UNIVERSITY OF CALIFORNIA

Los Angeles

Genetic Causes of Amyloid Fibrils and Their Structures

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Human Genetics

by

Gregory Marc Rosenberg

2023

ABSTRACT OF THE DISSERTATION

Genetic Causes of Amyloid Fibrils

and Their Structures

by

Gregory Marc Rosenberg

Doctor of Philosophy in Human Genetics

University of California, Los Angeles, 2023

Professor David S. Eisenberg, Co-Chair

Professor Daniel H. Geschwind, Co-Chair

The conversion of human proteins into the highly ordered fibrillar aggregates known as amyloid fibrils is implicated in a wide range of diseases. The number of proteins known to exhibit this activity is constantly growing and the molecular mechanisms that lead to amyloid aggregation are not fully understood. I focus on genetic mutations as a mechanism by which to predict novel amyloidogenic proteins. I also examine how mutations can explain some structural features of amyloid fibrils formed by mutant proteins. I developed a computational method for predicting previously unknown amyloidogenic proteins by calculating the propensity of disease mutations to induce amyloid aggregation. This method, called Identification of Mutations Promoting

Amyloidogenic Transitions (IMPAcT), identified protein TFG as a novel amyloidogenic protein. I biochemically validated the amyloidogenic nature of mutant protein TFG and determined the structures of the mutant amyloid fibrils, demonstrating the direct and indirect influence of each mutation on the structure and stability of the fibrils. Lastly, I composed a literature review of the mechanisms by which mutations drive the conversion to an amyloid fibril and influence the resulting molecular structures. This body of research expands our understanding of amyloid proteins and opens avenues for further study of the relationship between genetics and amyloids.

This dissertation includes three supplementary files.

Chapter 2 has two supplementary spreadsheets: Table S2-1 is a spreadsheet containing information regarding all the candidate mutations identified by the IMPAcT method; Table S2-2 is a spreadsheet containing the results of a statistical test on patterns in the kinds of mutations which were predicted to be amyloidogenic by the IMPAcT method.

Chapter 4 has one supplementary spreadsheet: Table S4-1 is a spreadsheet containing relevant information on all pathogenic amyloid proteins considered in the review.

The dissertation of Gregory Marc Rosenberg is approved.

Kathrin Plath

Todd O. Yeates

David S. Eisenberg, Committee Co-Chair

Daniel H. Geschwind, Committee Co-Chair

University of California, Los Angeles

2023

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

Chapter 2 Figures and Tables

Chapter 3 Figures and Tables

Chapter 4 Figures and Tables

ACKNOWLEDGEMENTS

My journey through my PhD was defined by an enormous amount of good fortune. The lab I chose to do my PhD research in is filled to the brim with wonderful, supportive people, many of whom became *de facto* mentors for me in addition to my extremely supportive, creative, and knowledgeable P.I. David S. Eisenberg. I am very grateful for the positive, nurturing environment I got to be a part of as a member of this lab.

I am also thankful for the support of my family and friends throughout my education, especially my mom Diane Rosenberg, my dad David Rosenberg, my brother Paxton Rosenberg, and my wife Myra Rosenberg. The continued and enthusiastic support I received from those close to me, including my colleagues at UCLA, helped me to thrive in academia. I have tremendous gratitude for all my friends, family, and colleagues who have been by my side throughout my PhD education.

I would also like to enumerate the specific contributions of the co-authors of the research included in this dissertation, without whom this work would not have been possible.

Chapter 2 is a version of Gregory M. Rosenberg, Kevin A. Murray, Lukasz Salwinski, Michael P. Hughes, Romany Abskharon, and David S. Eisenberg. Bioinformatic identification of previously unrecognized amyloidogenic proteins. J Biol Chem. (2022) Volume 298, Issue 5, DOI: 10.1016/j.jbc.2022.101920

Kevin A. Murray assisted in the formulation of the method used in the publication and assisted in protein expression and purification. Lukasz Salwinski gathered data from databases. Michael P. Hughes provided scripts for defining low-complexity domains in the human proteome and for accessing the ZipperDB database in a high-throughput fashion. Romany Abskharon assisted in protein expression and purification. David S. Eisenberg is the P.I.

Chapter 3 is a manuscript in preparation for publication. Its co-authors include Romany Abskharon, David R. Boyer, Peng Ge, Michael R. Sawaya, and David S. Eisenberg.

Romany Abskharon assisted in preparation of samples for cryo-electron microscopy, specifically preparation and freezing grids. David R. Boyer and Peng Ge assisted with acquiring microscope time for collecting data and guidance on how to use cryo-electron microscopy software. Michael R. Sawaya provided computer scripts used for structure refinement, calculation of solvation energy values, and for generating figures; he also provided aesthetic advice for some figures. David S. Eisenberg is the P.I.

Chapters 4 and 5 are parts of a literature review manuscript in preparation for publication. Their co-authors include Kevin A. Murray, Michael R. Sawaya, Yi Xiao Jiang, and David S. Eisenberg.

Kevin A. Murray assisted with the conception of the subjects the review would cover and made figure 4-8. Michael R. Sawaya assisted with the organization of the review and assisted with the analysis of amyloid fiber structures. He also provided aesthetic advice on figures. Yi Xiao Jiang assisted with proofreading, checking the accuracy of sections discussing proteins in his purview of research, and providing useful citations for those proteins. David S. Eisenberg is the P.I.

The research in this dissertation was supported by the National Institutes of Health (NIH) (grant numbers: R01AG048120 and R01AG07895), the NIH National Center for Advancing Translational Science (NCATS) UCLA CTSI Grant Number UL1TR001881, and NIH Grant Number R01GM123126. Some of this work was performed at the Stanford-SLAC Cryo-EM Center (S2C2), which is supported by the National Institutes of Health Common Fund Transformative High-Resolution Cryo-Electron Microscopy program (grant U24 GM129541). We also acknowledge the use of resources at the Electron Imaging Center for Nanomachines (EICN; supported by UCLA and by instrumentation grants from the NIH (1S10OD018111 and 1U24GM116792) and the NSF

(DBI-1338135 and DMR-1548924)). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Education

BS UCLA; Microbiology, Immunology, and Molecular Genetics June 2017

Honors and Awards

ASBMB 2023 Graduate Student or Postdoctoral Researcher Travel Award 2023

Microbiology, Immunology, and Molecular Genetics TA Award 2020

UCLA Dean's Honor List 2015-2017

LA Biomedical Research Institute Summer Fellowship 2013

Research Experience

PhD Research, Eisenberg Lab, UCLA 2018-Present

PhD Rotations, Geschwind Lab; Pajukanta Lab, UCLA 2017 to 2018

Undergraduate Research, Lusis Lab, UCLA 2015 to 2017

Los Angeles Biomedical Research Institute (Lundquist Institute), Torrance, CA 2013

Teaching Experience

UCLA, Los Angeles, CA Jan 2019 to Mar 2019; Jan 2020 to Mar 2020

Teaching Assistant, Microbiology, Immunology, and Molecular Genetics

Journal Publications

Murray KA, Hu CJ, Pan H, Lu J, Abskharon R, Bowler JT, **Rosenberg GM**, Williams CK, Elezi G, Balbirnie M, Faull KF, Vinters HV, Seidler PM, Eisenberg DS. **Small molecules disaggregate alpha-synuclein and prevent seeding from patient brain-derived fibrils**. Proc Natl Acad Sci U S A. 2023 Feb 14;120(7):e2217835120. doi: 10.1073/pnas.2217835120. Epub 2023 Feb 9. PMID: 36757890; PMCID: PMC9963379.

Murray KA, Hu CJ, Griner SL, Pan H, Bowler JT, Abskharon R, **Rosenberg GM**, Cheng X, Seidler PM, Eisenberg DS. **De novo designed protein inhibitors of amyloid aggregation and seeding**. Proc Natl Acad Sci U S A. 2022 Aug 23;119(34):e2206240119. doi: 10.1073/pnas.2206240119. Epub 2022 Aug 15. PMID: 35969734; PMCID: PMC9407671.

Rosenberg GM, Murray KA, Salwinski L, Hughes MP, Abskharon R, Eisenberg DS. **Bioinformatic identification of previously unrecognized amyloidogenic proteins**. J Biol Chem. 2022 May;298(5):101920. doi: 10.1016/j.jbc.2022.101920. Epub 2022 Apr 9. PMID: 35405097; PMCID: PMC9108986.

Tayeb-Fligelman E, Cheng X, Tai C, Bowler JT, Griner S, Sawaya MR, Seidler PM, Jiang YX, Lu J, **Rosenberg GM**, Salwinski L, Abskharon R, Zee CT, Hou K, Li Y, Boyer DR,

Murray KA, Falcon G, Anderson DH, Cascio D, Saelices L, Damoiseaux R, Guo F, Eisenberg DS. **Inhibition of amyloid formation of the Nucleoprotein of SARS-CoV-2**. bioRxiv [Preprint]. 2021 Mar 18:2021.03.05.434000. doi: 10.1101/2021.03.05.434000. PMID: 33688654; PMCID: PMC7941625.

Seidler PM, Boyer DR, Murray KA, Yang TP, Bentzel M, Sawaya MR, **Rosenberg G**, Cascio D, Williams CK, Newell KL, Ghetti B, DeTure MA, Dickson DW, Vinters HV, Eisenberg DS. **Structure-based inhibitors halt prion-like seeding by Alzheimer's disease-and tauopathy-derived brain tissue samples**. J Biol Chem. 2019 Nov 1;294(44):16451-16464. doi: 10.1074/jbc.RA119.009688. Epub 2019 Sep 19. PMID: 31537646; PMCID: PMC6827308.

Brumshtein B, Esswein SR, Sawaya MR, **Rosenberg G**, Ly AT, Landau M, Eisenberg DS. **Identification of two principal amyloid-driving segments in variable domains of Ig light chains in systemic light-chain amyloidosis**. J Biol Chem. 2018 Dec 21;293(51):19659-19671. doi: 10.1074/jbc.RA118.004142. Epub 2018 Oct 24. PMID: 30355736; PMCID: PMC6314132.

Seldin MM, Koplev S, Rajbhandari P, Vergnes L, **Rosenberg GM**, Meng Y, Pan C, Phuong TMN, Gharakhanian R, Che N, Mäkinen S, Shih DM, Civelek M, Parks BW, Kim ED, Norheim F, Chella Krishnan K, Hasin-Brumshtein Y, Mehrabian M, Laakso M, Drevon CA, Koistinen HA, Tontonoz P, Reue K, Cantor RM, Björkegren JLM, Lusis AJ. **A Strategy for Discovery of Endocrine Interactions with Application to Whole-Body Metabolism**. Cell Metab. 2018 May 1;27(5):1138-1155.e6. doi: 10.1016/j.cmet.2018.03.015. PMID: 29719227; PMCID: PMC5935137.

Journal Papers in Press

Garske KM, Kar A, Comenho C, Baillu B, Pan DZ, Bhagat YV, **Rosenberg G**, Koka A, Das SS, Miao Z, Sinsheimer JS, Kaprio J, Pietilainen KH, Pajukanta P. **Increased body mass index is linked to systemic inflammation through altered chromatin co-accessibility in human preadipocytes**. Nature Communications

Journal Papers in Preparation

Rosenberg GM, Abskharon R, Boyer DR, Ge P, Sawaya MR, Eisenberg DS. **Fibril structures of TFG protein mutants validate identification of TFG as a disease-related amyloid protein by the IMPAcT method**

Rosenberg GM, Murray KA, Sawaya MR, Jiang YX, Eisenberg DS. **Genetic causes of amyloid fibrils and their structures**

Presentations and Invited Lectures

Poster Presentation, “The Identification of Mutations Promoting Amyloidogenic Transitions (IMPAcT) method for identification of previously unrecognized amyloid disease-related proteins,” DiscoverBMB, Mar 28, 2023.

Chapter 1: Introduction

Amyloid formation is a component of the etiology of a wide variety of diseases and over 50 proteins are capable of forming amyloids. Amyloids are highly structured protein aggregates, and many years of research have gone into elucidating their common structural features as well as the unique features of individual amyloid proteins. Still, our knowledge of how amyloid formation contributes to disease is incomplete; we do not even have complete knowledge of every protein which can form amyloids in disease.

The goal of my research has been to discover unknown amyloidogenic proteins as well as uncover the mechanistic contribution of genetic factors, mainly Mendelian disease mutations, to the structure of amyloid fibrils and their relationship to disease etiology. The work included in this dissertation is the development and validation of a method to discover novel amyloid proteins based on the computational prediction of amyloidogenic disease mutations, cryo-electron microscopy structures of two amyloid fibrils of the same protein, each with a different disease mutation, and a literature review of the genetic causes of amyloid fibril formation which includes a catalogue of every known pathogenic amyloid and descriptions of the effects of and mechanisms behind genetic factors which lead to amyloid formation.

Chapter 2 of this dissertation is a scientific article describing my method of discovering novel amyloid proteins through a computational analysis of disease mutations on a subset of human proteins. This method has since been termed “Identification of Mutations Promoting Amyloidogenic Transitions”, or IMPAcT. The IMPAcT method works by compiling documented protein-coding disease mutations from publicly available databases, such as ClinVar, which occur in some subset of human proteins. In our case, proteins containing low-complexity domains were

chosen. The wild-type protein sequences and the mutant protein sequences are then scored by ZipperDB, a database which scores the propensity of a protein sequence to form an amyloid. Mutations which are calculated to convert a non-amyloidogenic sequence into one which is amyloidogenic are considered hits.

After implementing this method, I took steps to validate its output. Out of the proteins with mutations identified by the IMPAcT method, many were known to be amyloidogenic. One protein which was not known to be amyloidogenic was chosen for further validation: TFG. The low-complexity domain of TFG was expressed, purified, and tested for amyloidogenicity, and proved to indeed be amyloidogenic when the sequence contained the mutations identified by the IMPAcT method. The identification of TFG as a novel amyloidogenic protein validates the IMPAcT method and evinces an amyloid component of the etiology of the diseases caused by mutations in TFG.

The IMPAcT method is also versatile in that it can be repurposed with some different parameters to identify more novel candidates. Most apparently, the method can be applied to a different subset of the human proteome than the initial study. Also, the computation of amyloid propensity can be modulated by replacing the structural template that ZipperDB uses to calculate the amyloid propensity. The method may even be modified to judge the amyloidogenic effects of mutations in noncoding regions of the genome such as expression quantitative trait loci (eQTLs). I plan to implement some of these modifications and publish the results. This will further demonstrate the utility of this method to the field of amyloid research as well as expand the scope of the field to even more proteins.

Including the corresponding author and me, this article has six authors. Kevin Murray and I formulated the method together and I was responsible for implementing it through computer scripting. Lukasz Salwinski assisted in collecting data from mutation databases and I was responsible for parsing that data for my specific application. Michael Hughes provided scripts for identification of low-complexity domains as well as for accessing the ZipperDB database in a high-throughput fashion. Kevin Murray and Romany Abskharon provided guidance on expression and purification of protein constructs. I performed all other formulation of the overall research project, data analysis, experimentation, and writing, with oversight and direction from David Eisenberg.

Chapter 3 of this dissertation is a scientific article describing the cryo-electron microscopy (cryoEM) structure determination of two amyloid fibrils formed from recombinant protein constructs of the TFG low-complexity domain containing each of the two amyloidogenic mutations identified by the IMPAcT method. Through the determination of these structures, I describe how the different mutations directly contribute to the differences between the amyloid fibril cores. I also describe the structural features which are common to both structures, making them a hallmark of the TFG amyloid overall. I also compare the structures to other pathogenic and non-pathogenic amyloid structures and find that some features of the TFG structures are more similar to non-pathogenic amyloids, but certain stabilizing features, like glutamine ladders, may make some of the metrics used for comparison, like solvation energy, misleading. I also explore, based on the determined structures, why the mutants are able to form amyloids while the wild-type construct does not. The structures reveal how the mutations contribute to the formation of amyloid fibrils and thus provide a physical mechanistic explanation for the mutations' pathogenicity.

Including the corresponding author and me, this article has six authors. Romany Abskharon assisted in preparation of samples for imaging, namely freezing the cryoEM grids after loading samples of amyloid fibrils which I made. David Boyer and Peng Ge both assisted with acquiring time on electron microscopes for imaging my samples and provided expert guidance for utilizing software for processing of cryoEM data for structure determination of amyloid fibrils. Michael Sawaya provided the software I used to calculate solvation energy values and generate the figure containing that information. He also noticed some of the structural details which I describe in the article, namely the few extended β-strand conformations in the fibril core of the G269V mutant and the wave-like shape of the peptide backbone at the protofilament interface of the same mutant's fibril core. He also provided guidance on the aesthetics of some figures. I performed all other data analysis, experimentation, and writing, with oversight and direction from David Eisenberg.

Chapter 4 of this dissertation is a literature review on the genetic causes of amyloid formation and structure. In the review I catalogue every protein which is known to form pathogenic amyloids and classify them based on whether their wild-type form, mutant form, or both are known to form amyloid fibrils. I also compare the amyloid fibril structures of wild-type and mutant proteins for the several amyloid proteins for which both wild-type and mutant fibril structures have been determined. I also describe the biochemical mechanisms by which mutations contribute to amyloid formation and divide them into six different categories. The next section of the review examines haploinsufficiency in amyloid protein-coding genes and relates this to the inheritance patterns of most amyloidogenic disease mutations and how that can inform us of the mechanisms of amyloid pathogenicity. I end the review with a section about the scientific merit of determining the structures of amyloid fibrils formed by mutant proteins.

The review serves as a call to action for more research into the effects of genetic factors on the formation and structures of amyloid fibrils. This review is a unique and valuable addition to the literature of the amyloid field, especially due to the inclusive list of pathogenic amyloid proteins which is more comprehensive than any other review on a similar subject and the enumeration of amyloidogenic mechanisms of mutations.

Including the corresponding author and me, this review has five authors. Kevin Murray helped with the initial conception of the subjects the review would cover including helping to finalize the categories of amyloidogenic mutation mechanisms. He also made the figure for the mutation mechanism section of the review. Michael Sawaya helped with finalizing the overall organization of the review and assisted me with the comparisons of wild-type and mutant amyloid structures. He also provided advice on figure aesthetics, especially for the figures for the structure comparisons section. Yi Xiao Jiang assisted with proofreading and editing as well as providing important references for some of the amyloid proteins he has studied in his research. I wrote the entire text of the review, gathered all the references except for those provided by Yi Xiao, and made all the figures except for the one made by Kevin, with oversight and direction from David Eisenberg.

Chapter 5 of this dissertation is a resource document which is related to the literature review of chapter 4. This document contains short profiles describing each pathogenic amyloid protein included in the review in chapter 4. Each profile includes information on the protein relevant to the topics discussed in chapter 4: a description of the protein's function, the amyloid disease(s) it causes, the evidence which confirms its amyloid nature, and the mechanisms by which its

mutations, if any, induce amyloidogenicity. These amyloid protein profiles serve as a comprehensive resource for information on every known pathogenic amyloid protein.

I composed the document in its entirety and gathered every reference except for some collected by Yi Xiao for certain proteins.

In all, the research included in this dissertation is significant due to the development and validation of a method for discovery of novel amyloid proteins, the structure determination of two mutant amyloid fibrils which demonstrate the influence of disease mutations on amyloid structures, and a comprehensive review of pathogenic amyloid proteins and how their mutations contribute to their amyloidogenicity. All the research described here is original and was performed mainly by me, with contributions from other authors for their expertise in certain techniques. This work fulfills the scientific goals of my graduate studies and makes a significant contribution to science, and in particular to the study of amyloid proteins.

Chapter 2: Bioinformatic identification of TRK-fused gene protein (TFG) as a previously unrecognized amyloidogenic protein

Gregory M. Rosenberg, Kevin A. Murray, Lukasz Salwinski, Michael P. Hughes, Romany Abskharon, David S. Eisenberg

Abstract

Low-complexity domains of proteins have been shown to self-associate, and pathogenic mutations within these domains often drive the proteins into amyloid aggregation associated with disease. These domains may be especially susceptible to amyloidogenic mutations because they are commonly intrinsically disordered and function in self-association. The question therefore arises whether a search for pathogenic mutations in low-complexity domains of the human proteome can lead to identification of other proteins associated with amyloid disease. Here, we take a computational approach to identify documented pathogenic mutations within low-complexity domains which may favor amyloid formation. Using this approach, we identify numerous known amyloidogenic mutations, including several such mutations within proteins previously unidentified as amyloidogenic. Among the latter group, we focus on two mutations within the TRK-fused gene protein (TFG), known to play roles in protein secretion and innate immunity, which are associated with two different peripheral neuropathies. We show that both mutations increase the propensity of TFG to form amyloid fibrils. We therefore conclude that TFG is a novel amyloid protein and propose that the diseases associated with its mutant forms may be amyloidoses.

Introduction

Low-complexity domains (LCDs) are common, but functionally mysterious regions of proteins in the human proteome of which several are associated with amyloidoses(1–3). Low-complexity

domains are characterized by long segments made up of relatively low sequence diversity and are also commonly intrinsically disordered. LCDs are thought to be integral to the self-association of some proteins involved in RNA-binding, the formation of membraneless organelles, and the self-association of intermediate filament proteins(4–9), but not all proteins with LCDs exhibit these functions. Whereas subsequent dissociation of these complexes is a hallmark of the normal function of LCDs, proteins with LCDs may become prone to aggregate irreversibly into pathogenic amyloids due to missense mutations which encourage protein misfolding(3, 10, 11). Some examples of amyloidgenic LCD-containing proteins are FUS, TDP43, and HNRNPA1 which are all associated with ALS.

Because LCDs are often disordered, cross-β structures are present in LCD condensates(12), and many proteins containing them form amyloids in disease, we speculate that these domains may be more susceptible to mutations that cause the formation of an amyloid to be energetically favorable. Under this assumption, we chose to focus our search for novel amyloidogenic proteins on those proteins which contain a LCD (**Figure 2-1**). This search for unidentified amyloidogenic proteins based on pathogenic mutations expands on previous work(9, 13) by considering a larger subset of the human proteome. Our approach is agnostic to details about the queried proteins (besides identifying LCDs based on amino acid sequence) such as their functions or the diseases with which they are associated. Also, while our approach does identify many known amyloidogenic proteins, our focus is solely on those that have never been documented to form amyloid fibers either *in vivo* or *in vitro*.

