.

Accession pages

Each database entry can be investigated in detail by accessing its dynamically generated entry page via browsing, searching or direct URL links (Figure 1).

On the top of each entry page, the entry ID is displayed along with the name of the protein, the International Society of Amyloidosis (ISA) recommended name of the derived amyloid (36) with smaller letters (if relevant), and the classification of the regions (pathogenic, functional prion, functional amyloid, unknown, biologically not relevant). This section also provides the alternative names, the source organism and the precursor/parent protein with its respective alternative names (where relevant), the related publications and direct links to download the data in TXT, FASTA or JSON formats. Below this section a brief, amyloid state-centric description of the entry is provided followed by the corresponding UniProt functional description that describes the parent protein if not explicitly stated otherwise. The next two sections provide information on the functional classification of the amyloid form and a list of the experimental techniques by which it was discovered/ studied in the listed publications.

The investigated protein sequence (as provided by the authors) is displayed below, with the corresponding UniProt accession and residue boundaries provided in brackets and amyloidogenic regions colored according to their category's color code.

Below the data visualization, the actual amyloidogenic region and prion domain sequences are displayed in addition to their residue indices (corresponding to the entry sequence and the UniProt sequence) and links to the corresponding literature references. If in different publications overlapping segments were claimed amyloidogenic, those were merged to obtain our AmyPro amyloidogenic
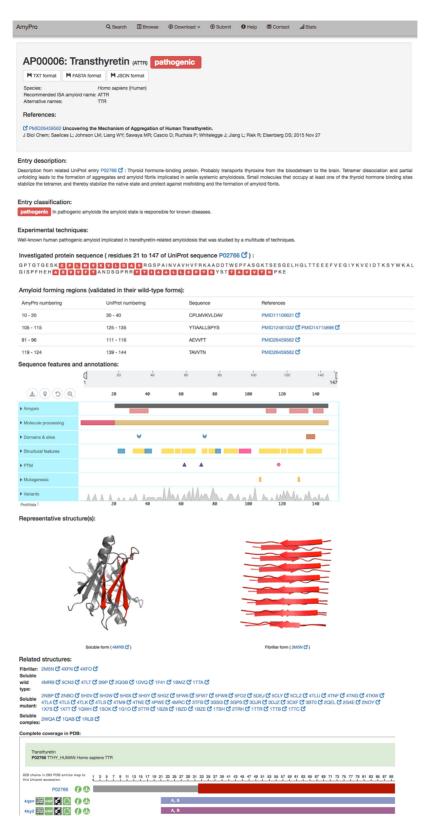
Figure 1. AmyPro entry screen. AmyPro provides all the relevant information on specific entry pages. In addition to displaying data, these pages also offer direct download links (TXT, FASTA and JSON formats) as well as data visualization by mapping the amyloidogenic regions or prion domain of the given entry protein onto available structural data and functional annotations (imported from PDBe and UniProt, respectively).

segments and all the supporting publications are provided. The bottom section of the entry pages show important UniProt features after pressing the 'Click to load Feature Viewer' button, using the recently released visualization tool of UniProt, ProtVista (37), extending it with AmyPro data. Using this tool, the AmyPro sequence and amyloidogenic regions can easily be compared to various features, such as protein processing data, structural features, PTMs and variations (each corresponding to an extendable line in the table). When applicable, relevant PDB IDs grouped into fibrillar (including structures of short cross-β spines with varied steric zippers), precursor and soluble states, also distinguishing wild-type, segment, mutant and complex forms within the soluble state are listed below. These PDB IDs link to the corresponding PDBe/RCSB/PDBj entry pages. Representative PDB structures of the soluble wild-type and/or fibrillar forms are also displayed (in the absence of the full-length soluble wild-type sequence, a relevant segment of the protein may be displayed), with the amyloidogenic regions highlighted. Finally, at the bottom of the entry page, a web component is displayed that maps every available PDB structure to the UniProt ID, while also displaying Pfam entities. This web component was developed and is maintained by PDBe in their component library (https://www.ebi.ac.uk/pdbe/pdb-component-library/).

DISCUSSION

AmyPro is a novel, carefully curated database of amyloid precursor proteins and their amyloidogenic sequence regions that aims to provide an up-to-date view of the entire amylome. It contains significantly more amyloid-forming proteins than any of its peers, mainly due to incorporating validated functional amyloids and prion proteins from all kingdoms of life for the first time. AmyPro will be regularly updated, relying not only on data submissions from the research community but also on regular internal updates based on scanning newly published amyloid literature. By storing the most comprehensive list of amyloid fibril-forming proteins published so far in combination with a useful set of features facilitating their efficient analysis, we anticipate AmyPro to become central to amyloid research, and to have major impact on its progress. To achieve this goal we are dedicated to ensure the long-term availability of the database.

In particular, AmyPro has huge potential in helping to elucidate the sequential determinants of amyloid fibril formation. Furthermore, the provided functional classifications might enable refining these sequential determinants according to the biological functions of amyloid fibrils and understanding the functional relevance of their differences. Better definition of the sequence determinants of amyloid fibril formation will also help in understanding and predicting the effects of mutations within amyloidogenic sequence fragments and relate them to the underlying molecular mechanisms, which in turn may enable the development of novel methods for more accurate computational identification of amyloidogenic regions within proteins.