Here we advance computational screening methods to identify mutations which may cause a functional LCD to become amyloidogenic(9). We define an amyloid as an irreversible, fibrous protein aggregate with a cross-β sheet scaffold. The common methods of experimental

identification of amyloid are the binding of amyloidophilic dyes such as Congo red or Thioflavin T (ThT) and x-ray diffraction revealing the ~4.7-4.8 Å separation of β-strands and ~10 Å separation of β-sheets. Amyloids are found in a wide range of diseases from Alzheimer's to type II diabetes(14). These amyloidoses are characterized by deposition of insoluble amyloid aggregates which, by mechanisms not completely understood, lead to cellular injury, tissue damage, and organ dysfunction(15). If amyloid deposition drives disease, it is crucial to identify the protein responsible to develop reliable treatments for the disease. We propose that many diseases have yet undiscovered amyloid components to their etiology. In this work, our algorithm identifies numerous known amyloidogenic mutations as well as many mutations not previously associated with any amyloidoses. Among the second group of mutations, we demonstrate that two from the protein TFG increase the amyloid propensity of the protein. The identification of this and other potential undiscovered amyloid proteins is important for understanding the pathogenesis and expanding the treatment options of their associated diseases.

Results

To identify LCDs, we first applied the SEG algorithm(16) to the human proteome to categorize amino acid segments as either high complexity or low complexity. We then conservatively defined a low-complexity domain as any low-complexity segment of at least 35 amino acids with leeway for 5 interrupting high complexity amino acids in a row. Under these criteria, 3,251 human proteins contain at least one LCD. We then scoured Uniprot Knowledgebase (UniprotKB)(17), Online Mendelian Inheritance in Man (OMIM)(18), and ClinVar(19) for pathogenic missense mutations within the LCDs of these proteins and found 738 documented disease-related mutations. This set of mutations was collected while remaining agnostic to the functional consequences for the affected protein, so among these mutations are some that are pathogenic because they increase

amyloid propensity, but also many that are pathogenic for various other reasons unrelated to amyloidogenicity.

Among the residues that make up LCDs, prolines are the most common, followed by glycine, serine, and alanine (**Figure 2-2A**). Glycine is by far the most common residue to be replaced in pathologies, with 501 of the 738 disease-related mutations being changed from glycine (**Figure 2-2B**). The most common mutation is from glycine to arginine, followed in descending order by glycine to aspartate, glycine to valine, and glycine to serine. The next most common residue to be replaced in pathologies is arginine with 57 mutations. This is followed by proline mutations, making up 50 of the documented disease-related mutations. These findings suggest that glycine residues are especially important in maintaining normal function of human LCDs.

Next, we sought to identify the mutations from the set of 738 pathogenic mutations that increase the propensity of a functional sequence to form a steric zipper, the common adhesive protein motif driving amyloid(20). To achieve this, we used ZipperDB, a database which predicts the fibril-forming propensity of segments within proteins(21). ZipperDB evaluates the energetic fit of 6-residue segments in the conformation of a steric zipper. Therefore, for each mutation we analyzed two 11-residue sequences centered on the mutated residue: one containing the wild-type residue and one containing the mutant residue. For each of these sequences, all 6 possible hexamers containing the residue of interest were assigned energy values by ZipperDB (**Figure 2-3, Figure 2-4A**). Proline residues are not energetically favorable in β sheets, so segments containing this residue tend to have very high positive energy scores. Since proline is a common residue in LCDs (**Figure 2-2A**), this skews a significant portion of the data to high positive values and **Figure 2-4A** does not include this skewed portion of the plot (**Figure S2-1**).

In detail, we identified amyloidogenic mutations as follows: we predict a mutation to be amyloidogenic if a hexamer containing the wild-type residue has an energy value greater than the ZipperDB threshold of -23 kcal/mol (therefore with lower amyloid propensity) and the corresponding hexamer containing the mutant residue has an energy value lower than or equal to -23 kcal/mol (therefore with higher amyloid propensity) (**Figure 2-3, Figure 2-4A, Table S2-1**). Mutations which generated these hexamers were predicted to cause a gain-of-function amyloid propensity. This group contains 88 mutations, the most common changes being glycine to serine, glycine to valine, glycine to cysteine, and proline to leucine (**Figure 2-4B**).

Conspicuously, changes to charged amino acids are greatly underrepresented in this predicted-amyloidogenic set. For example, glycine to arginine changes are completely absent whereas they are the most common type of pathogenic mutation in LCDs in general. This is unsurprising, however, since amino acid side chains on the interior of a steric zipper need to be able to pack closely together as well as to stack on top of each other along the fiber axis and these charged side chains will repel each other making amyloid formation energetically unfavorable.

Furthermore, statistical analysis reveals that glycine is the only residue that, when mutated, has statistically significant differences in whether the mutation is predicted to be amyloidogenic or not depending on which residue it mutates into (**Table S2-2**). In other words, for glycine only, the amino acid it mutates into is significant in determining whether the mutation is predicted to be amyloidogenic or not. Taken together, there are many documented mutations within LCDs that may drive a functional sequence to become amyloidogenic, but mutations from glycine and proline are more likely than others to be of this kind.

To validate that our approach is capable of identifying amyloidogenic mutations, we combed the list for known amyloidogenic mutations and found several. Two of the listed mutations in hnRNPA1 (D314V and D314N) and the mutation in hnRNPA2B1 (D302V) have been experimentally shown to induce fiber formation(3). One of the listed mutations in KRT8 (G62C) has been demonstrated to enhance aggregation propensity(9). Desmin, a protein which can form amyloid fibers in myofibrillar myopathy(22), has three mutations in the list (S2I, S46F, and S46Y) that are associated with myofibrillar myopathy and have been shown to cause abnormal aggregation(7). A mutation in PABPN1 (G12A) mimics a pathogenic polyalanine expansion(23) and an extended polyalanine segment in this protein has been shown to induce fiber formation(24). Known amyloidogenic proteins TDP43 and FUS also have mutations that appear on the list, but so far, these mutations have not been experimentally tested for increased amyloid propensity. These examples confirm that our approach can identify at least some mutations that contribute to the formation of amyloid fibers.

To determine if our approach identifies novel amyloid mutations, we analyzed one of the pathogenically altered proteins, TRK-fused gene protein (TFG). As part of its native function, TFG self-associates into octameric oligomers and its LCD facilitates these octamers to form larger complexes(25). The two mutations in TFG which were identified by our method, G269V and P285L, have been associated with Charcot-Marie-Tooth disease type 2 and hereditary motor and sensory neuropathy with proximal dominant involvement, respectively(26, 27), and both mutations were shown to result in abnormal aggregation of the protein. We also found that the WT protein and the protein with the P285L mutation are able to phase separate *in vitro* in the presence of a crowding agent, but with the G269V mutation the protein forms amorphous aggregates (**Figure S2-2**).

We expressed and purified the LCD of TFG (residues 237-327), fused with mCherry to increase solubility, in three forms: the wild-type sequence (WT), containing the G269V mutation (G269V), and containing the P285L mutation (P285L). Each construct was shaken at 37°C for 138 hours with thioflavin T, a dye which fluoresces in the presence of amyloid fibers(28, 29). Both mutant constructs demonstrated strong ThT fluorescence while the WT construct did not (**Figure 2-5A**). The presence of fibers from the mutant constructs and the absence of fibers from the WT construct were confirmed by electron microscopy (**Figure 2-5B**). Both mutant fibers displayed an apparent twist, typical of amyloid fibers.

The amyloid nature of the TFG 237-327 fibers was also confirmed through X-ray diffraction. Drops of solutions containing the fibers were suspended between glass rods and allowed to dry, which aligns the fibers between the rods. The fibers were placed in an X-ray beam with the length of the fibers perpendicular to the direction of the beam. The resulting diffraction pattern for both mutant fibers displayed distinct rings at resolutions representing the characteristic dimensions of an amyloid fiber: 10 Å representing the intersheet spacing and 4.7 Å representing the interstrand spacing (**Figure 2-5C**). We observed nearly the same behavior in the full-length protein, with the differences being faster fiber formation, likely due to lower solubility, and the WT sequence was able to form fibers, although at a significantly slower rate (**Figure S2-3**). In short, we found TFG exhibits amyloid behavior when pathogenically altered, as predicted by our bioinformatic approach.

Discussion

Interpretation of predicted amyloid-driving mutations in terms of atomic structure

The most common residues within LCDs to be replaced in pathologies are glycine, arginine, and proline, all of which are known to be important for the normal function of many LCDs especially in

regard to the regulation of phase-separation properties(30–32). Due to their functional importance in LCDs, altering these residues can result in pathology for various reasons unrelated to amyloidogenesis, but in this study we are most interested in pathogenic mutations in LCDs that result in increased amyloidogenicity specifically, and the trends in predicted amyloidogenic mutations in LCDs are slightly different than overall pathogenic mutations in LCDs, with proline being the second most common residue to have mutations predicted to be amyloidogenic rather than arginine.

The observed frequency of predicted amyloid-driving mutations (**Figure 2-4B**) can to some extent be understood in terms of atomic structures. Mutational replacement of a residue can destabilize the native state, favoring a conformational change leading to a pathogenic loss of native function; alternatively, replacement can favor formation of amyloid, leading to pathogenic amyloid. Both effects are possible with replacement of glycine, which we identified as the most commonly replaced residue associated with disease (**Figure 2-2B**) and with the most replacements predicted to be amyloidogenic (**Figure 2-4B**). Glycine residues confer flexibility to the peptide backbone which is important in maintaining the liquid properties of phase-separated protein droplets(30). Glycine's lack of a sidechain larger than a hydrogen atom disfavors secondary structure, since it grants an extra flexible peptide bond, and hence does not preferentially form α-helices or β-sheets(33). This means that glycine residues are less likely than other residues to contribute to the stability of an amyloid fiber core which is typically enriched with β-sheets. We previously demonstrated that glycines can introduce kinks into the backbone of beta-strands in amyloid structures derived from low-complexity segments(34), which may partially destabilize kinked backbones in low-complexity amyloid structures. We have also found that glycines lead to extended beta-strand motifs in these low-complexity amyloid structures, which may also contribute to their lability(35). Considering only single-nucleotide variant missense mutations,

glycine can potentially mutate into serine, alanine, glutamic acid, arginine, valine, aspartic acid, cysteine, and tryptophan (**Figure S2-4**). The most common glycine mutations which we predicted to be amyloidogenic are glycine to valine, glycine to cysteine, and glycine to serine.

Glycine to valine is the second most common glycine mutation predicted to be amyloidogenic. Valine has a relatively bulky side chain which is branched at the β carbon, and side chains like this prefer to form β-sheet secondary structures. Changes from glycine to valine would likely facilitate the formation of a steric zipper for this reason, especially if the surrounding sequence is already somewhat amyloid-prone other than being broken by the glycine residue and, as is the case for some glycine-rich LCD proteins (FUS and TDP43), forms metastable complexes with copies of itself as part of its function. One of the TFG mutations which increases amyloid propensity is a glycine to valine mutation (G269V).

Glycine to cysteine is the third most common glycine mutation which is predicted to induce amyloid propensity. The cysteine side chain is a thiol group which, when oxidized, can form disulfide bonds with other oxidized cysteine side chains. These covalent bonds normally greatly increase the stability of globular proteins, but these are normally intramolecular bonds. Intermolecular disulfide bonds can potentially stabilize cross-β interactions and contribute to the formation of amyloid fibers(36, 37). This intermolecular disulfide bonding is especially likely if the glycine to cysteine mutation creates a sequence with only a single cysteine in a region that is routinely exposed to copies of itself, as in many LCDs in which cysteines are not particularly common (**Figure 2-2A**). A glycine to cysteine mutation in KRT8 (G62C) has been shown to increase amyloid propensity(9).

Glycine to serine is the most common residue change out of the glycine mutations which are predicted to increase amyloid propensity. It is unclear exactly what benefit a serine residue would bring to the structure of an amyloid fiber since it has a polar side chain which is not able to form hydrogen bonds with itself when stacked like glutamine side chains. Nonpolar residues are more common on the interior of pathogenic amyloids, but polar residues do sometimes exist in steric zippers. One alternative role for serine is possibly facilitating the formation of a turn and contributing to the dagger-like fold which is seen in some structures of amyloid cores(20). Glycine to serine mutations in FUS which are associated with ALS (G206S, G191S) appear in our list of potentially amyloidogenic mutations.

After glycine mutations, proline mutations are the most common mutations predicted to be amyloidogenic in our set of interest (**Figure 2-4B**). Like glycine, proline is thought of as a secondary structure-breaker. Proline side chains break up β-strands due to steric restrictions of their phi and psi angles imposed by their unique side chains, which forms a bond with the α-amino nitrogen of the peptide backbone(38). Besides breaking the secondary structure necessary to form a steric zipper, the proline side chain also disfavors the formation of amyloid fibrils by removing the peptide backbone nitrogen's availability for interstrand hydrogen bonding, which normally is a major contributor to the stability of the overall fiber(7). With only single-nucleotide variant missense mutations, proline can potentially mutate into serine, alanine, arginine, leucine, threonine, glutamine, and histidine (**Figure S2-4**).

The most common proline mutation predicted to be amyloidogenic is from proline to leucine. Leucine is a nonpolar side chain with a branched gamma carbon. Packing of hydrophobic residues in the core of an amyloid fiber tends to increase stability and is preferred to polar residues in pathogenic amyloids(20). This explains how a mutation to leucine could contribute to the

formation of an amyloid fiber. One of the TFG mutations which increases amyloid propensity is a proline to leucine mutation (P285L).

It is important to consider the possibility that the distribution of common mutations predicted to be amyloidogenic may be sequence context-dependent rather than a result of which residue changes are most amyloidogenic in general. In other words, the reduced amino acid diversity in individual LCDs may limit which mutations actually increase the likelihood of a steric zipper within their sequence context. This could explain phenomena like proline to leucine mutations being more common in the list than proline to threonine mutations, the latter of which would have comparatively higher β-sheet propensity yet is much less common. Another consideration is codon limitations. Single-nucleotide missense mutations only allow for a limited number of amino acid changes and some are more likely than others due to similarities in codons and codon number (**Figure S2-4**). This explains why proline to leucine mutations are the most common type of proline mutation among the mutations predicted to be amyloidogenic, but there are no glycine to leucine mutations, since glycine to leucine mutations are impossible with only a single nucleotide change. It also makes it more significant that glycine to cysteine mutations are commonly predicted to be amyloidogenic even though there are fewer ways for single nucleotide changes to result in that mutation compared to glycine to alanine mutations (**Figure S2-4**). The underlying reasons for the distribution of mutations in this list requires further study.

Use of ZipperDB to assess mutations most likely to be amyloidogenic

We used ZipperDB to score sequences on their propensity to form a steric zipper, the core of amyloid fibers. ZipperDB threads sequences onto a peptide backbone based on the crystal structure of NNQQNY, a fibril-forming peptide from the yeast sup35 prion protein and generates an energy score. It is possible to utilize a different peptide backbone for steric zipper predictions,

but NNQQNY is the default and the one used for all the existing segments in the database. ZipperDB is not the only existing method for predicting amyloid fibers, but it is useful for high-throughput applications and is structure-based rather than sequence-based. Different amyloid-prediction tools can be variable in their predictions. To demonstrate this, we used AMYLPRED2(39) to predict amyloidogenic regions in TFG and its mutants (**Figure S2-5**). AMYLPRED2 employs up to 11 different amyloid-prediction methods and outputs their consensus. We ran AMYLPRED2 using 10 methods and there was enough consensus for a high-confidence prediction of an amyloid segment containing the G269V mutation, but not the P285L mutation. ZipperDB's scoring system is calibrated against experimental amyloid structures and has proven very reliable in predicting sequences which can form fibers(3, 21, 40, 41). Though, there are some important drawbacks to note. ZipperDB only considers homozippers while many amyloid fibers contain heterozipper interfaces at their cores, which can contribute to underprediction of amyloid-forming segments. Also, ZipperDB does not consider the sequence context of each segment it analyzes, so segments that may form fibers in theory may not actually be able to interact due to being buried in the interior of the protein or some other interference from surrounding segments, which can contribute to overprediction of amyloid-forming segments. These considerations mean that our method of identification has the potential to miss amyloids that would have been better identified by other methods and also include erroneous amyloid predictions.

Validation of our algorithm to identify amyloidogenic proteins

Our analysis has proven able to identify mutations which grant amyloidogenic gain of function. Within the list of mutations predicted to be amyloidogenic, there were many which have been previously demonstrated to promote amyloid fibrillation, as well as many mutations which have unknown structural consequences. Also, there were many mutations in proteins known to be able

to form amyloids, but the mutations themselves have no documentation on their biochemical consequences. It is also important to note that even if the mutation has the potential to cause the protein to form an amyloid, this does not necessarily mean that the protein will form an amyloid under physiological conditions in disease. This makes it difficult to gauge the specificity and sensitivity of our method. These considerations also factored into the model protein we used to validate our method, TFG, since TFG had previously been shown to aggregate when pathogenically altered in both cell models and tissue biopsies from diseased patients. However, other proteins in our list have also been demonstrated to aggregate when pathogenically altered but have not been shown to be amyloids, namely LMNA and CHCHD10(42, 43). We are currently analyzing the behavior of these proteins in regard to amyloid formation. Some other interesting proteins from our list of potential amyloid mutations include proteins that have the gene ontology (GO) molecular function term “identical protein binding”. This term encompasses all the known amyloid proteins from the list along with TFG and LMNA, but also other interesting proteins such as UBQLN2, which is involved in some forms of ALS(44), and GRM6, in which mutations can lead to night blindness due to disrupted trafficking of the protein(45).

The LCD of TFG was able to form amyloid fibers only when containing mutations, in line with what our method predicts (**Figure 2-5A,B**). However, the WT sequence of the full length protein is able to form amyloid fibers along with the mutant sequences, albeit at a much slower rate (**Figure S2-3**). This is not contradictory to our prediction, since the mutants show increased fiber-forming propensity in both contexts, but the discrepancy is interesting and warrants explanation. In the WT sequence there exist many segments predicted to be able to form a steric zipper, inside and outside the LCD, and any of these segments could drive its amyloid formation. There may have been differences in solubility between the full length protein and the LCD alone, especially since the LCD was conjugated to a molecule of mCherry, and if the full length protein is less soluble it

may have been more prone to forming fibers than the LCD alone in general. In the same vein, the WT sequence of the LCD may have been able to form fibers if given more time or dissolved in different buffer conditions. Another potential explanation is the inclusion of the PB1 domain in the full length protein which functions as a mediator of homo-oligomerization for TFG(25). This domain, not present in the mCherry-LCD constructs, may have facilitated self-interaction of the protein which promoted fiber formation. Regardless of the cause of the discrepancy, behavior of both the full length protein and the LCD conjugated to mCherry were consistent with our predictions.

Summary

In this study we combined documented disease-causing mutations with structure-based computation to predict amyloidogenic mutations. This method was validated by the identification of known amyloidogenic mutations as well as demonstrating the formation of amyloid fibers from sequences with mutations not previously identified as amyloidogenic. Our analysis has revealed many possible unidentified amyloid proteins which need to be validated biochemically.

Experimental Procedures

Low-Complexity Region Prediction

Amino acid sequences in the human proteome were evaluated for low complexity using SEG reference with default settings: window length = 12, trigger complexity 2.2, extension complexity 2.5. A sequence was determined to be a low-complexity region if it contained at least 35 residues scored as low complexity with at most 5 interrupting non-low-complexity residues.

Protein Expression and Purification

Recombinant TFG(237-327) for the WT, G269V, and P285L forms was purified using a pHis-parallel-mCherry vector, using a previously described method(6). Briefly, protein was overexpressed in BL21(DE3) Gold E. coli cells. Cultures were grown to an OD_{600}=0.4-0.8, then induced with 0.5M IPTG overnight. Cells were pelleted by centrifugation and the clarified lysate was purified by Ni-NTA columns followed by size exclusion chromatography and dialyzed into PBS.

Recombinant full length TFG was purified similarly except using a pet28b+ vector with a His tag but no mCherry and being dialyzed into buffer containing 20mM Tris pH8 and 150mM NaCl.

Phase Separation Assay

Recombinant full length TFG constructs were dissolved to 10μM concentration in buffer containing 25mM Tris pH 7.4, 150mM KCl, 2.5% v/v glycerol, and 10% w/v PEG 8000. Protein was first added to a microcentrifuge tube and diluted by adding the buffer on top of it. The total solution volume was 80μL. 3 aliquots of 20μL were then pipetted into a NUNC 384-well clear-bottom microplate and imaged immediately using differential interference contrast (DIC) microscopy.

In Vitro Aggregation Assay

Wild-type and mutant TFG LCD was diluted to 50μM in 1X PBS containing ThT at 40μM to a final volume of 150uL in black Nunc 96-well optical bottom plates (Thermo Scientific). A single PTFE bead (0.125 inch diameter) was added to each well to facilitate agitation. Plates were incubated in a microplate reader (FLUOstar OMEGA, BMG Labtech) for ~138 hours at 37^{o}C with 700 rpm double orbital shaking. Fluorescent measurements were recorded every 15 minutes using λ ex = 440 nm and λ em = 480 nm. This was performed with n=6 technical replicates.

Aggregation assays with full-length TFG were performed with the same method, except the PBS was replaced with buffer containing 20mM Tris pH8 and 150mM NaCl.

Transmission Electron Microscopy

10 uL of aggregated wild-type and mutant TFG samples (taken from in vitro aggregation experiments) was spotted onto carbon film on 150 mesh copper grids (Electron Microscopy Sciences) and incubated for 4 minutes. Grids were stained with 10 uL uranyl acetate solution (2% w/v in water) for 4 minutes. Excess solution was removed by blotting and air dried for 4 minutes. TEM images were acquired with a JOEL 100CX TEM electron microscope at 100 kV.

X-Ray Fiber Diffraction

Aggregated samples of TFG were centrifuged at 15,000 rpm for 30 minutes and buffer was exchanged with water twice. Samples were suspended between two siliconated glass capillaries ~1 mm apart, forming a bridge between the two capillaries. Sample was allowed to dry and the capillary with the dried aggregate was mounted on an in-house x-ray diffraction machine and diffracted with x-rays for 8 minutes, with the diffraction pattern collected on a CCD detector.

Data availability

The data that support the findings of this study are available from the corresponding author, D.S.E., upon reasonable request.

Acknowledgements

We thank Myra Rosenberg for aesthetic help with figures.

Author contributions

Project was conceived and designed by G.M.R and D.S.E with contribution from K.A.M. Mutation database data collection was performed by L.S. Bioinformatics analysis was performed by G.M.R and M.P.H. Protein expression and purification was performed by G.M.R., K.A.M. and R.A. G.M.R. performed in vitro aggregation and phase separation assays. X-ray fiber diffraction was performed by G.M.R. Electron microscopy was performed by G.M.R. Manuscript was written by G.M.R and D.S.E with contribution from all other authors.

Funding and additional information

This research was supported by NIH National Center for Advancing Translational Science (NCATS) UCLA CTSI Grant Number UL1TR001881, NIH Grant Number R01GM123126, and NIH Grant Number 2R01AG048120. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest

DSE is SAB Chair and equity holder in ADRx, Inc.

Tables and Figures

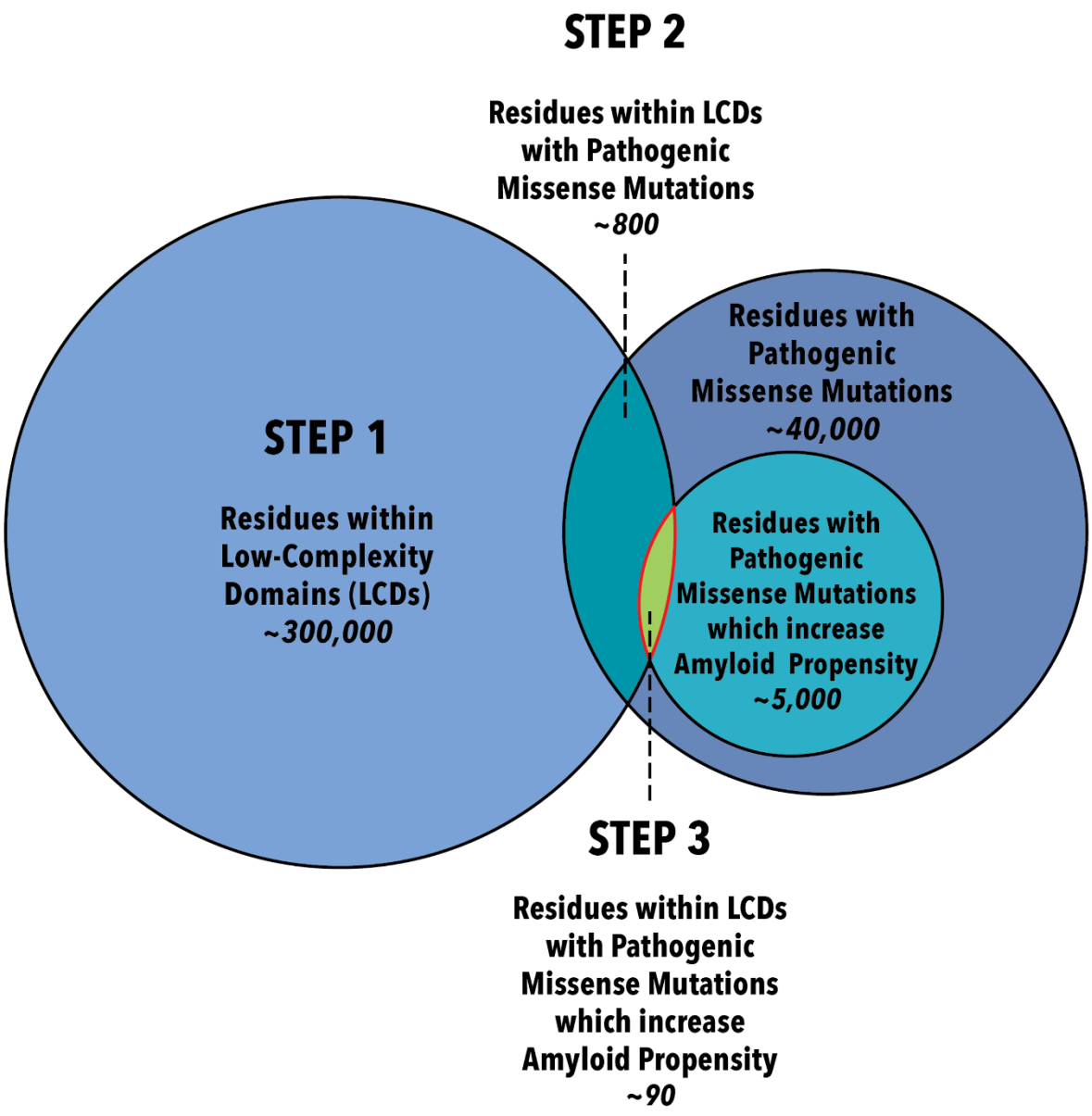


Figure 2-1 Schematic representation of our algorithm for identification of previously unrecognized amyloidogenic proteins. These proteins contain the mutated residues in the set at the intersection of all three circles: residues within LCDs with pathogenic missense mutations which increase amyloid propensity. We determined these residues in the three steps shown. Estimates for the number of residues represented in steps 1-3 are derived from this study. To estimate the number of residues with pathogenic missense mutations we used Simple ClinVar (https://simple-clinvar.broadinstitute.org/). To calculate the estimate for the number of missense

mutations which increase amyloid propensity, we extrapolated the percentage of mutations within LCDs which increase amyloid propensity (88/732; ~12%) to our estimate of total known pathogenic missense mutations in humans (~40,000) and rounded up.

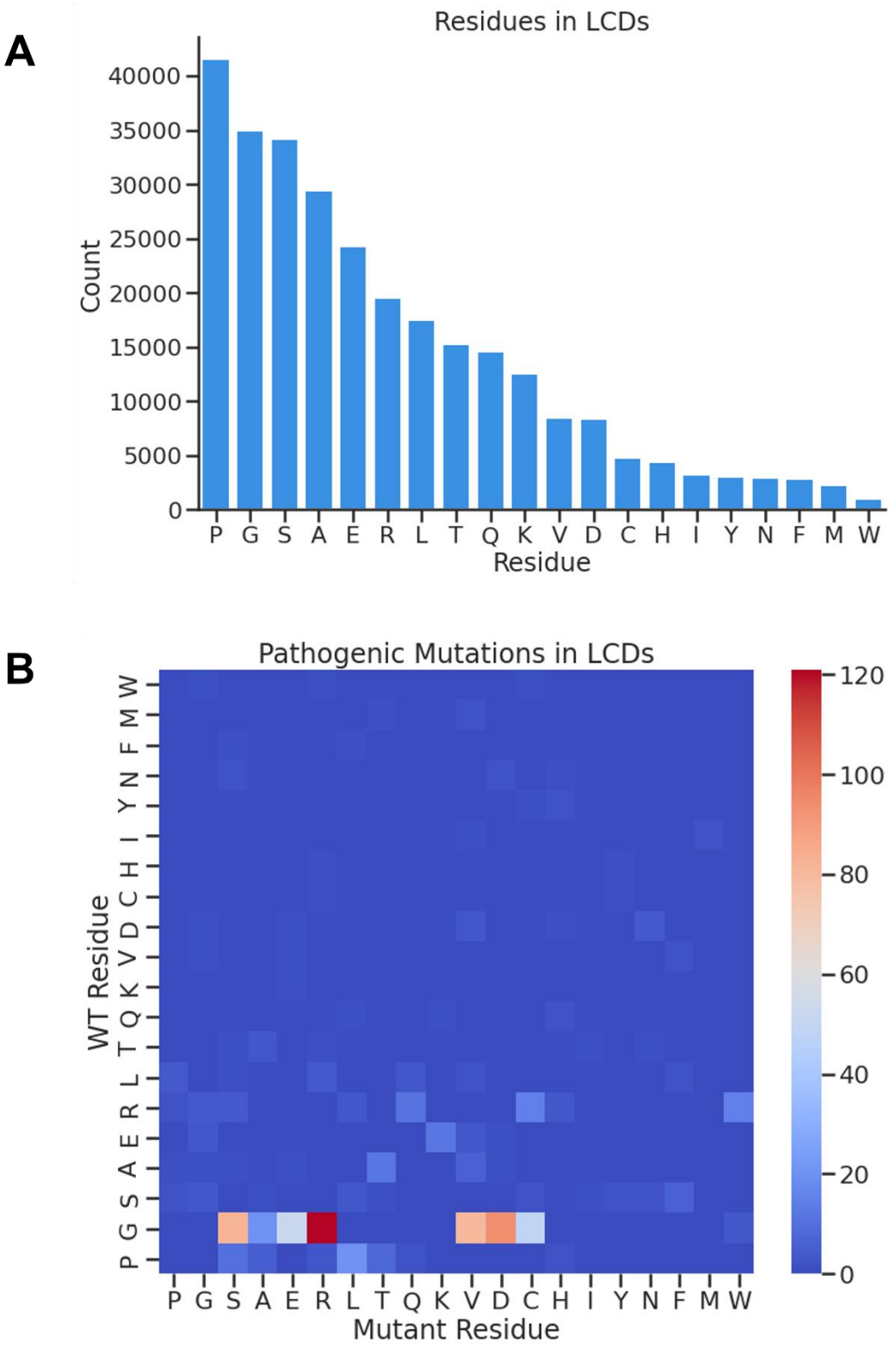

A
Residues in LCDs
Count
40000
35000
30000
25000
20000
15000
10000
5000
0
P G S A E R L T Q K V D C H I Y N F M W
Residue
B
Pathogenic Mutations in LCDs
WT Residue
P G S A E R L T Q K V D C H I Y N F M W
Mutant Residue
P G S A E R L T Q K V D C H I Y N F M W
120
100
80
60
40
20
0

Figure 2-2 Census of amino acid residues present in LCDs. A) Counts of all residues in LCDs in the human proteome. B) Heatmap displaying counts of all LCD residues involved in documented pathogenic missense mutations and which residues they change into. Note that many residue changes are not possible from single-nucleotide variants, which accounts for many of the data points of 0 in the heatmap.

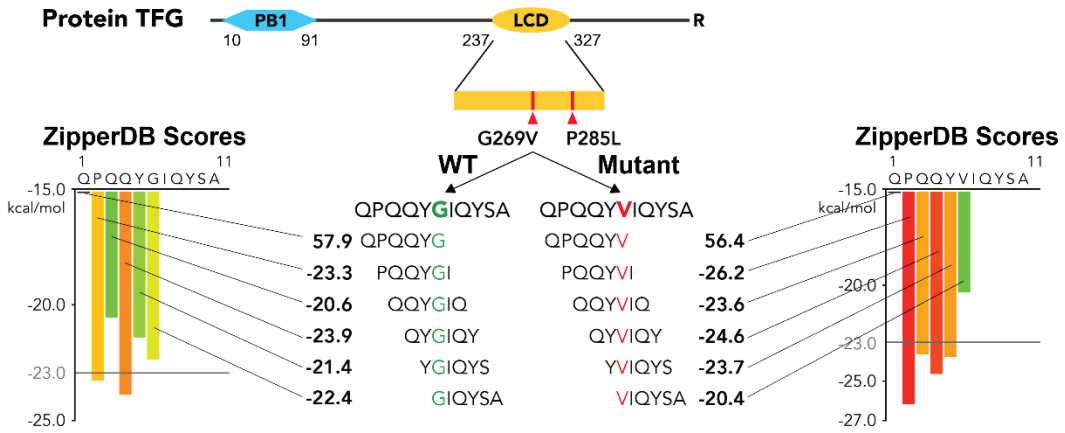
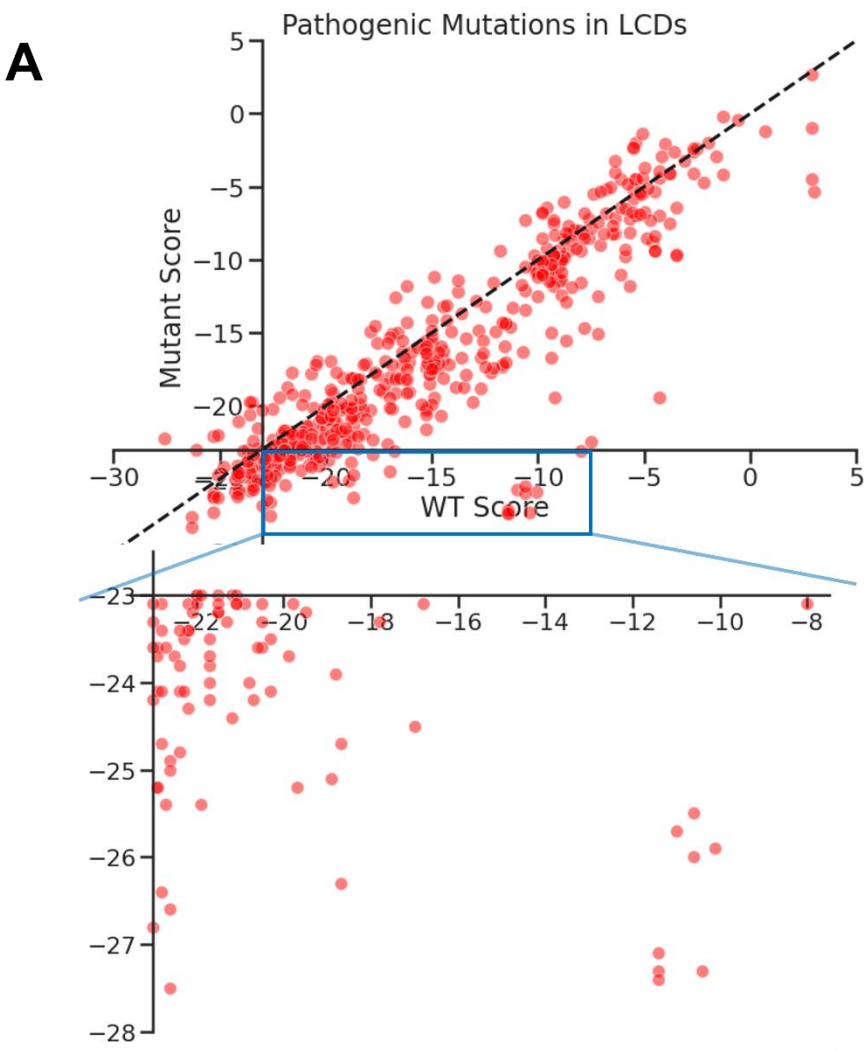


Figure 2-3 Schematic summary of the methodology by which we discovered amyloidogenic mutations. Top: Diagram of Protein TFG which contains a PB1 domain and a low complexity domain. We investigated only mutations in the low complexity domain. Bottom: Analyzing the sequence context of the mutant residue. We calculated Rosetta energy scores using ZipperDB for every hexamer containing the wild-type residue as well as the mutant residue. Each wild-type hexamer is compared to its corresponding mutant hexamer. Wild-type scores greater than -23.0 which correspond to a mutant score less than -23.0 imply greater amyloid propensity and are of the most interest.


A
Pathogenic Mutations in LCDs
Mutant Score
WT Score
5
0
−5
−10
−15
−20
−30
−25
−20
−15
−10
−5
0
5
−22
−20
−18
−16
−14
−12
−10
−8
−23
−24
−25
−26
−27
−28

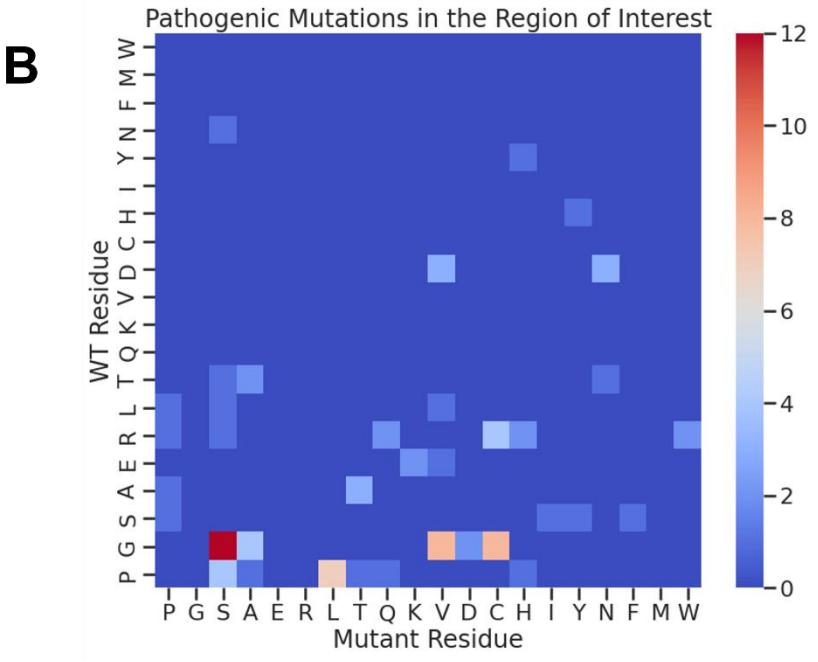

B
Pathogenic Mutations in the Region of Interest
WT Residue
Mutant Residue
P G S A E R L T Q K V D C H I Y N F M W
12
10
8
6
4
2
0

Figure 2-4 Mutant residues in the LCDs throughout the human proteome with greater propensity to form amyloid than their corresponding WT residue. A) Energy scores of wild-type and mutant segments in LCDs computed by ZipperDB. Because each mutation generates 6 possible score pairs, only the score pair which mapped to the “region of interest” (inset) or with the greatest negative change from WT to mutant score is plotted for each mutation. The dashed line shows mutations which do not affect the ZipperDB score. The x and y intercepts are both at -23.0 kcal/mol-of-segment, the ZipperDB threshold for a predicted amyloid-forming steric zipper. Inset contains a zoomed view of the lower right quadrant of the plot which is the “region of interest” containing points corresponding to a wild-type segment with a score above -23.0 kcal/mol-of-segment and a mutant segment with a score below -23.0 kcal/mol-of-segment, indicating a mutation that increases the amyloid propensity. B) Heatmap displaying counts of the kinds of mutational changes in the “region of interest”

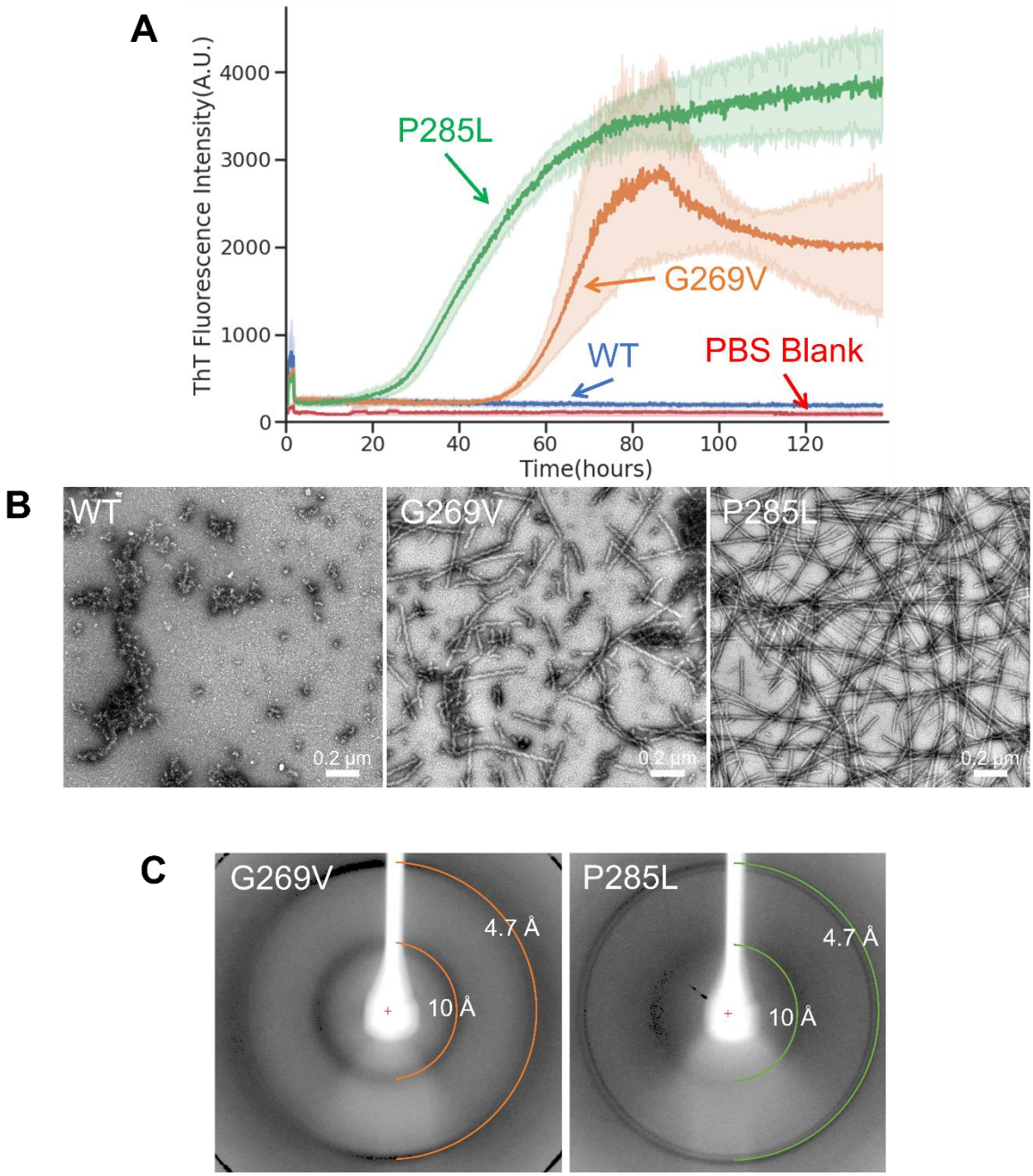


Figure 2-5 Amyloid properties of the LCD of protein TFG. A) Time dependent ThT fluorescence for TFG LCD mutants. G269V and P285L are documented pathogenic mutations of TFG. All constructs are at 50µM concentration in PBS with ThT at 40µM concentration. Each construct has n=6 technical replicates, except the PBS blank which has n=3 technical replicates, and y-axis values represent the mean ThT fluorescence value of all replicates for each construct.