AVAILABILITY

AmyPro is available at http://amypro.net and is open to submissions from the scientific community. We kindly encourage users to submit newly identified amyloid fibril-forming proteins to AmyPro using our online submission form (http://amypro.net/#/submit) or to contact us about new information on the already existing entries.

REFERENCES

1. Knowles,T.P., Vendruscolo,M. and Dobson,C.M. (2014) The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.*, **15**, 384–396.
2. Lotz,G.P. and Legleiter,J. (2013) The role of amyloidogenic protein oligomerization in neurodegenerative disease. *J. Mol. Med. (Berl)*, **91**, 653–664.
3. Mukherjee,A., Morales-Scheihing,D., Butler,P.C. and Soto,C. (2015) Type 2 diabetes as a protein misfolding disease. *Trends Mol. Med.*, **21**, 439–449.
4. Obici,L., Franceschini,G., Calabresi,L., Giorgetti,S., Stoppini,M., Merlini,G. and Bellotti,V. (2006) Structure, function and amyloidogenic propensity of apolipoprotein A-I. *Amyloid*, **13**, 191–205.
5. Nelson,R., Sawaya,M.R., Balbirnie,M., Madsen,A.O., Riekel,C., Grothe,R. and Eisenberg,D. (2005) Structure of the cross-beta spine of amyloid-like fibrils. *Nature*, **435**, 773–778.
6. Sawaya,M.R., Sambashivan,S., Nelson,R., Ivanova,M.I., Sievers,S.A., Apostol,M.I., Thompson,M.J., Balbirnie,M., Wiltzius,J.J., McFarlane,H.T. *et al.* (2007) Atomic structures of amyloid cross-beta spines reveal varied steric zippers. *Nature*, **447**, 453–457.
7. Fowler,D.M., Koulov,A.V., Balch,W.E. and Kelly,J.W. (2007) Functional amyloid–from bacteria to humans. *Trends Biochem. Sci.*, **32**, 217–224.
8. Pham,C.L., Kwan,A.H. and Sunde,M. (2014) Functional amyloid: widespread in Nature, diverse in purpose. *Essays Biochem.*, **56**, 207–219.
9. Lembre,P., Vendrely,C. and Martino,P.D. (2014) Identification of an amyloidogenic peptide from the Bap protein of Staphylococcus epidermidis. *Protein Pept. Lett.*, **21**, 75–79.
10. Otoo,H.N., Lee,K.G., Qiu,W. and Lipke,P.N. (2008) Candida albicans Als adhesins have conserved amyloid-forming sequences. *Eukaryot. Cell*, **7**, 776–782.
11. Romero,D., Aguilar,C., Losick,R. and Kolter,R. (2010) Amyloid fibers provide structural integrity to Bacillus subtilis biofilms. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 2230–2234.
12. Wang,X., Zhou,Y., Ren,J.J., Hammer,N.D. and Chapman,M.R. (2010) Gatekeeper residues in the major curlin subunit modulate bacterial amyloid fiber biogenesis. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 163–168.
13. Bieler,S., Estrada,L., Lagos,R., Baeza,M., Castilla,J. and Soto,C. (2005) Amyloid formation modulates the biological activity of a bacterial protein. *J. Biol. Chem.*, **280**, 26880–26885.
14. Gour,S., Kaushik,V., Kumar,V., Bhat,P., Yadav,S.C. and Yadav,J.K. (2016) Antimicrobial peptide (Cn-AMP2) from liquid endosperm of Cocos nucifera forms amyloid-like fibrillar structure. *J. Pept. Sci.*, **22**, 201–207.

15. Maji,S.K., Perrin,M.H., Sawaya,M.R., Jessberger,S., Vadodaria,K., Rissman,R.A., Singru,P.S., Nilsson,K.P., Simon,R., Schubert,D. *et al.* (2009) Functional amyloids as natural storage of peptide hormones in pituitary secretory granules. *Science*, **325**, 328–332.

16. Louros,N.N., Iconomidou,V.A., Giannelou,P. and Hamodrakas,S.J. (2013) Structural analysis of peptide-analogues of human Zona Pellucida ZP1 protein with amyloidogenic properties: insights into mammalian Zona Pellucida formation. *PLoS One*, **8**, e73258.

17. Louros,N.N., Petronikolou,N., Karamanos,T., Cordopatis,P., Iconomidou,V.A. and Hamodrakas,S.J. (2014) Structural studies of "aggregation-prone" peptide-analogues of teleostean egg chorion ZPB proteins. *Biopolymers*, **102**, 427–436.

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

19. Tartaglia,G.G. and Vendruscolo,M. (2010) Proteome-level interplay between folding and aggregation propensities of proteins. *J. Mol. Biol.*, **402**, 919–928.