B) Electron micrographs of the samples at the endpoint of the ThT curves. Fibers were present only in the mutant constructs.C) X-ray fiber diffraction of TFG fibers. Rings are present at 4.7 Å and 10 Å spacing with distinct wedges, indicative of cross-β structure

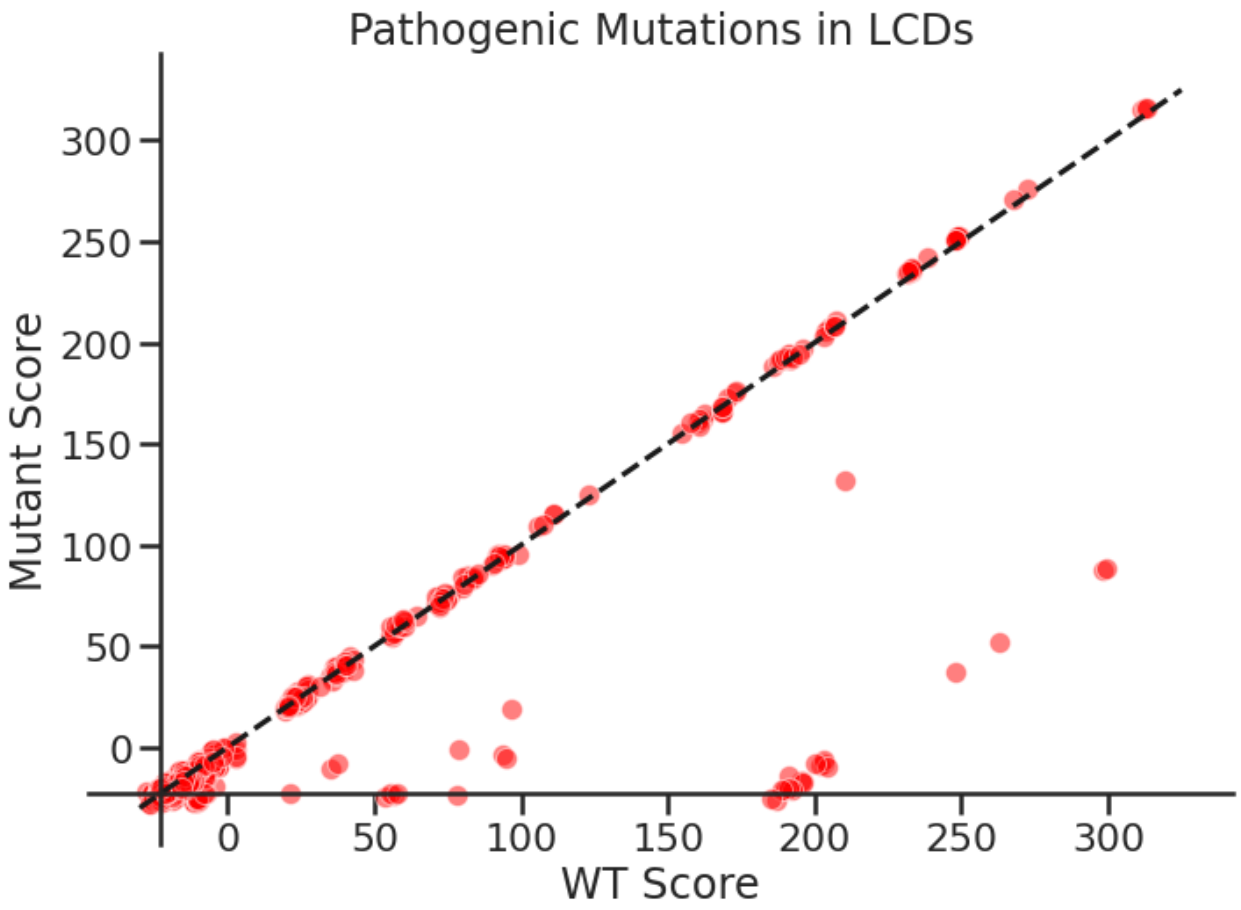


Figure S2-1 Full plot of ZipperDB scores for WT and mutant sequences. The vast majority of mutations with very high positive ZipperDB scores for the WT and/or mutant score are due to sequences containing prolines which are unlikely to form steric zippers.

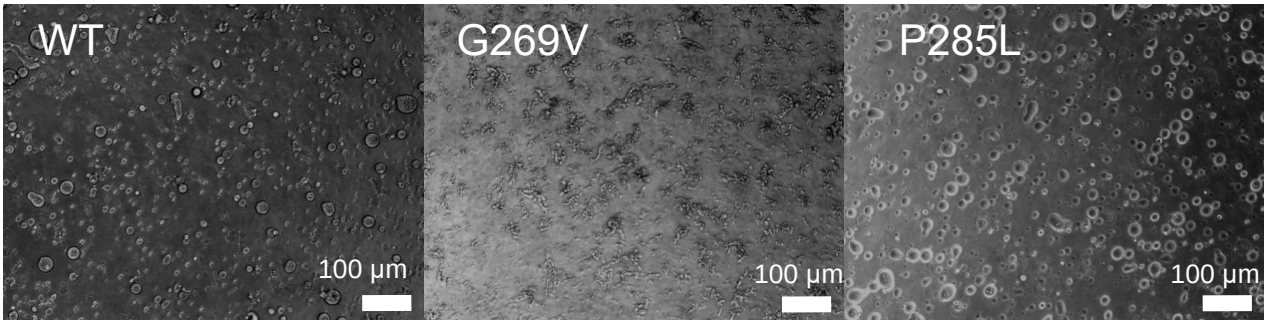


Figure S2-2 Phase separation properties of full length TFG mutants. Constructs were diluted to 10µM concentration in buffer containing 25mM Tris pH 7.4, 150mM KCl, 2.5% v/v glycerol, and 10% w/v PEG 8000. Differential interference contrast (DIC) microscopy images were taken at 10x magnification immediately after adding buffer to the proteins. WT and P285L separated into many large, roughly circular droplets while G269V formed amorphous aggregates.

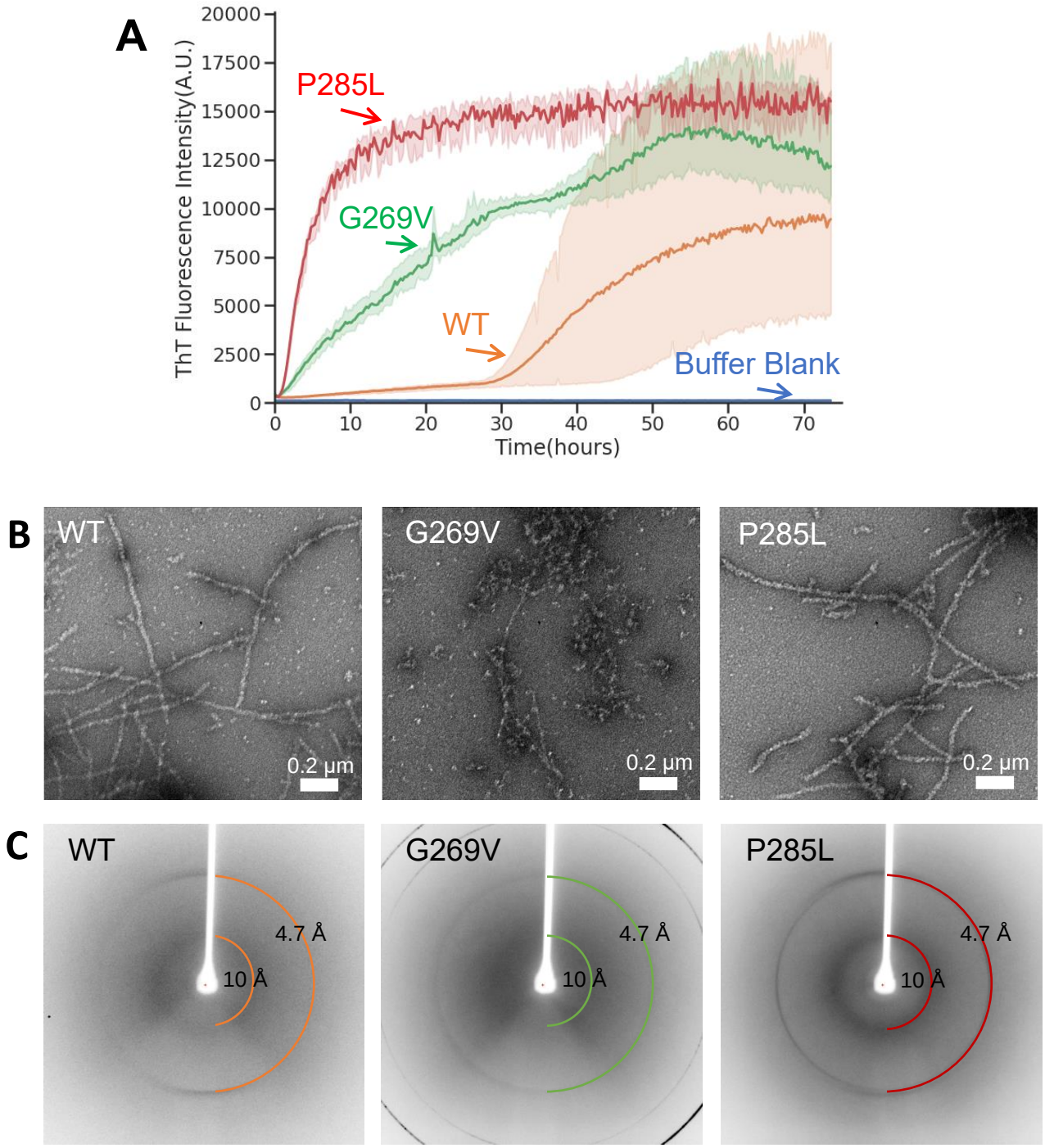


Figure S2-3 Amyloid properties of full-length protein TFG. A) Time dependent ThT fluorescence over time for full length TFG mutants. G269V and P285L are documented pathogenic mutations of TFG. All constructs are at 50µM concentration in buffer containing 20mM Tris pH8 and 150mM NaCl with ThT at 40µM concentration. Each construct has n=3 technical

replicates and y-axis values represent the mean ThT fluorescence value of all replicates for each construct. B) Electron micrographs of the samples at the endpoint of the ThT curves. Fibers were present in all constructs. C) X-ray fiber diffraction of TFG fibers. Rings are present at 4.7 Å and 10 Å spacing with distinct wedges, indicative of cross-β structure

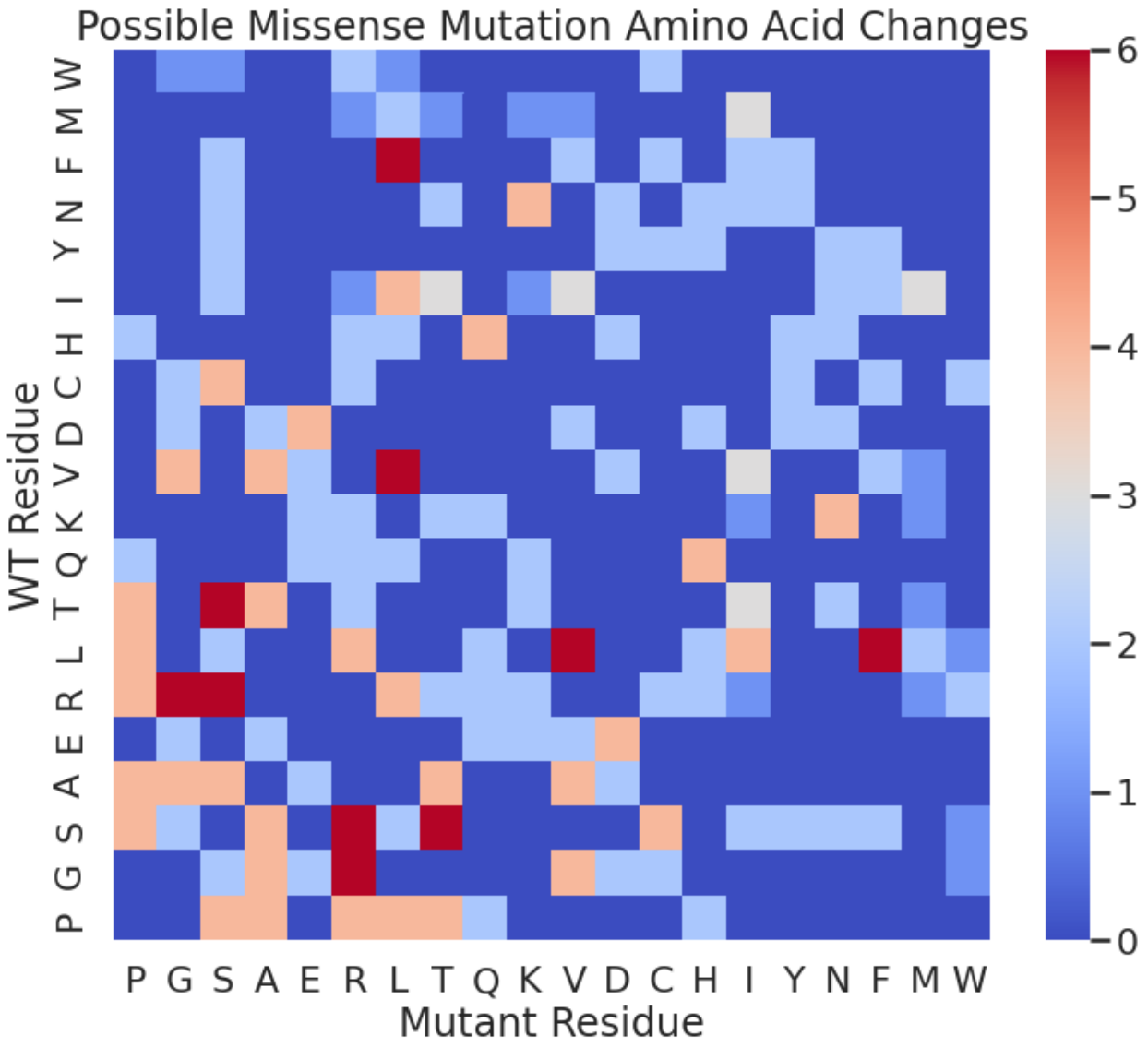


Figure S2-4 Counts of possible missense mutations by single nucleotide variants. The values in each cell represent the number of ways a codon for the WT residue could be mutated into a codon for the mutant residue through single nucleotide variants.

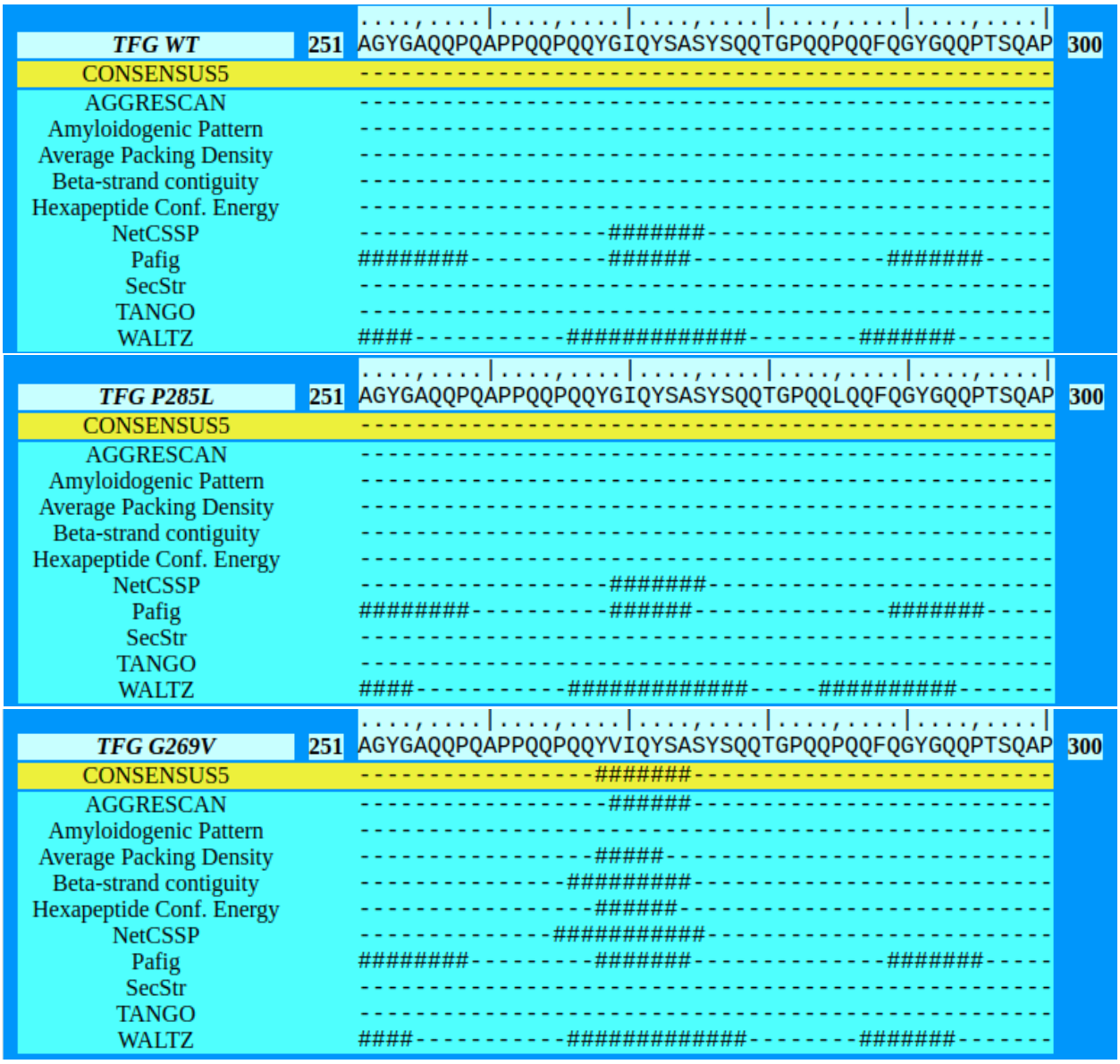


Figure S2-5 Output of AMYLPRED2 for amino acids 251-300 of the TFG WT sequence (top), TFG with the P285L mutation (middle), and TFG with the G269V mutation (bottom). Number signs represent positions predicted to be a part of an amyloidogenic region. The light blue box shows the predicted amyloid regions for each prediction tool incorporated into AMYLPRED2. The yellow box shows regions for which at least 5 out of the 10 prediction methods overlapped in their predicted amyloid segments. Only the G269V mutation had a consensus between enough of the amyloid predictors to cause AMYLPRED2 to have a hit in this region. The only change between

the WT and P285L sequence was a slight change in the output of WALTZ with one of its predicted amyloid segments extending to include the leucine at position 285 in the mutant sequence. Images taken from http://thalis.biol.uoa.gr/AMYLPRED2/

Table S2-1 Mutations in LCDs which alter a non-amyloidogenic WT segment into an amyloidogenic mutant segment. The WT and mutant scores included in the table are representative of the hexamer pair with the most dramatic negative change from the WT to the mutant score. WT and mutant scores are in units kcal/mol. The table is sorted by gene name in alphabetical order.

Table S2-2 Comparison of distribution of mutation types predicted to be amyloidogenic or not for each WT residue. P Values obtained through Fishers exact test. P Values could not be calculated for some initial values where 100% of the mutations were predicted to be amyloidogenic or not. Hypothesis tested is whether there is a difference between the distributions of the mutations that were predicted to be amyloidogenic versus not amyloidogenic.

References

1. Nolan, M., Talbot, K., and Ansorge, O. (2016) Pathogenesis of FUS-associated ALS and FTD: insights from rodent models. *Acta Neuropathol Commun*. **4**, 99

2. Pesiridis, G. S., Lee, V. M.-Y., and Trojanowski, J. Q. (2009) Mutations in TDP-43 link glycine-rich domain functions to amyotrophic lateral sclerosis. *Hum Mol Genet*. **18**, R156-162

3. Kim, H. J., Kim, N. C., Wang, Y.-D., Scarborough, E. A., Moore, J., Diaz, Z., MacLea, K. S., Freibaum, B., Li, S., Molliex, A., Kanagaraj, A. P., Carter, R., Boylan, K. B., Wojtas, A. M., Rademakers, R., Pinkus, J. L., Greenberg, S. A., Trojanowski, J. Q., Traynor, B. J., Smith, B. N., Topp, S., Gkazi, A.-S., Miller, J., Shaw, C. E., Kottlors, M., Kirschner, J., Pestronk, A., Li, Y.

R., Ford, A. F., Gitler, A. D., Benatar, M., King, O. D., Kimonis, V. E., Ross, E. D., Weihl, C. C., Shorter, J., and Taylor, J. P. (2013) Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature*. **495**, 467–473

4. Nott, T. J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowietz, A., Craggs, T. D., Bazett-Jones, D. P., Pawson, T., Forman-Kay, J. D., and Baldwin, A. J. (2015) Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol Cell*. **57**, 936–947

5. Elbaum-Garfinkle, S., Kim, Y., Szczepaniak, K., Chen, C. C.-H., Eckmann, C. R., Myong, S., and Brangwynne, C. P. (2015) The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proc Natl Acad Sci U S A*. **112**, 7189–7194

6. Kato, M., Han, T. W., Xie, S., Shi, K., Du, X., Wu, L. C., Mirzaei, H., Goldsmith, E. J., Longgood, J., Pei, J., Grishin, N. V., Frantz, D. E., Schneider, J. W., Chen, S., Li, L., Sawaya, M. R., Eisenberg, D., Tycko, R., and McKnight, S. L. (2012) Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell*. **149**, 753–767

7. Zhou, X., Lin, Y., Kato, M., Mori, E., Liszczak, G., Sutherland, L., Sysoev, V. O., Murray, D. T., Tycko, R., and McKnight, S. L. (2021) Transiently structured head domains control intermediate filament assembly. *Proc Natl Acad Sci U S A*. **118**, e2022121118

8. Sysoev, V. O., Kato, M., Sutherland, L., Hu, R., McKnight, S. L., and Murray, D. T. (2020) Dynamic structural order of a low-complexity domain facilitates assembly of intermediate filaments. *Proc Natl Acad Sci U S A*. **117**, 23510–23518

9. Murray, K. A., Hu, C. J., Seidler, P., Hughes, M. P., Salwinski, L., Sawaya, M., Pan, H., and Eisenberg, D. S. (2022) Identifying amyloid-related diseases by mapping mutations in low-complexity protein domains to pathologies. *Nat. Struct. Mol. Biol.*

10. Murray, D. T., Zhou, X., Kato, M., Xiang, S., Tycko, R., and McKnight, S. L. (2018) Structural characterization of the D290V mutation site in hnRNPA2 low-complexity-domain polymers. *Proc Natl Acad Sci U S A*. **115**, E9782–E9791

11. Nomura, T., Watanabe, S., Kaneko, K., Yamanaka, K., Nukina, N., and Furukawa, Y. (2014) Intranuclear aggregation of mutant FUS/TLS as a molecular pathomechanism of amyotrophic lateral sclerosis. *J Biol Chem*. **289**, 1192–1202

12. Guenther, E. L., Cao, Q., Trinh, H., Lu, J., Sawaya, M. R., Cascio, D., Boyer, D. R., Rodriguez, J. A., Hughes, M. P., and Eisenberg, D. S. (2018) Atomic structures of TDP-43 LCD segments and insights into reversible or pathogenic aggregation. *Nat Struct Mol Biol*. **25**, 463–471

13. Kato, M., Zhou, X., and McKnight, S. L. (2022) How do protein domains of low sequence complexity work? *RNA*. **28**, 3–15

14. Rambaran, R. N., and Serpell, L. C. (2008) Amyloid fibrils: abnormal protein assembly. *Prion*. **2**, 112–117

15. Merlini, G., Seldin, D. C., and Gertz, M. A. (2011) Amyloidosis: pathogenesis and new therapeutic options. *J Clin Oncol*. **29**, 1924–1933

16. Wootton, J. C., and Federhen, S. (1993) Statistics of local complexity in amino acid sequences and sequence databases. *Computers & Chemistry*. **17**, 149–163

17. UniProt Consortium (2021) UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res*. **49**, D480–D489

18. Amberger, J. S., Bocchini, C. A., Scott, A. F., and Hamosh, A. (2019) OMIM.org: leveraging knowledge across phenotype-gene relationships. *Nucleic Acids Res*. **47**, D1038–D1043

19. Landrum, M. J., Lee, J. M., Benson, M., Brown, G. R., Chao, C., Chitipiralla, S., Gu, B., Hart, J., Hoffman, D., Jang, W., Karapetyan, K., Katz, K., Liu, C., Maddipatla, Z., Malheiro, A., McDaniel, K., Ovetsky, M., Riley, G., Zhou, G., Holmes, J. B., Kattman, B. L., and Maglott, D. R. (2018) ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res*. **46**, D1062–D1067

20. Sawaya, M. R., Hughes, M. P., Rodriguez, J. A., Riek, R., and Eisenberg, D. S. (2021) The expanding amyloid family: Structure, stability, function, and pathogenesis. *Cell*. **184**, 4857–4873

21. Goldschmidt, L., Teng, P. K., Riek, R., and Eisenberg, D. (2010) Identifying the amylome, proteins capable of forming amyloid-like fibrils. *Proc Natl Acad Sci U S A*. **107**, 3487–3492

22. Kedia, N., Arhzaouy, K., Pittman, S. K., Sun, Y., Batchelor, M., Weihl, C. C., and Bieschke, J. (2019) Desmin forms toxic, seeding-competent amyloid aggregates that persist in muscle fibers. *Proc Natl Acad Sci U S A*. **116**, 16835–16840

23. Robinson, D. O., Wills, A. J., Hammans, S. R., Read, S. P., and Sillibourne, J. (2006) Oculopharyngeal muscular dystrophy: a point mutation which mimics the effect of the PABPN1 gene triplet repeat expansion mutation. *J Med Genet*. **43**, e23

24. Scheuermann, T., Schulz, B., Blume, A., Wahle, E., Rudolph, R., and Schwarz, E. (2003) Trinucleotide expansions leading to an extended poly-L-alanine segment in the poly (A) binding protein PABPN1 cause fibril formation. *Protein Sci*. **12**, 2685–2692

25. Johnson, A., Bhattacharya, N., Hanna, M., Pennington, J. G., Schuh, A. L., Wang, L., Otegui, M. S., Stagg, S. M., and Audhya, A. (2015) TFG clusters COPII-coated transport carriers and promotes early secretory pathway organization. *EMBO J*. **34**, 811–827

26. Tsai, P.-C., Huang, Y.-H., Guo, Y.-C., Wu, H.-T., Lin, K.-P., Tsai, Y.-S., Liao, Y.-C., Liu, Y.-T., Liu, T.-T., Kao, L.-S., Yet, S.-F., Fann, M.-J., Soong, B.-W., and Lee, Y.-C. (2014) A novel TFG mutation causes Charcot-Marie-Tooth disease type 2 and impairs TFG function. *Neurology*. **83**, 903–912

27. Ishiura, H., Sako, W., Yoshida, M., Kawarai, T., Tanabe, O., Goto, J., Takahashi, Y., Date, H., Mitsui, J., Ahsan, B., Ichikawa, Y., Iwata, A., Yoshino, H., Izumi, Y., Fujita, K., Maeda, K., Goto, S., Koizumi, H., Morigaki, R., Ikemura, M., Yamauchi, N., Murayama, S., Nicholson, G. A., Ito, H., Sobue, G., Nakagawa, M., Kaji, R., and Tsuji, S. (2012) The TRK-fused gene is mutated in hereditary motor and sensory neuropathy with proximal dominant involvement. *Am J Hum Genet*. **91**, 320–329

28. LeVine, H. (1999) Quantification of beta-sheet amyloid fibril structures with thioflavin T. *Methods Enzymol*. **309**, 274–284

29. Nilsson, M. R. (2004) Techniques to study amyloid fibril formation in vitro. *Methods*. **34**, 151–160

30. Wang, J., Choi, J.-M., Holehouse, A. S., Lee, H. O., Zhang, X., Jahnel, M., Maharana, S., Lemaitre, R., Pozniakovsky, A., Drechsel, D., Poser, I., Pappu, R. V., Alberti, S., and Hyman, A. A. (2018) A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell*. **174**, 688-699.e16

31. Ryan, V. H., Dignon, G. L., Zerze, G. H., Chabata, C. V., Silva, R., Conicella, A. E., Amaya, J., Burke, K. A., Mittal, J., and Fawzi, N. L. (2018) Mechanistic View of hnRNPA2 Low-

Complexity Domain Structure, Interactions, and Phase Separation Altered by Mutation and Arginine Methylation. *Mol Cell*. **69**, 465-479.e7

32. Franzmann, T. M., and Alberti, S. (2019) Prion-like low-complexity sequences: Key regulators of protein solubility and phase behavior. *J Biol Chem*. **294**, 7128–7136

33. Levitt, M. (1978) Conformational preferences of amino acids in globular proteins. *Biochemistry*. **17**, 4277–4285

34. Hughes, M. P., Sawaya, M. R., Boyer, D. R., Goldschmidt, L., Rodriguez, J. A., Cascio, D., Chong, L., Gonen, T., and Eisenberg, D. S. (2018) Atomic structures of low-complexity protein segments reveal kinked β sheets that assemble networks. *Science*. **359**, 698–701

35. Murray, K. A., Evans, D., Hughes, M. P., Sawaya, M. R., Hu, C. J., Houk, K. N., and Eisenberg, D. (2022) Extended β-Strands Contribute to Reversible Amyloid Formation. *ACS Nano*. 10.1021/acsnano.1c08043

36. Li, Y., Yan, J., Zhang, X., and Huang, K. (2013) Disulfide bonds in amyloidogenesis diseases related proteins. *Proteins*. **81**, 1862–1873

37. Göbl, C., Morris, V. K., van Dam, L., Visscher, M., Polderman, P. E., Hartlmüller, C., de Ruiter, H., Hora, M., Liesinger, L., Birner-Gruenberger, R., Vos, H. R., Reif, B., Madl, T., and Dansen, T. B. (2020) Cysteine oxidation triggers amyloid fibril formation of the tumor suppressor p16INK4A. *Redox Biol*. **28**, 101316

38. Li, S. C., Goto, N. K., Williams, K. A., and Deber, C. M. (1996) Alpha-helical, but not beta-sheet, propensity of proline is determined by peptide environment. *Proc Natl Acad Sci U S A*. **93**, 6676–6681

39. 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

40. Brumshtein, B., Esswein, S. R., Sawaya, M. R., Rosenberg, G., Ly, A. T., Landau, M., and Eisenberg, D. S. (2018) Identification of two principal amyloid-driving segments in variable domains of Ig light chains in systemic light-chain amyloidosis. *J Biol Chem*. **293**, 19659–19671

41. Seidler, P. M., Boyer, D. R., Murray, K. A., Yang, T. P., Bentzel, M., Sawaya, M. R., Rosenberg, G., Cascio, D., Williams, C. K., Newell, K. L., Ghetti, B., DeTure, M. A., Dickson, D. W., Vinters, H. V., and Eisenberg, D. S. (2019) Structure-based inhibitors halt prion-like seeding by Alzheimer’s disease-and tauopathy-derived brain tissue samples. *J Biol Chem*. **294**, 16451–16464

42. Scharner, J., Brown, C. A., Bower, M., Iannaccone, S. T., Khatri, I. A., Escolar, D., Gordon, E., Felice, K., Crowe, C. A., Grosmann, C., Meriggioli, M. N., Asamoah, A., Gordon, O., Gnocchi, V. F., Ellis, J. A., Mendell, J. R., and Zammit, P. S. (2011) Novel LMNA mutations in patients with Emery-Dreifuss muscular dystrophy and functional characterization of four LMNA mutations. *Hum Mutat*. **32**, 152–167

43. Keith, J. L., Swinkin, E., Gao, A., Alminawi, S., Zhang, M., McGoldrick, P., McKeever, P., Robertson, J., Rogaeva, E., and Zinman, L. (2020) Neuropathologic description of CHCHD10 mutated amyotrophic lateral sclerosis. *Neurol Genet*. **6**, e394

44. Deng, H.-X., Chen, W., Hong, S.-T., Boycott, K. M., Gorrie, G. H., Siddique, N., Yang, Y., Fecto, F., Shi, Y., Zhai, H., Jiang, H., Hirano, M., Rampersaud, E., Jansen, G. H., Donkervoort, S., Bigio, E. H., Brooks, B. R., Ajroud, K., Sufit, R. L., Haines, J. L., Mugnaini, E., Pericak-Vance, M. A., and Siddique, T. (2011) Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature*. **477**, 211–215

45. Zeitz, C., Forster, U., Neidhardt, J., Feil, S., Kälin, S., Leifert, D., Flor, P. J., and Berger, W. (2007) Night blindness-associated mutations in the ligand-binding, cysteine-rich, and intracellular domains of the metabotropic glutamate receptor 6 abolish protein trafficking. *Hum Mutat*. **28**, 771–780

Chapter 3: Fibril structures of TFG protein mutants validate identification of TFG as a disease-related amyloid protein by the IMPAcT method

Gregory M. Rosenberg, Romany Abskharon, David R. Boyer, Peng Ge, Michael R. Sawaya, David S. Eisenberg

Abstract

We previously presented a bioinformatic method for identifying diseases that arise from a mutation in a protein's low-complexity domain that drives the protein into pathogenic amyloid fibrils. One protein so identified was the tropomyosin-receptor kinase fused gene protein (TRK-fused gene protein or TFG). Mutations in TFG are associated with degenerative neurological conditions. Here we present experimental evidence that confirms our prediction that these conditions are amyloid related. We find the low-complexity domain of TFG containing the disease related mutations G269V or P285L forms amyloid fibrils, and we have determined their structures using cryo-electron microscopy. These structures are unmistakably amyloid in nature and confirm the propensity of the mutant TFG low-complexity domain to form amyloid fibrils. Also, despite resulting from a pathogenic mutation, the fibril structures bear some similarities to other amyloid structures which are thought to be non-pathogenic and even functional, but there are other factors that support these structures' relevance to disease, including increased propensity to form amyloid compared to the WT sequence, structure-stabilizing influence from the mutant residues themselves, and double-protofilament amyloid cores. Our findings elucidate two potentially disease-relevant structures of a previously unknown amyloid and also provide new perspectives on factors that contribute to the pathogenicity of amyloid fibers.

Introduction

Previously we developed a method to search for pathogenic mutations which induce amyloid aggregation of an otherwise functional protein(1). We have since named this method the Identification of Mutations Promoting Amyloidogenic Transitions method, or the IMPAcT method. Briefly, the method entails comparing the computed amyloidogenic propensity of mutant protein sequences to their wild-type (WT) counterparts and retaining those mutations which generate an amyloidogenic sequence from a non-amyloidogenic WT sequence. With the assumption that low-complexity domains (LCDs) would be prime candidates for amyloidogenic gain of function due to commonly functioning in protein self-association and often being intrinsically disordered, we applied this method to every LCD in the human proteome and identified many candidate mutations in proteins not previously shown to be amyloidogenic. We demonstrated that two of these mutations, both in the TRK-fused gene protein (TFG), do indeed induce fibrillar aggregates with amyloid characteristics.

TFG functions in the early secretory pathway by regulating the transport of secretory vesicles between the endoplasmic reticulum (ER) and ER-Golgi intermediate compartments (ERGIC)(2). More specifically, TFG brings together secretory vesicles from the ER by interacting with the conserved coat protein complex II (COPII) on the surface of the vesicles(3). This interaction persists until the vesicles tether and fuse with the ERGIC membrane. Thus, the TFG-COPII interaction is crucial for preventing the diffusion of COPII coated vesicles away from the ER-ERGIC interface, an event which would derail protein secretion. To perform its role in clustering vesicles, TFG self-assembles into homomeric cup-shaped octamers via interactions encoded in its N-terminal Phox and Bem1p (PB1) domain and a subsequent coiled-coil motif. However, the majority of TFG's 400 amino acids compose an intrinsically disordered region. Within this region is an LCD (residues 237-327), rich in proline, tyrosine, and glutamine (**Figure 3-1A**), which functions to further polymerize these octamers into larger assemblies(2).

Mutations in the TFG LCD are associated with degenerative neurological disorders including Charcot-Marie-Tooth disease type 2 (CMT2) and hereditary motor and sensory neuropathy with proximal dominant involvement (HMSN-P)(4, 5). CMT2 is a group of axonal neuropathies, typically with autosomal dominant inheritance, characterized by progressive axonal degeneration and deterioration of muscle strength leading to distal wasting, weakness, sensory loss with reduced tendon reflexes, and foot deformity(6). HMSN-P is also a slow, progressive axonal degeneration disease with autosomal dominant inheritance and is characterized by painful muscle cramps and muscle weakness in early stages, sensory disturbances (dysesthesia, hypesthesia, and hypopallesthesia), and, in late stages, virtually no muscle strength in the extremities making patients unable to walk and bedridden with impaired respiration similar to ALS(7). The mutations associated with these diseases have also been shown to lead to aggregation of TFG in cells(4, 5).

Here we determined the cryo-electron microscopy (cryoEM) structure of the amyloid fibers produced by the TFG LCD with both of the disease mutants identified by the IMPAcT method, G269V and P285L, which are associated with CMT2(4) and HMSN-P(5), respectively. These structures further validate the amyloid propensity of this protein, elucidate a potential molecular basis for the effects of these mutations on disease, and add nuance to the paradigm of structural features which evince either pathogenic or non-pathogenic amyloids.

Results

Fibril Formation of the TFG LCD

We expressed and purified the LCD of TFG (residues 237-327) conjugated to a molecule of mCherry, for enhanced solubility, with the WT TFG sequence (mC-TFG-LCD), containing the

G269V mutation (mC-TFG-LCD-G269V), or containing the P285L mutation (mC-TFG-LCD-P285L). We shook the constructs at 37˚C for 138 hours in the presence of the amyloidophilic dye Thioflavin T (ThT) to monitor amyloid fibril formation *in vitro*. mC-TFG-LCD-G269V and mC-TFG-LCD-P285L formed ThT-positive aggregates while mC-TFG-LCD did not (**Figure 3-1B**). The mutant protein fibrils were visualized with negative stain electron microscopy (EM) (**Figure 3-1C,E**) and the cross-β nature of the fibrils was confirmed via X-ray diffraction of aligned and dried fibrils (**Figure 3-1D,F**).

Cryo-EM Structure of mC-TFG-LCD Mutant Fibrils

To understand the relationship of the G269V and P285L disease mutations to the formation of fibrils and further confirm their amyloid nature, we determined the cryo-EM structure of the mC-TFG-LCD-G269V fibrils to a resolution of 2.8 Å (**Figure 3-2A**) and the mC-TFG-LCD-P285L fibrils to a resolution of 2.6 Å (**Figure 3-2B**). Data collection and refinement statistics for mC-TFG-LCD-G269V and mC-TFG-LCD-P285L are summarized in **Table 3-1**.

The mC-TFG-LCD-G269V fibrils have a uniform morphology consisting of two protofilaments of the same sequence (although one has three more residues resolved than the other) but slightly different structures and interacting through a class 4 zipper interface(8, 9) (parallel, face-to-back, up-down packing) (**Figure 3-2A**). The fibril has a pitch of 578 Å, a left-handed helical twist of -3.07º, and a helical rise of 4.93 Å. As expected from the initial biochemical evidence, these features are all consistent with identifying this structure as an amyloid fibril. The fibril core contains two protofilaments, one with a wide core of 29 residues (P262- G290) and one with a narrow core of 26 residues (P265-G290) (3 fewer residues resolved) (**Figure 3-2A**). Each protofilament consists of a single dagger-shaped β-arch and constitutes less than 30% of the LCD (**Figure 3-1A**).

The disease-relevant mutation G269V is present and buried in the interiors of both protofilaments. V269 resides at the center of the longest β-strand of each protofilament (**Figure 3-2A**) and fits tightly with side chains from the opposite arm of the β-arch. Thus, V269 excludes water from the interior of the protofilaments and contributes to a heterotypic steric zipper. Moreover, as a participant in the only hydrophobic interaction in the interior of each protofilament, V269 contributes strongly to the energetic stability of the β-arch and at the protofilament interface.

The mC-TFG-LCD-P285L fibrils have two distinct polymorphs: a narrow, untwisting fibril (making up the minority of fibrils) whose structure could not be determined, and the major polymorph which consists of two identical protofilaments related by a pseudo 2_1 screw axis forming a class 1 zipper interface (parallel, face-to-face, up-up packing) (**Figure 3-2B**). The fibril has a pitch of 877 Å, a left-handed helical twist of 179.02^{o}, and a helical rise of 2.39Å. As expected, these features align with the biochemical evidence of amyloid fibrils. The protofilaments both consist of 35 residues (P265-A299) and form shapes reminiscent of the lower-case letters “d” and ”p”.

The disease-relevant mutation P285L is present and buried in the protofilament interior. L285 is a member of the longest β-strand of the protofilament core, fills in the space left by the glycine residue across from it, and forms a hydrophobic interaction with I270. Water molecules resolved in the interior of the fiber core reveal that L285 forms a barrier between solvent-accessible surface area and a glutamine zipper. Q283 is interacting with 3 water molecules while the zipper made up of Q267, Q287, and Q289 are protected by L285 from solvent infiltration.

Structural Differences are Directly Attributable to Effects of Mutant Residues

The fibril cores formed by mC-TFG-LCD-G269V and mC-TFG-LCD-P285L have distinctly different folds. The most consistent feature between the two is the glutamine steric zipper formed by residues Q267, Q287, and Q289 (**Figure S3-1**). Although, only the narrow protofilament of the mC-TFG-LCD-G269V fibril exhibits this structural motif, while the wide protofilament has a different arrangement of these glutamine residues in relation to each other. Besides this similarity, the portion of the core from G/V269 to P/L285 has a distinct shape in each structure and also the mC-TFG-LCD-P285L fibril core has several extra C-terminal residues resolved (Y291-A299) and these residues form both intra- and inter-protofilament interactions.

The difference in shape of the 269-285 segment is due to the different residues terminating this section in either structure. In the mC-TFG-LCD-G269V structure, there is a valine at position 269 which participates in a β-strand from Q266 to S273, but in the mC-TFG-LCD-P285L structure there is a glycine at position 269, which allows more flexibility in the peptide backbone, creating a slight kink which disrupts the β-strand and flips I270 from solvent-facing to the protofilament interior allowing it to interact with L285, the mutant residue. This kink changes the shape of this segment from a dagger-like fold to a wider, flattened out turn which has a solvent-accessible channel. The G269V mutation contributes to a longer β-strand which constrains the core to a more dagger-shaped fold, while the P285L mutation stabilizes the kink centered on G269 by introducing a hydrophobic interaction with I270.

The shape difference of the 269-285 segment also contributes to the presence or absence of Y291-A299 in the resolved fibril core. This is because, in the mC-TFG-LCD-P285L structure, residues T296-Q298 form a short polar zipper with residues Q278-T280 on the opposite protofilament. Q278-T280 are constituents of a short β-strand which, in the mC-TFG-LCD-P285L structure, is almost orthogonal to the central protofilament interface, while in the mC-TFG-LCD-

G269V structure it is almost parallel to the protofilament interface. This makes the 296-298 segment available to interact with residues 296-298 only in the P285L structure because of the shape of the entire 269-285 segment, and this shape is attributable to which mutation the protein has.

The last major difference between the two structures is that the mC-TFG-LCD-G269V structure has asymmetric protofilaments not related by a screw axis, while the mC-TFG-LCD-P285L structure has symmetrical protofilaments related by a pseudo 2_1 screw axis. This symmetry difference is difficult to explain, but a difference in the structures of the protofilament interfaces can actually be attributed to the difference in mutations. The narrow protofilament of the mC-TFG-LCD-G269V fibril has two proline residues in the segment which makes up the protofilament interface. Looking down the fibril axis, these proline residues seem to be in a relatively straight segment resembling a typical β-strand, but when viewed orthogonal to the fibril axis this segment has some obvious buckling causing the peptide backbone to have a wave-like shape (**Figure 3-3**). This warp in the backbone shape, along with slight opposite angling of the planes of each individual protofilament, cause the zipper interface to be somewhat staggered in the mC-TFG-LCD-G269V fibril despite not being related by a screw axis. The G269V mutation may stabilize this unusual configuration by forming a hydrophobic interaction with P285 and thus forcing the P282-P285 segment to conform to a β-sheet-like interaction in one dimension and warp in another dimension. The mC-TFG-LCD-P285L fibril does not exhibit this buckle in the peptide backbone because the corresponding proline residues (P282 and P285) are either involved in a turn which does not exist in the mC-TFG-LCD-G269V fibril (P282 participates in a turn) or mutated to a different residue (P285 is mutated to leucine). Thus, both structures exhibit a staggered zipper interface, but the staggering is related to two different structural mechanisms which can be indirectly linked to the mutant residues.