20. Pawlicki,S., Le Bechec,A. and Delamarche,C. (2008) AMYPdb: a database dedicated to amyloid precursor proteins. *BMC Bioinformatics*, **9**, 273.

21. Beerten,J., Van Durme,J., Gallardo,R., Capriotti,E., Serpell,L., Rousseau,F. and Schymkowitz,J. (2015) WALTZ-DB: a benchmark database of amyloidogenic hexapeptides. *Bioinformatics*, **31**, 1698–1700.

22. Wozniak,P.P. and Kotulska,M. (2015) AmyLoad: website dedicated to amyloidogenic protein fragments. *Bioinformatics*, **31**, 3395–3397.

23. Thangakani,A.M., Nagarajan,R., Kumar,S., Sakthivel,R., Velmurugan,D. and Gromiha,M.M. (2016) CPAD, curated protein aggregation database: a repository of manually curated experimental data on protein and peptide aggregation. *PLoS One*, **11**, e0152949.

24. UniProt,C. (2015) UniProt: a hub for protein information. *Nucleic Acids Res.*, **43**, D204–D212.

25. Tsolis,A.C., Papandreou,N.C., Iconomidou,V.A. and Hamodrakas,S.J. (2013) A consensus method for the prediction of 'aggregation-prone' peptides in globular proteins. *PLoS One*, **8**, e54175.

26. Iconomidou,V.A., Leontis,A., Hoenger,A. and Hamodrakas,S.J. (2013) Identification of a novel 'aggregation-prone'/'amyloidogenic determinant' peptide in the sequence of the highly amyloidogenic human calcitonin. *FEBS Lett.*, **587**, 569–574.

27. Saelices,L., Johnson,L.M., Liang,W.Y., Sawaya,M.R., Cascio,D., Ruchala,P., Whitelegge,J., Jiang,L., Riek,R. and Eisenberg,D.S. (2015) Uncovering the mechanism of aggregation of human transthyretin. *J. Biol. Chem.*, **290**, 28932–28943.

28. Tsiolaki,P.L., Hamodrakas,S.J. and Iconomidou,V.A. (2015) The pentapeptide LQVVR plays a pivotal role in human cystatin C fibrillization. *FEBS Lett.*, **589**, 159–164.

29. Tsiolaki,P.L., Louros,N.N., Hamodrakas,S.J. and Iconomidou,V.A. (2015) Exploring the 'aggregation-prone' core of human Cystatin C: a structural study. *J. Struct. Biol.*, **191**, 272–280.

30. Wong,Y.Q., Binger,K.J., Howlett,G.J. and Griffin,M.D. (2012) Identification of an amyloid fibril forming peptide comprising residues 46–59 of apolipoprotein A-I. *FEBS Lett.*, **586**, 1754–1758.

31. Hernandez-Santoyo,A., del Pozo Yauner,L., Fuentes-Silva,D., Ortiz,E., Rudino-Pinera,E., Sanchez-Lopez,R., Horjales,E., Becerril,B. and Rodriguez-Romero,A. (2010) A single mutation at the sheet switch region results in conformational changes favoring lambda6 light-chain fibrillogenesis. *J. Mol. Biol.*, **396**, 280–292.

32. Sorensen,C.S., Runager,K., Scavenius,C., Jensen,M.M., Nielsen,N.S., Christiansen,G., Petersen,S.V., Karring,H., Sanggaard,K.W. and Enghild,J.J. (2015) Fibril core of transforming growth factor beta-induced protein (TGFBIp) facilitates aggregation of corneal TGFBIp. *Biochemistry*, **54**, 2943–2956.

33. Vidal,R., Frangione,B., Rostagno,A., Mead,S., Revesz,T., Plant,G. and Ghiso,J. (1999) A stop-codon mutation in the BRI gene associated with familial British dementia. *Nature*, **399**, 776–781.

34. Vidal,R., Revesz,T., Rostagno,A., Kim,E., Holton,J.L., Bek,T., Bojsen-Moller,M., Braendgaard,H., Plant,G., Ghiso,J. *et al.* (2000) A decamer duplication in the 3' region of the BRI gene originates an amyloid peptide that is associated with dementia in a Danish kindred. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 4920–4925.

35. Velankar,S., van Ginkel,G., Alhroub,Y., Battle,G.M., Berrisford,J.M., Conroy,M.J., Dana,J.M., Gore,S.P., Gutmanas,A., Haslam,P. *et al.* (2016) PDBe: improved accessibility of macromolecular structure data from PDB and EMDB. *Nucleic Acids Res.*, **44**, D385–D395.

36. Sipe,J.D., Benson,M.D., Buxbaum,J.N., Ikeda,S.I., Merlini,G., Saraiva,M.J. and Westermark,P. (2016) Amyloid fibril proteins and amyloidosis: chemical identification and clinical classification International Society of Amyloidosis 2016 Nomenclature Guidelines. *Amyloid*, **23**, 209–213.

37. Watkins,X., Garcia,L.J., Pundir,S., Martin,M.J. and UniProt Consortium,T. (2017) ProtVista: visualization of protein sequence annotations. *Bioinformatics*, **33**, 2040–2041.