Energetic Analysis Reveals Resemblance to Reversible, Functional Amyloids

We calculated the atomic solvation energy of the mC-TFG-LCD-G269V and mC-TFG-LCD-P285L fibrils using the coordinates of the ordered fibril cores (**Figure 3-4A,C**) and compared their values to other amyloid fibril structures. The solvation energy per residue is -0.32 kcal mol^{-1} for mC-TFG-LCD-G269V and -0.33 kcal mol^{-1} for mC-TFG-LCD-P285L which are both similar to the reversible hnRNPA2 WT LCD fibril (-0.34 kcal mol^{-1} per residue)(10). The solvation energy per layer is -17.6 kcal mol^{-1} for mC-TFG-LCD-G269V and -23.2 kcal mol^{-1} for mC-TFG-LCD-P285L which are also similar to the hnRNPA2 WT LCD (-19.5 kcal mol^{-1})(10). Other amyloid structures with similar calculated energies belong to the FUS-LCD (PDB ID: 5w3n; energy per residue: -0.20 kcal mol^{-1}; energy per layer: -12.2 kcal mol^{-1})(11), the fungal prion Het-s (PDB ID: 2rnm; energy per residue: -0.30 kcal mol^{-1}; energy per layer: -18.8 kcal mol^{-1})(11), and a heparin-induced recombinant tau fiber “tau 4R snake” (PDB ID: 6qjh; energy per residue: -0.30 kcal mol^{-1}; energy per layer: -17.7 kcal mol^{-1})(11). The FUS-LCD structure is ostensibly functional and reversible(12), but Het-s, while functional, is not reversible(13), and tau 4R snake does not seem to be functional, but it is not representative of the structures of tau fibers extracted from patient brains(14) and its reversibility is unknown. Definitively pathogenic structures tend to have significantly more stable solvation energy values, such as for TTR fibers (PDB ID: 6sdz; energy per residue: -0.68 kcal mol^{-1}; energy per layer: -62.1 kcal mol^{-1})(11) and serum amyloid A (PDB ID: 6mst; energy per residue: -0.65 kcal mol^{-1}; energy per layer: -68.8 kcal mol^{-1})(11). Thus, based on solvation energy, we would group the mC-TFG-LCD-G269V and mC-TFG-LCD-P285L amyloid fibers with functional and/or reversible amyloids rather than pathogenic ones.

One explanation why our energetic algorithm predicts relatively poor stability for the fibril cores is the abundance of polar residues that form the inter and intramolecular steric zippers in each

structure, mainly glutamine residues (**Figure 3-4B,D**). Pathogenic amyloids are better stabilized by hydrophobic interactions in their cores than by polar interactions(11). The dominance of glutamine in the fibril cores resembles the glutamine-rich β-arches of the reversible, functional drosophila amyloid Orb2(15) (**Figure S3-2**). However, glutamine zippers are also found in pathogenic amyloids such as in some TDP-43 fibrils(16) and aggregates of huntingtin(17). In both mutant fibrils, the glutamine side chains in the interior of each protofilament are positioned such that they form hydrogen bonds with the glutamines above and below them along the fibril axis (**Figure S3-3**). Also, in the mC-TFG-LCD-G269V fibril, Q278 of the narrow protofilament and Q264 of the wide protofilament form a hydrogen bond between the two protofilaments in addition to inter-layer hydrogen bonds (**Figure S3-3A**). In the mC-TFG-LCD-P285L fibril, an intra-protofilament hydrogen bond ladder is formed between Q286 and Q294 (**Figure 3-3B**). Interestingly, in the mC-TFG-LCD-P285L structure, one of the interior glutamine residues does not participate in interlayer hydrogen bonding because it is being solvated by water molecules. Overall, this hydrogen bonding activity may compensate for the destabilizing effect of their hydrophilicity.

Another destabilizing feature of the mutant fibrils is the abundance of proline residues within the fibril cores. Prolines weaken β-sheet formation owing to proline's inability to donate a hydrogen bond to a backbone carbonyl oxygen in an adjacent layer of the fibril—a consequence of proline's unique side chain being covalently bonded to the peptide backbone amide. In the mC-TFG-LCD-G269V fibril, the wider protofilament core contains four prolines and the narrower one contains three; one of these prolines (P285 in both protofilaments) is directly across from the mutant valine residue, participating in a hydrophobic interaction, thus the valine is compensating for the structural weakness incurred by the presence of the proline residue. In the mC-TFG-LCD-P285L fibril, both protofilaments contain three prolines and one of them (P282) participates in the

protofilament interface and thus is able to somewhat contribute to the structural stability of the fibril. The TFG LCD as a whole is relatively proline-rich, containing 14 prolines among its 90 amino acids. The presence of these prolines within the fibril core would significantly destabilize this fibril due to its lack of hydrogen bonding capability(18) but may be compensated for by the hydrogen bonding of the glutamine side chains, as described above.

One feature that distinguishes the mutant TFG fibrils from reversible amyloids, however, is the presence of two protofilaments. Most reversible amyloid structures consist of only a single protofilament which may make them more amenable to disassembly(11). The double-protofilament structure also bolsters stability by compensating for the meager energy per chain (~-9.0 kcal mol^{-1} for mC-TFG-LCD-G269V and ~-11.5 kcal mol^{-1} for mC-TFG-LCD-P285L) which is poor due to the previously mentioned energetic penalty of desolvating glutamine side chains as well as the relatively low number of amino acids in each protofilament. The bundling of two protofilaments is a feature more common to pathogenic amyloids, and every structure of an amyloid known to be conditionally labile has a single-protofilament core(11).

To empirically test the stability of the mutant TFG fibrils, we heated aliquots of the fibrils at a range of temperatures from 40°C to 100°C for 10 minutes each. For the mC-TFG-LCD-G269V fibrils, we found that the fibrils remained present after being heated to 70°C, but dissociated after being heated to 80°C (**Figure 3-5A**). For the mC-TFG-LCD-P285L fibrils, fibrous aggregates were visible up to 80°C and disappeared only at 90°C, although even at 60°C the fibrils appeared bristly (**Figure 3-5B**). The same trend was observed in response to treatment with varying concentrations of guanidine hydrochloride (GdnHCl), a chemical denaturant (**Figure S3-4**). For mC-TFG-LCD-G269V, fibrils began dissociating but were still abundant in 2M GdnHCl, fibrils were extremely sparse in 2.5M GdnHCl, and all fibrils had disappeared in 3M GdnHCl (**Figure S3-**

4A,B). For mC-TFG-LCD-P285L, fibrils remained abundant, but bristly, in 2M GdnHCl, appeared further dissociated but still abundant in 2.5M GdnHCl, but extremely sparse in 3M GdnHCl (**Figure S3-4C,D**). These results suggests that both mutant fibrils are relatively stable and difficult to dissociate.

Asymmetric C5 Hydrogen Bonding Capability

One feature of the mC-TFG-LCD-G269V fibril, but not the mC-TFG-LCD-P285L fibril, which straddles the line between resembling reversible or irreversible amyloid fibrils is its capability to form C5 hydrogen bonds, but only within one of its two protofilaments. C5 hydrogen bonds are intraresidue interactions between the amide proton and carbonyl oxygen of an individual amino acid(19). These bonds are favored in extended β-strand conformations called EAGLS (extended amyloid-like glycine-rich low-complexity segments) where |φ| and |ψ| angles are both > 150˚(20). The mC-TFG-LCD-G269V fibril exhibits two residues which conform to this extended (unpleated) β-strand conformation: S273 and Q278, but only in the wide protofilament; the corresponding residues in the narrow protofilament exhibit backbone torsion angles consistent with a pleated β-strand (**Figure S3-5**). C5 hydrogen bonds would compete with interlayer hydrogen bonding along the fibril axis, which can contribute to the reversibility of fibrils which exhibit extended β-sheets by reducing the energetic penalty of amyloid dissociation(20). It is thus an important feature that only one of the mC-TFG-LCD-G269V protofilaments has C5 hydrogen bonding capabilities, and this difference between the two protofilaments also contributes to their slight differences in shape, with the narrow protofilament being slightly more curved and having the positions of two glutamine side chains (Q267 and Q287) swapped relative to the wide protofilament. Despite the fiber-destabilizing effects of C5 hydrogen bonds, since this feature is not common to both protofilaments, it is not clear how much it contributes to reversibility or irreversibility of this mutant TFG LCD fibril in particular.

G269V and P285L Mutations Contribute to Structural Stability

Since these amyloid fibril structures result from particular pathogenic mutations (G269V and P285L) and the WT sequence did not form fibrils, it is relevant to ascertain how these amino acid changes contribute to the formation of these particular structures. In regard to the G269V mutation, valine preferentially forms β-strand secondary structure and so contributes to the stability of steric zippers(21). In the mC-TFG-LCD-G269V structure, V269 forms a hydrophobic interaction with the proline residue across from it (P285) and this is the only hydrophobic interaction in the interior of the individual protofilaments. In regard to the P285L mutation, leucine is not particularly prone to β-strand secondary structure, more likely participating in α-helix secondary structure, but it is much more amenable to being in a β-strand than the WT proline residue. Leucine is also capable of interlayer hydrogen bonding of the peptide backbone while proline is not. In the mC-TFG-LCD-P285L structure, L285 forms a hydrophobic interaction with the isoleucine residue across from it (I270) which helps stabilize the fold of the protofilament. Since hydrophobic interactions are more stabilizing for amyloid fibrils than polar interactions, the hydrophobic interactions formed by the mutant residues in both structures may be integral to the formation and persistence of these structures.

To understand why the WT sequence may not form these fibril structure, we modeled the WT sequence into the structure of the mC-TFG-LCD-G269V fibril and the mC-TFG-LCD-P285L fibril (**Figure S3-6**). When we do this with the mC-TFG-LCD-G269V fibril, we see that the potentially important hydrophobic interaction between residues 269 and 285 that the mutant valine residue participated in is abolished and a gap is open where water can fit and potentially destabilize the adjacent glutamine zippers (**Figure S3-6A**). The missing piece of the steric zipper at the core of the two protofilaments is thus detrimental to the formation of the fibril structure we observe for the

G269V mutant. As for the mC-TFG-LCD-P285L structure, when the mutant leucine is replaced with the WT proline, P285 may still be able to form a hydrophobic interaction with I270, but not in such a way that blocks solvent from accessing the adjacent glutamine zipper like leucine is able to do (**Figure S3-6B**). This gap may preclude the mC-TFG-LCD-P285L structure from forming from the WT sequence, similar to the mC-TFG-LCD-G269V structure. These considerations explain the relevance of the mutations to the observed structures and why they are not observed for the WT sequence.

Discussion

Validation of the IMPAcT method

The IMPAcT method identifies amyloidogenic mutations by comparing the amyloid propensity of wild-type and mutant protein sequences(1). Mutations which convert a nonamyloidogenic sequence to amyloidogenic are considered hits. The fibril structures presented in this work are formed by a protein with either of two of the mutations identified by the IMPAcT method as being amyloidogenic. The presence of the mutant residues in both of the resolved fibril cores is an indication of the IMPAcT method's ability to predict these sequences' amyloid propensities, and more generally the efficacy of the IMPAcT method to identify previously unrecognized amyloid-related conditions. Noting this, the IMPAcT method identified many other mutations as amyloidogenic in proteins which have not previously been recognized as amyloidogenic. Some of the other interesting candidate proteins are LMNA (mutations: G602S and R624H) and CHCHD10 (mutation: P34S), which have been shown to aggregate when mutated(22, 23), and KRT74 (mutation: D482N) since other keratin proteins are amyloidogenic(24–27) and the KRT74 mutation is associated with hypotrichosis of the scalp(28) which is a disorder than is sometimes associated with amyloidosis of another protein called corneodesmosin(29).

The IMPAcT method can also be repeated with different parameters to obtain even more candidates. For example, in our usage we applied the method to proteins with LCDs and to mutations only within LCDs, but this could be expanded to a different subset of the human proteome such as all proteins with intrinsically disordered regions or even potentially the entire human proteome. Also, the method employs the ZipperDB database which bases its scoring method on the energetic fit of a protein sequence to a particular steric zipper structure. The structural template which ZipperDB uses can be modified to different archetypal steric zipper classes or even a custom steric zipper structure. Our validation of the initial implementation of the IMPAcT method is encouraging for the veracity of its other predictions and for the merit of implementing the method in the future with a broader scope or modified parameters.

Functional or Pathogenic Amyloids

Since the fibril structures of TFG are derived from protein sequences containing disease-associated mutations, it may have been expected that the structures would resemble pathogenic amyloids. However, what we see instead are many similarities to reversible and/or functional amyloids. These similarities are comparable values of solvation energy, polar steric zippers, extended β-strands, and the presence of only a single polymorph(10). Despite these similarities, the structures do share similarities to pathogenic amyloids: multiple protofilaments and a near lack of kinked segments called LARKS(30). On top of this, some of those features associated with reversible and/or functional amyloids are not unique to reversible and/or functional amyloids (poor solvation energy in irreversible Het-s fibers; polar steric zippers in TDP-43 fibers) and not all functional amyloids are reversible (Het-s, RIPK1/3(31), Pmel17(32)). Also, it is conceivable that a reversible amyloid fibril can be pathogenic. This distinction between functional and reversible amyloids is important because features that suggest a fibril structure belongs to one of these groups do not compound to suggest it belongs to both groups. For example, the C5 hydrogen

bonding capability of mC-TFG-LCD-G269V could be considered evidence of reversibility, but this does not automatically make it evidence of functionality as well.

Functional amyloids, reversible or not, have associated cellular mechanisms that have evolved to manage their aggregation and maintain their functions(33). So, cellular machinery may not be as well-equipped to mount a response to amyloid fibrils made from a mutant protein which does not form them as part of its normal function, regardless of the potential reversibility of the fibril. Since TFG's normal function does not seem to require the formation of amyloid-like fibrils, the mutant fibrils' commonalities with functional or reversible amyloids may be misleading when judging their pathogenicity.

Another consideration is how much these mutant structures reflect potentially transient structures formed by the LCD during the normal functioning of TFG. The protein has been shown to be capable of phase separation(1) and its function has been proposed to require the formation of a membraneless compartment between the ER and ERGIC in which COPII-coated vesicles move through(3). Phase separation behavior and the formation of membraneless organelles, especially when facilitated by intrinsically disordered regions of proteins, are thought to involve the otherwise disordered regions taking on secondary structure which includes hydrogen bonding between β-strands(18). Formation of amyloids is mediated by the formation of expansive hydrogen bond networks between layered β-strands, so the ability of TFG to form amyloid fibers supports the proposition that the native function of its LCD involves the formation of β-strands connected by hydrogen bonds. Our observation that TFG forms amyloid fibrils is in line with numerous other examples of functional phase-separating proteins which also form amyloid fibrils(11, 12, 34, 35), which contribute to the same conclusion that β-strand secondary structure and hydrogen bonding are important for liquid-liquid phase separation of some proteins.

Potential Mechanisms of Pathogenicity

The underlying reason for the association between amyloid fibrils and disease is not fully understood. Amyloid fibrils themselves may be cytotoxic(36, 37), or oligomeric intermediates between monomers and fibers may be the pathogenic species(38, 39). Indeed, both could be true, with some diseases resulting from either aggregation state, depending on the protein or mutations involved. If fibrils of mutant TFG are able to dissociate, it is less likely that fiber-related cytotoxicity is the basis of the mutation's disease-association. However, regardless of the reversibility of the fibrils, if the mutant protein preferentially forms an amyloid fibril rather than its functional structure, this would manifest indistinguishably from a more mundane loss-of-function mutation. TFG knockout mouse models develop neuromuscular junction degeneration(40), so it is likely that even if the only effect of the TFG mutations is a preference for an inert or reversible amyloid fibril, this would be sufficient to explain the observed disease state. This is in addition to the possibility that the mutant TFG fibrils are indeed irreversible in cells and cytotoxic. There are multiple ways the formation of these fibrils could cause disease, and elucidating the actual mechanism requires more biochemical study of the protein, but loss of function due to aberrant amyloid formation seems sufficient to explain the mutation's role in disease.

mCherry is Unlikely to Influence Fibril Core Structure

The addition of an mCherry tag to the LCD constructs was intended to ensure solubility and discourage amorphous aggregation, similar to previous work on similar constructs(10). This tag was not cleaved off at any point during experimentation, so it is relevant to ask if this tag affected the formation of the fibrils we observed. It is unlikely that this tag significantly influenced the formation of fibers based on three lines of reasoning. First, the WT construct was not able to form any fibers, so it is unlikely that the mCherry is what drove fibril formation in the mutant construct.

In fact, mCherry may have increased solubility and discouraged self-interaction relative to the full-length protein since the full-length protein has a larger disordered region as well as a PB1 domain which functions in self-oligomerization of TFG. This evidence suggests that this construct would be selective for fiber structures with greater energetic favorability than the untagged full-length protein since the full-length protein may be more prone to interact with itself and aggregate. Second, the end of the mCherry tag is 34 amino acids away from the residues which make up the fibril cores in our structures. These 34 amino acids include a short 6-residue linker and 28 residues of the TFG LCD which is intrinsically disordered (25 for the wide protofilaent of mC-TFG-LCD-G269V, making that sequence 31 amino acids away from mCherry). This distance allows enough flexibility for the interaction of the residues included in the fibril cores to be uninfluenced by the mCherry tag. Third, the mCherry tag is completely unresolved in the final structures. This means there is no single preferred interactions between the mCherry and the fibril cores.

Summary

Here we present two near-atomic resolution structures of the amyloid fibril core formed by the LCD of TFG with either of two pathogenic mutations. These structures are a further validation of the prediction by the IMPAcT method that these particular disease-related mutations are amyloid-promoting. The mutated residues were found to facilitate the stability of the fibril core structures and thus fibril formation may be an explanation for their relevance in disease. The fibrils also have similarities to functional and reversible amyloid fibrils, implying that these features are not sufficient to explain the difference between pathogenic amyloids and functional/reversible amyloids.

Materials and Methods

Low-Complexity Region Prediction

The amino acid sequence of TFG was evaluated for low complexity using SEG(41) with default settings: window length = 12, trigger complexity 2.2, extension complexity 2.5. A sequence was determined to be a low-complexity domain if it contained at least 35 residues scored as low complexity with at most 5 interrupting non-low-complexity residues.

Protein Expression and Purification

Recombinant TFG(237-327) for the WT, G269V, and P285L forms was purified using a pHis-parallel-mCherry vector, using a previously described method(35). Briefly, protein was overexpressed in BL21(DE3) Gold E. coli cells. Cultures were grown to an OD600=0.4-0.8, then induced with 0.5M IPTG overnight. Cells were pelleted by centrifugation and the clarified lysate was purified by Ni-NTA columns followed by size exclusion chromatography and dialyzed into PBS.

In Vitro Aggregation Assay

Wild-type and mutant TFG LCD were diluted to 50μM in 1X PBS containing ThT at 40μM to a final volume of 150μL in black Nunc 96-well optical bottom plates (Thermo Scientific). A single PTFE bead (0.125 inch diameter) was added to each well to facilitate agitation. Plates were incubated in a microplate reader (FLUOstar OMEGA, BMG Labtech) for ~138 hours at 37˚C with 700 rpm double orbital shaking. Fluorescent measurements were recorded every 15 minutes using λ ex = 440 nm and λ em = 480 nm. This was performed with n=3 technical replicates of the mutant and n=6 technical replicates of the WT.

Transmission Electron Microscopy

10μL of aggregated wild-type and mutant TFG samples (taken from in vitro aggregation experiments) was spotted onto carbon film on 150 mesh copper grids (Electron Microscopy

Sciences) and incubated for 4 minutes. Grids were stained with 10μL uranyl acetate solution (2% w/v in water) for 4 minutes. Excess solution was removed by blotting and air dried for 4 minutes. TEM images were acquired with a JOEL 100CX TEM electron microscope at 100 kV.

X-Ray Fiber Diffraction

Aggregated samples of TFG were centrifuged at 15,000 rpm for 30 minutes and buffer was exchanged with water twice. Samples were suspended between two siliconized glass capillaries ~1 mm apart, forming a bridge between the two capillaries. The sample was allowed to dry and the capillary was used to mount the aggregate in an X-ray beam from a Rigaku FR-E rotating anode x-ray generator for 8 minutes. The diffraction pattern was collected on an R-axis IV++ imaging plate detector.

Heat Sensitivity Assay

TFG fibrils were formed at 37˚C as described above. At the endpoint of shaking, the sample was diluted 1:5 in PBS and divided into 7 aliquots of 10μL each. Each aliquot was individually heated to either 40˚C, 50˚C, 60˚C, 70˚C, 80˚C, 90˚C, or 100˚C for 10 minutes using a BioRad T100 Thermocycler. The entire aliquot was then used to prepare a TEM grid as described above.

Guanidine Hydrochloride Sensitivity Assay

TFG fibrils were formed as described above. The fibrils were diluted 1:5 in solutions of GdnHCl dissolved in PBS such that solutions of 0M, 2M, 2.5M, 3M, 3.5M, 4M, and 4.5M GdnHCl were made containing equal concentrations of TFG fibrils for a final volume of 187.5μL for each GdnHCl concentration. These solutions were transferred in triplicate in 50μL aliquots to a black Nunc 384-well optical bottom plate (Thermo Scientific) to continue shaking under identical conditions to the

fibril formation step, except without any PTFE beads, for 2.25 hours. TEM grids were prepared from these samples as described above.

CryoEM Data Collection, Reconstruction, and Model Building

Two and half microliters of mC-TFG-LCD-G269V or mC-TFG-LCD-P285L fibril solution were applied to Quantifoil 1.2/1.3 electron microscope grid which was glow-discharged for 4 min. Grids were blotted with filter paper to remove excess sample and plunge frozen into liquid ethane using a Vitrobot Mark IV (FEI).

For mC-TFG-LCD-G269V, the cryo-EM dataset was collected on a Titan Krios (Thermo Fisher Scientific) microscope equipped with a Bioquantum K3 (Gatan) camera located at the Stanford-SLAC Cryo-EM Center (S^2C^2). The microscope was operated at 300 kV acceleration voltage and a slit width of 20 eV. Movies were acquired with a nominal pixel size of 0.86 Å (x105,000 nominal magnification) with a dose per frame of ~1.3 $e^-/Å^2$. 40 frames were recorded for each movie (total dose per movie was 50 $e^-/Å^2$). Automated data collection was driven by EPU automation software package (Thermo Fisher Scientific). Defocus was set to −1.8 to −2.6 µm. A total of 15,192 movies were collected.

For mC-TFG-LCD-P285L, cryoEM dataset was collected on a Titan Krios (Thermo Fisher Scientific) microscope equipped with a Bioquantum K3 (Gatan) camera. The microscope was operated at 300 kV acceleration voltage and a slit width of 20 eV. Cryo-EM movies were collected at a nominal magnification of x81,000 with the camera operated at its super-resolution mode, yielding a pixel size of 0.539 Å. Automated data collection was done with SerialEM(42). Defocus was targeted to −1.8 to −2.6 µm. A total of 7,920 movies were collected.

For both datasets, motion correction and dose weighting was performed using Unblur(43) and contrast transfer function estimation was performed using CTFFIND 4.1.8(44). For both datasets, a subset of fibril particles was picked manually using EMAN2 e2helixboxer.py(45), which was used to train crYOLO(46). The last automatically picked the rest of the particles. We used RELION(47, 48) to perform particle extraction, 2D classification, helical reconstruction, and 3D refinement. Helical reconstruction was performed with a cylindrical reference(47). The 3D classification was performed using three classes and by manually controlling the tau_fudge factor and healpix_order the particles were separated into good and bad classes. To obtain a higher-resolution reconstruction, the best particles from 1024-pixel box 3D classification were selected, and helical tubes corresponding to good particles were extracted using a box size of 432 pixels for mC-TFG-LCD-G269V fibrils or 384 pixels for mC-TFG-LCD-P285L fibrils. After several more rounds of 3D classification with refinement of helical twist and rise, we then used the final subset of particles to perform high-resolution gold-standard refinement. The final overall resolution was estimated to be 2.8 Å for mC-TFG-LCD-G269V fibrils and 2.6 Å for mC-TFG-LCD-P285L fibrils based on the 0.143 Fourier shell correlation (FSC) resolution cutoff. We then sharpened the maps using phenix.auto_sharpen(49).

We built de novo atomic models into the sharpened maps using Coot(50). We refined the structures using phenix.real_space_refine(51) and validated them using phenix.comprehensive_validation(52, 53).

Solvation Energy Calculation

The solvation energy was calculated based on an algorithm described previously(11, 54). The solvation energy for each residue was calculated by the sum of the products of the area buried for each atom and the corresponding atomic solvation parameters. The overall energy was

calculated by the sum of energies of all residues. The colors assigned to each residue in the solvation energy map represent the sum of energies over all atoms in the residue.

Acknowledgments

This research was supported by the National Institutes of Health (NIH) (grant numbers: R01AG048120 and R01AG07895). Some of this work was performed at the Stanford-SLAC Cryo-EM Center (S2C2), which is supported by the National Institutes of Health Common Fund Transformative High-Resolution Cryo-Electron Microscopy program (grant U24 GM129541). We also acknowledge the use of resources at the Electron Imaging Center for Nanomachines (EICN; supported by UCLA and by instrumentation grants from the NIH (1S10OD018111 and 1U24GM116792) and the NSF (DBI-1338135 and DMR-1548924))

Tables and Figures

Table 3-1: Cryo-EM data collection, refinement and validation statistics

	mC-TFG-LCD-G269V	**mC-TFG-LCD-P285L**
Data collection and processing		
Magnification	x105,000	X81,000
Voltage (kV)	300	300
Electron exposure (e–/Å^2)	50	50
Defocus range (μm)	-1.8 to -2.6	-1.8 to -2.6
Pixel size (Å)	0.86 (counting mode)	0.539 (super mode)
Symmetry imposed	C1; Helical	C1; Helical
Initial particle images (no.)	384,324	690,373
Final particle images (no.)	31,643	193,640

Map resolution (Å)	2.8	2.6
FSC threshold	0.143	0.143
Map resolution range (Å)	200-2.8	200-2.6
Helical Twist (°)	-3.07	179.02
Helical rise (Å)	4.93	2.39
Refinement		
Initial model used (PDB code)	De Novo	De Novo
Model resolution (Å)	2.8	2.6
FSC threshold	0.5	0.5
Model resolution range (Å)	200-2.8	200-2.6
Map sharpening *B* factor ($Å^2$)	103.86	143.57
Model composition		
Non-hydrogen atoms	6705	5780
Protein residues	825	700
B factors ($Å^2$)		
Protein	44.55	14.16
R.m.s. deviations		
Bond lengths (Å)	0.003	0.007
Bond angles (°)	0.561	0.894
Validation		
MolProbity score	1.37	1.95
Clashscore	6.75	10.67
Poor rotamers (%)	0.00	0.00
Ramachandran plot		
Favored (%)	98.04	93.94
Allowed (%)	1.96	6.06
Disallowed (%)	0.00	0.00

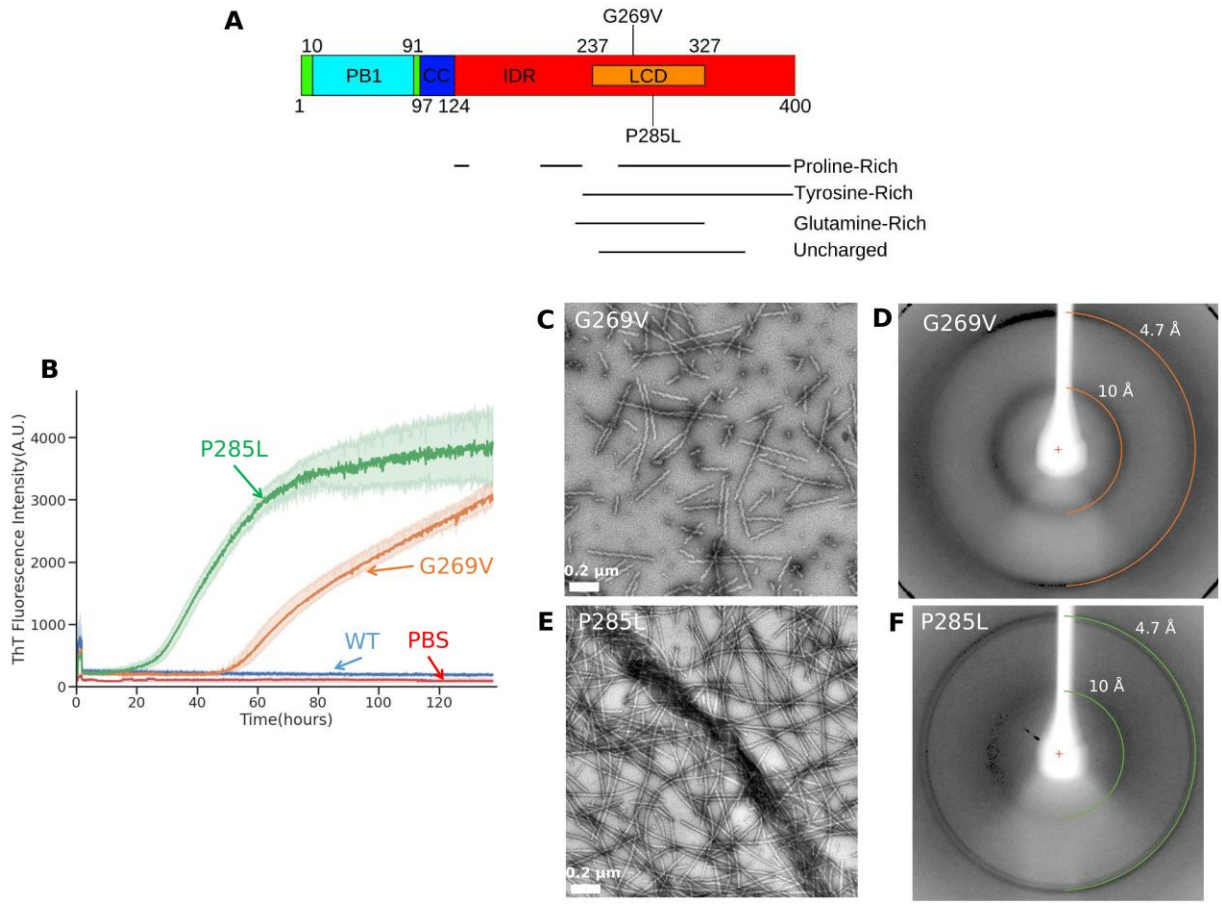


Figure 3-1 The TFG LCD forms amyloid fibrils. A) Domain structure of the full-length TFG protein. The protein contains an ordered PB1 domain and coiled-coil motif (residues 10-91 and 97-124, respectively) and the rest of the protein is intrinsically disordered. An LCD was identified using SEG31 (residues 237-327) which was the portion of the protein which was conjugated to mCerry and expressed and purified. Amyloid-promoting mutations G269V and P285L are also labeled. The ordered fibril cores of both mutant fibrils encompass both of these mutation locations. Regions of the protein especially enriched for certain amino acids are also indicated. B) Time dependent ThT fluorescence for the WT TFG LCD sequence conjugated to a molecule of mCherry (mC-TFG-LCD, labeled “WT”), the TFG LCD with the G269V mutation conjugated to a molecule of mCherry (mC-TFG-LCD-G269V, labeled “G269V”), the TFG LCD with the P285L mutation conjugated to a molecule of mCherry (mC-TFG-LCD-P285L, labeled “P285L”), and PBS buffer (labeled “PBS”). All constructs are at 50µM concentration in PBS with ThT at 40µM

concentration. The mC-TFG-LCD-G269V construct and the PBS blank have n=3 technical replicates and the mC-TFG-LCD and mC-TFG-LCD-P285L construct have n=6 technical replicates. The y-axis values represent the mean ThT fluorescence value of all replicates for each construct. C,E) Electron micrograph of the mC-TFG-LCD-G269V sample (C) and the mC-TFG-LCD-P285L sample (E) at the endpoint of the ThT assay. The mC-TFG-LCD construct did not form fibrils. D,F) X-ray fiber diffraction of mC-TFG-LCD-G269V fibrils (D) and mC-TFG-LCD-P285L fibrils (F). Rings are present at 4.7 Å and 10 Å spacing with distinct wedges, indicative of cross-β structure.

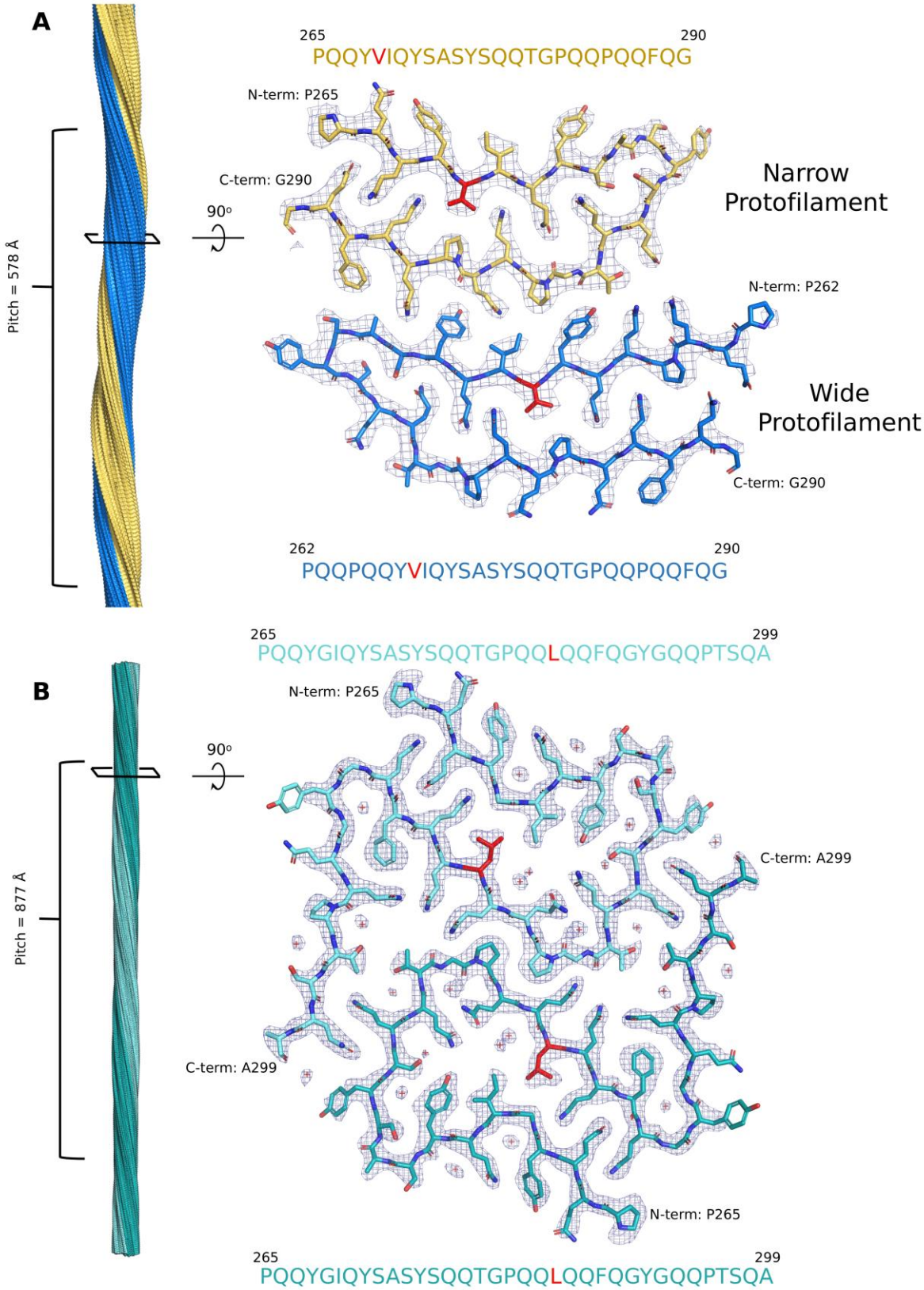

A
265
290
PQQYVIQYSASYSQQTGPQQPQQFQG
N-term: P265
C-term: G290
90°
Narrow
Protofilament
N-term: P262
Wide
Protofilament
C-term: G290
Pitch = 578 Å
262
290
PQQPQQYVIQYSASYSQQTGPQQPQQFQG
B
265
299
PQQYGIQYSASYSQQTGPQQLQQFQGYGQQPTSQA
N-term: P265
90°
C-term: A299
C-term: A299
N-term: P265
Pitch = 877 Å
265
299
PQQYGIQYSASYSQQTGPQQLQQFQGYGQQPTSQA

Figure 3-2 CryoEM structures of the mC-TFG-LCD-G269V and mC-TFG-LCD-P285L amyloid fibril cores. A) Left: mC-TFG-LCD-G269V fibril reconstruction showing a left-handed twist and the pitch. Right: Density map and atomic model of one layer of the fibril viewed down the fibril axis. The G269V mutation is red. The protofilament with 3 extra residues resolved (P262-Q264) is referred to as the “wide” protofilament and the other protofilament, starting at residue 265, is accordingly referred to as the “narrow” protofilament. B) Left: mC-TFG-LCD-P285L fibril reconstruction showing a left-handed twist and the pitch. Right: Density map and atomic model of one layer of the fibril viewed down the fibril axis. The P285L mutation is red. Red crosses represent modeled water molecules.

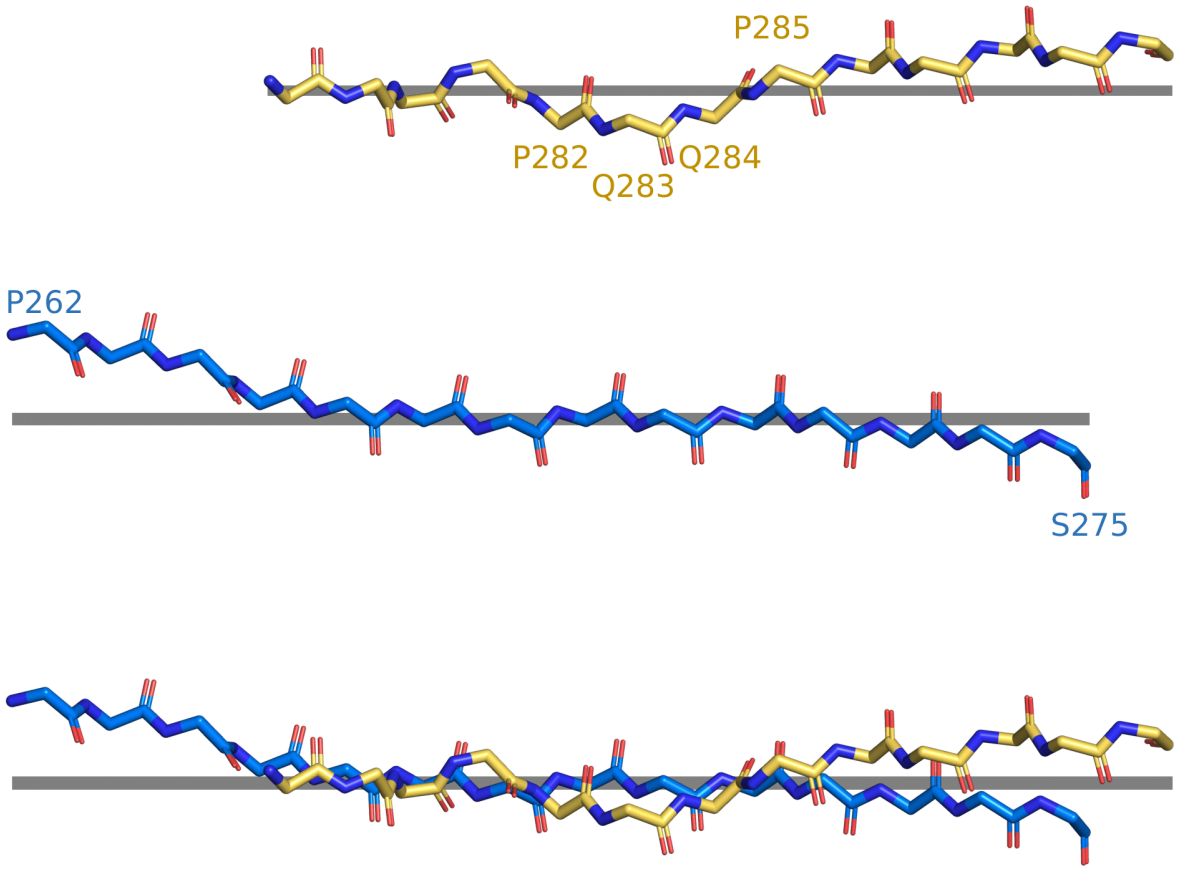


Figure 3-3 Protofilament interface of mC-TFG-LCD-G269V fibril. Top: The peptide backbone of the narrow protofilament of the mC-TFG-LCD-G269V fibril which constitutes the protofilament interface (residues 278-290) viewed orthogonal to the fibril axis with a line representing the plane perpendicular to the fibril axis. Residues P282-P285 are labeled to highlight the backbone warp they constitute. Middle: The peptide backbone of the wide protofilament of the mC-TFG-LCD-G269V fibril which constitutes the protofilament interface (residues 262-275) viewed

orthogonal to the fibril axis with a line representing the plane perpendicular to the fibril axis. Bottom: The peptide backbones of the aforementioned segments of the narrow and wide protofilaments overlayed relative to each other as they exist in the mC-TFG-LCD-G269V fibril structure with a line representing the plane perpendicular to the fibril axis. The peptide backbones of each protofilament are only in plane with each other at a few spots, and notably not at the center of the interface due to the backbone warp of the narrow protofilament at residues 282-285.

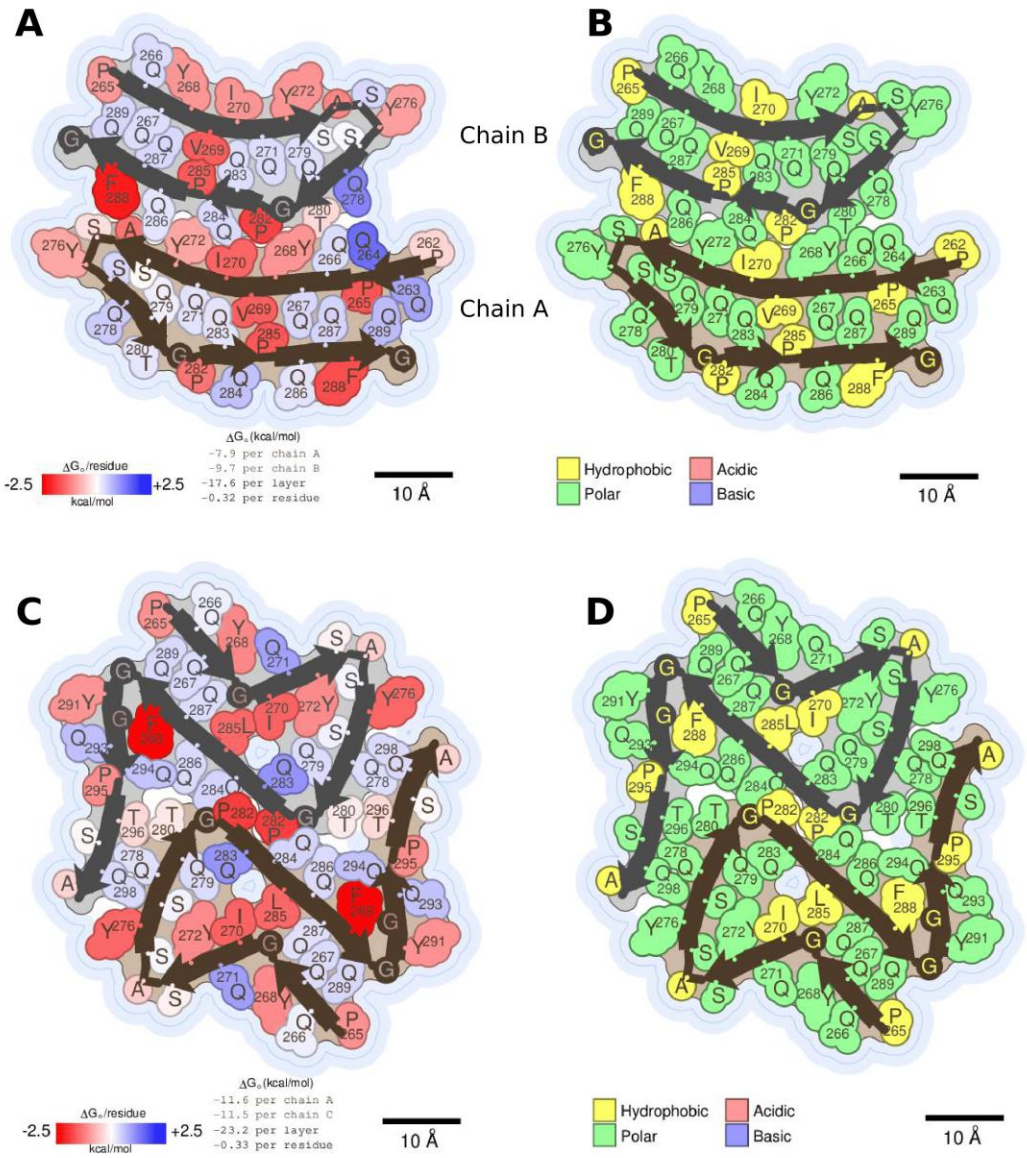


Figure 3-4 Solvation energy maps and polarity maps of the mutant TFG LCD amyloid fibrils. A,C) Solvation energy map of the mC-TFG-LCD-G269V amyloid fibril (A) and mC-TFG-LCD-P285L amyloid fibril (C). Residues are colored according to their stabilization energies. Deeper blue (positive) is unfavorable for amyloid assembly and deeper red (negative) is favorable. The thin dark blue line represents the solvent-accessible surface. Energy values are listed below the structure illustration. For the mC-TFG-LCD-G269V fibril, Chain A is the wide

protofilament and chain B is the narrow protofilament in the illustration. B,D) Polarity map of the mC-TFG-LCD-G269V amyloid fibril (B) and mC-TFG-LCD-P285L amyloid fibril (D). Residues are colored according to whether they are polar, hydrophobic, acidic, or basic. The thin dark blue line represents the solvent-accessible surface.

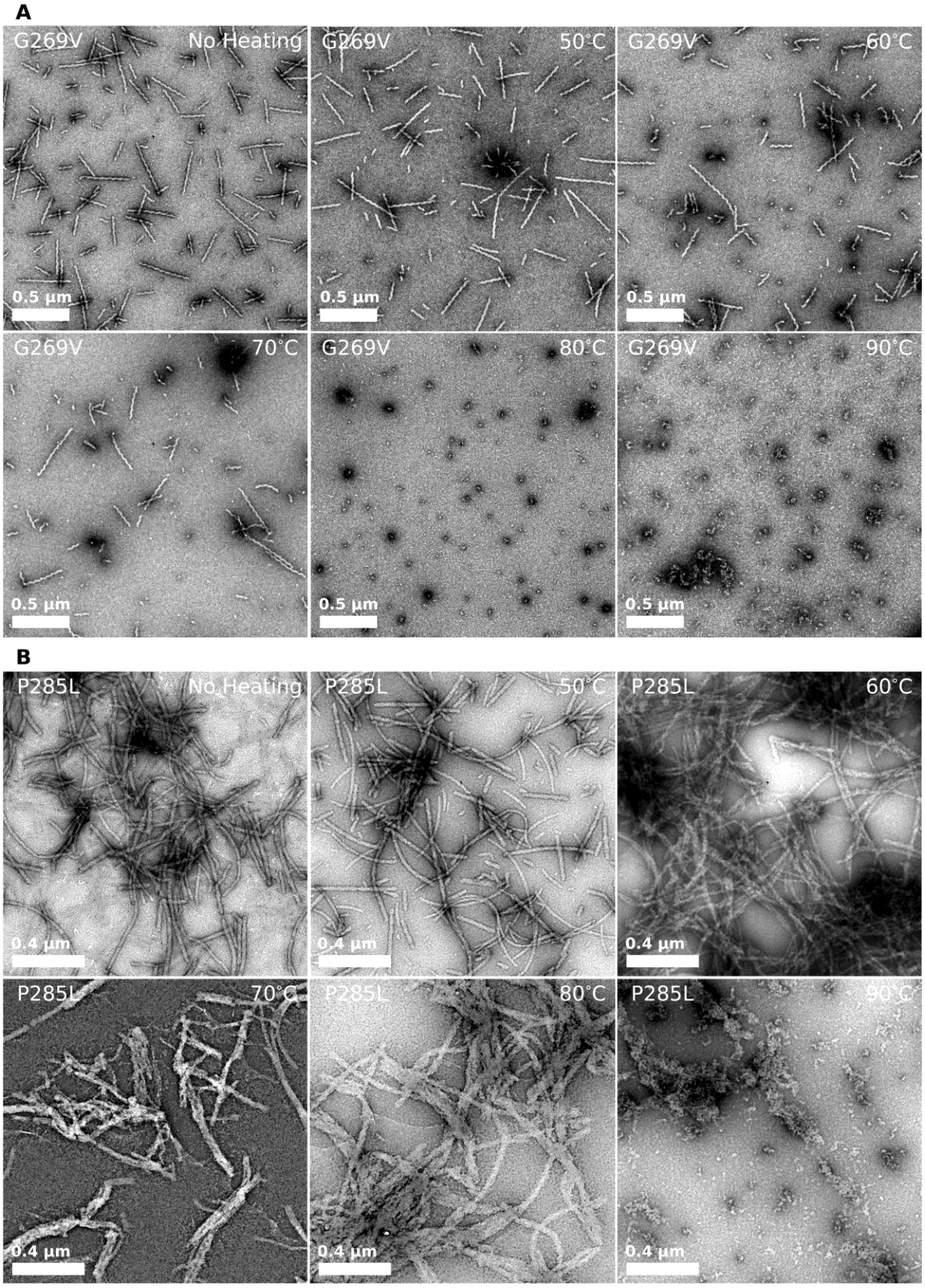

A
G269V
No Heating
0.5 µm
G269V
50˚C
0.5 µm
G269V
60˚C
0.5 µm
G269V
70˚C
0.5 µm
G269V
80˚C
0.5 µm
G269V
90˚C
0.5 µm
B
P285L
No Heating
0.4 µm
P285L
50˚C
0.4 µm
P285L
60˚C
0.4 µm
P285L
70˚C
0.4 µm
P285L
80˚C
0.4 µm
P285L
90˚C
0.4 µm

Figure 3-5 Heat stability of mutant TFG LCD amyloid fibers. Amyloid fibers formed at 50μm monomer concentration at 37°C were diluted 1:5 in PBS and distributed into 10μL aliquots. Each aliquot was heated to the specified temperature (40°C to 100°C in intervals of 10°C) for 10 minutes each and then each sampled was prepared for negative stain EM. A) Representative images of mC-TFG-LCD-G269V fibrils heated to 50°C to 90°C. Fibrils remain clearly visible in all samples heated below 80°C, but are completely absent in samples heated to 80°C and above. B) Representative images of mC-TFG-LCD-P285L fibrils heated to 50°C to 90°C. Fibrils remain clearly visible in all samples heated below 90°C, but are completely absent in samples heated to 90°C and above.

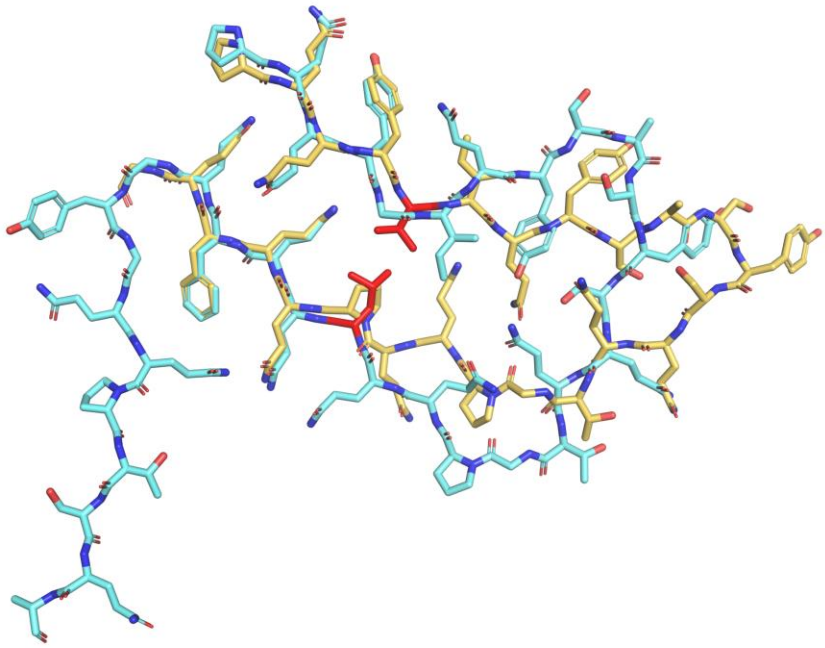


Figure S3-1 Superposition of the narrow protofilament of the mC-TFG-LCD-G269V fibril (yellow) and one of the protofilaments of the mC-TFG-LCD-P285L fibril (blue) demonstrating the similarity of the glutamine zipper formed by residues Q267, Q287, and Q289

in both structures, but the divergence of the structural similarity past the location of the mutations (mutant residues from each structure are colored red).

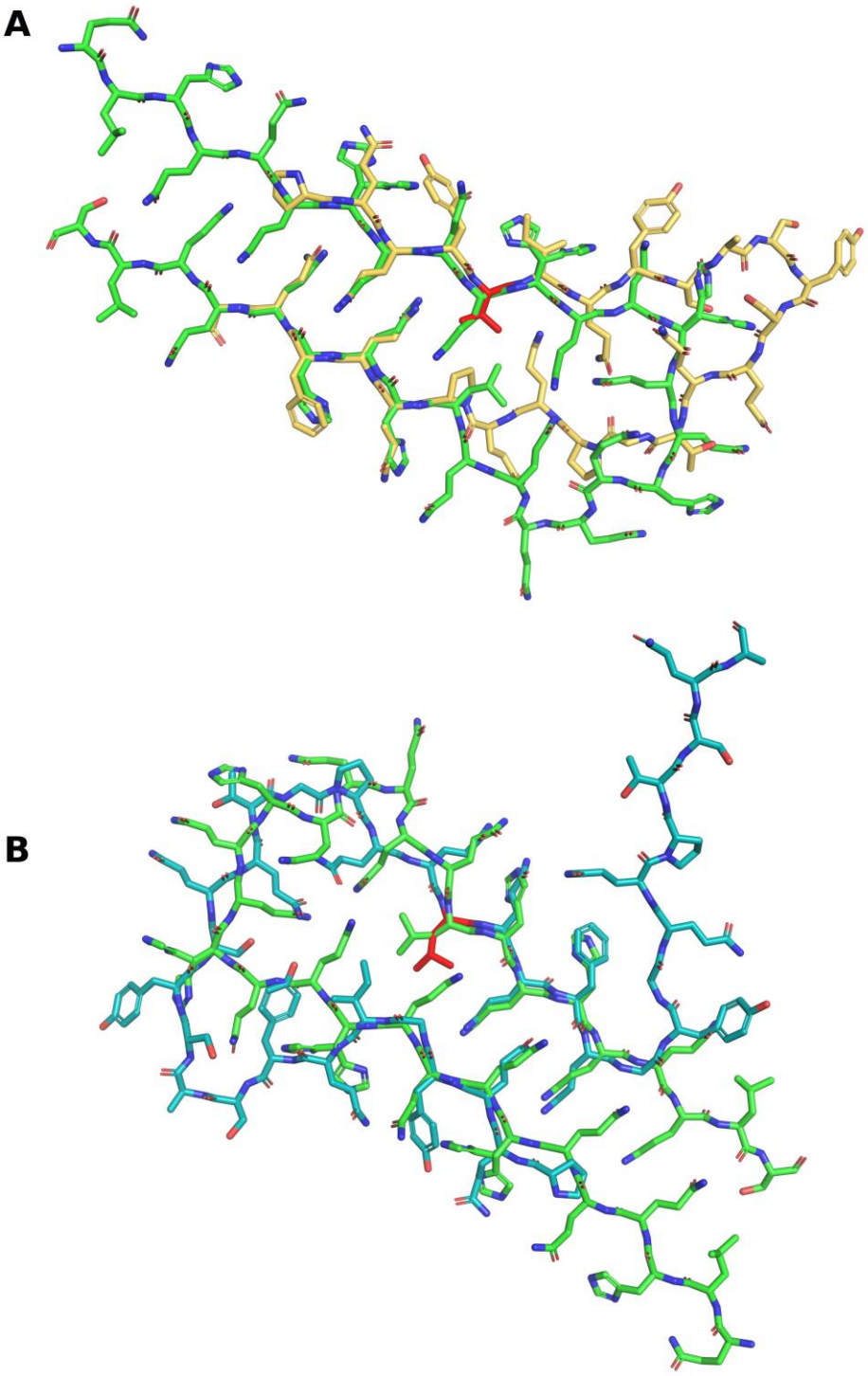

A
B

Figure S3-2 A) Superposition of the narrow filament of the mC-TFG-LCD-G269V fibril (yellow) and one of the protofilaments of the orb2A amyloid fibril (green). B) Superposition of one protofilament of the mC-TFG-LCD-P285L fibril (blue) and one of the protofilaments of the orb2A amyloid fibril (green). Both superpositions are aligned so the glutamine zipper present in both mutant TFG LCD fibril structures (formed by residues Q267, Q287, and Q289) overlaps one of the glutamine zippers of the orb2A fibril structure (in this case residues Q30, Q47, and Q49).

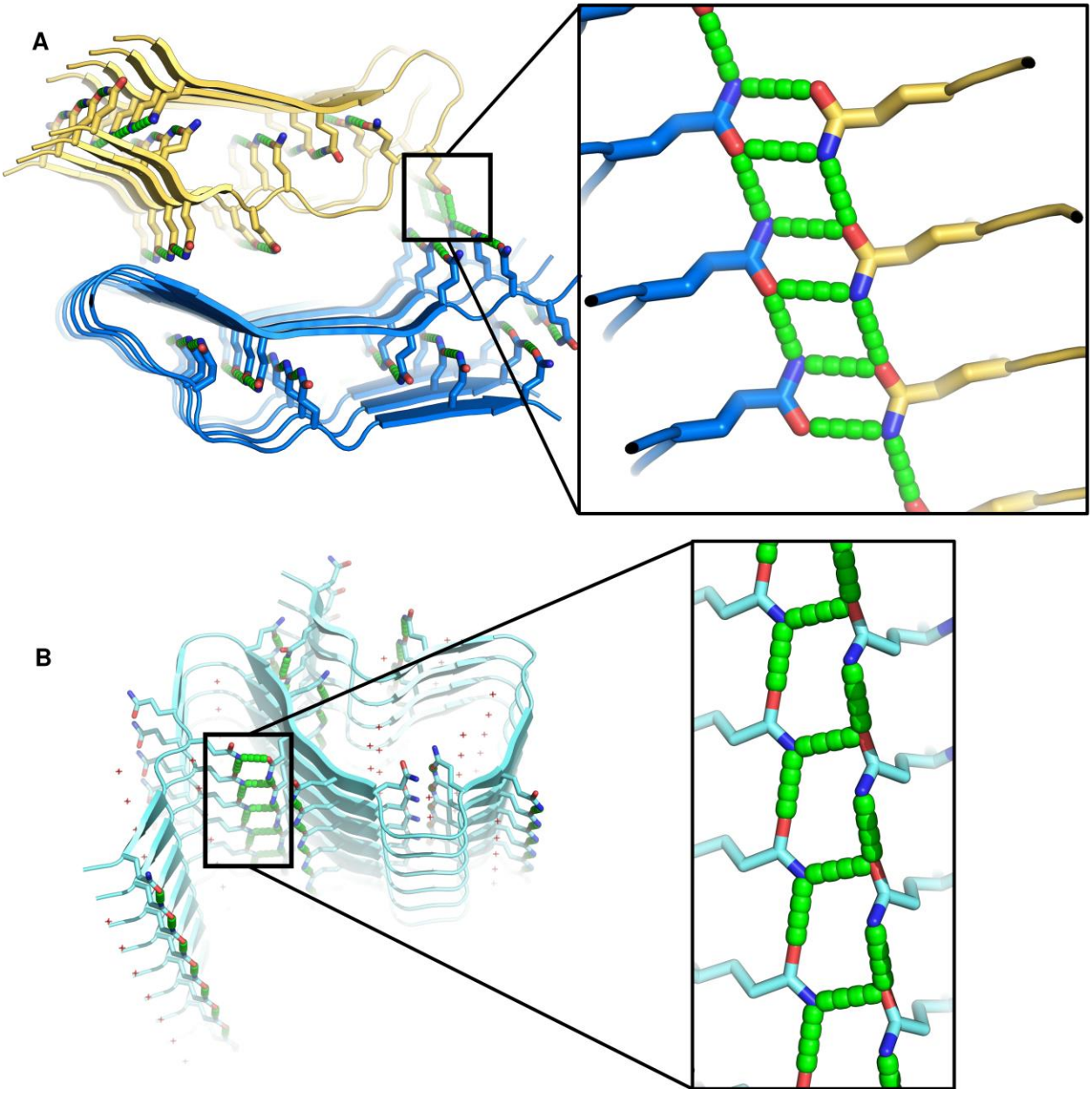


Figure S3-3 A) Cartoon representation of the mC-TFG-LCD-G269V fibril showing only the glutamine side chains which form inter-layer hydrogen bonds (hydrogen bonds are represented by green connections). Inset: glutamine side chains which form inter-protofilament hydrogen bonds as well as inter-layer hydrogen bonds. B) Cartoon representation of one protofilament of the mC-TFG-LCD-P285L fibril showing all glutamine side chains (hydrogen bonds are represented by green connections). Two glutamine side chains are not making interlayer hydrogen bonds: one is facing solvent and one is facing the interior of the protofilament but is solvated by three water molecules. Inset: glutamine side chains which form intra-protofilament hydrogen bonds as well as inter-layer hydrogen bonds.

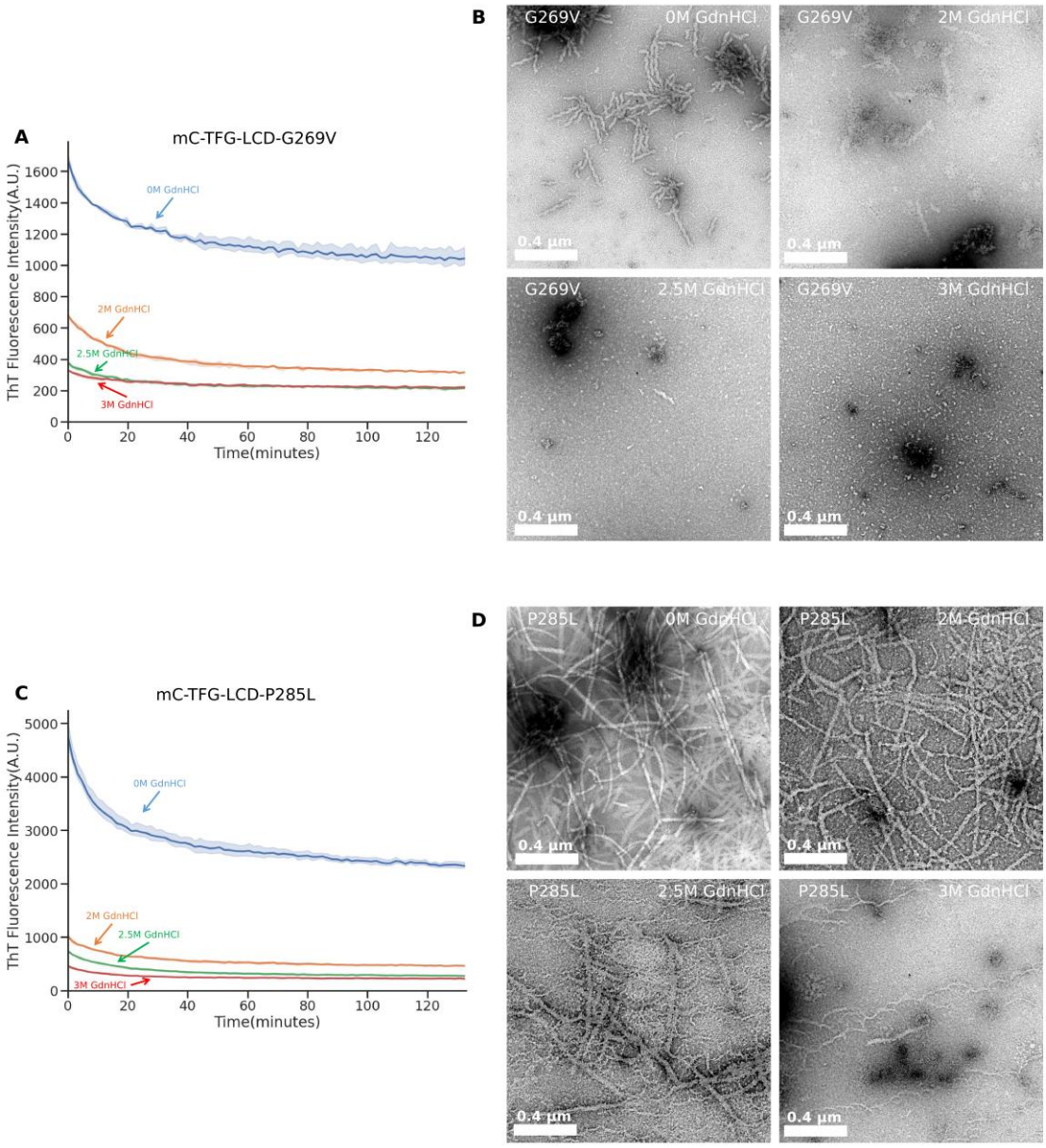


Figure S3-4 A,C) Time dependent ThT fluorescence of mC-TFG-LCD-G269V fibrils (A) and mC-TFG-LCD-P285L fibrils (C) in solutions of varying concentrations of guanidine hydrochloride (GdnHCl) and 20µM ThT: 0M, 2M, 2.5M, or 3M GdnHCl. Fibrils were generated at 50µM concentration in PBS with ThT at 20µM concentration and diluted 1:5 in the GdnHCl solutions. B,D) Representative TEM images of samples taken from the endpoint of the ThT readings. For the mC-TFG-LCD-G269V fibrils (B), fibrils remained abundant in 2M GdnHCl, but were

extremely sparse in 2.5M GdnHCl, and were completely absent in 3M GdnHCl. For the mC-TFG-LCD-P285L fibrils (D), fibrils remained abundant in both 2M and 2.5M GdnHCl, but were extremely sparse in 3M GdnHCl.

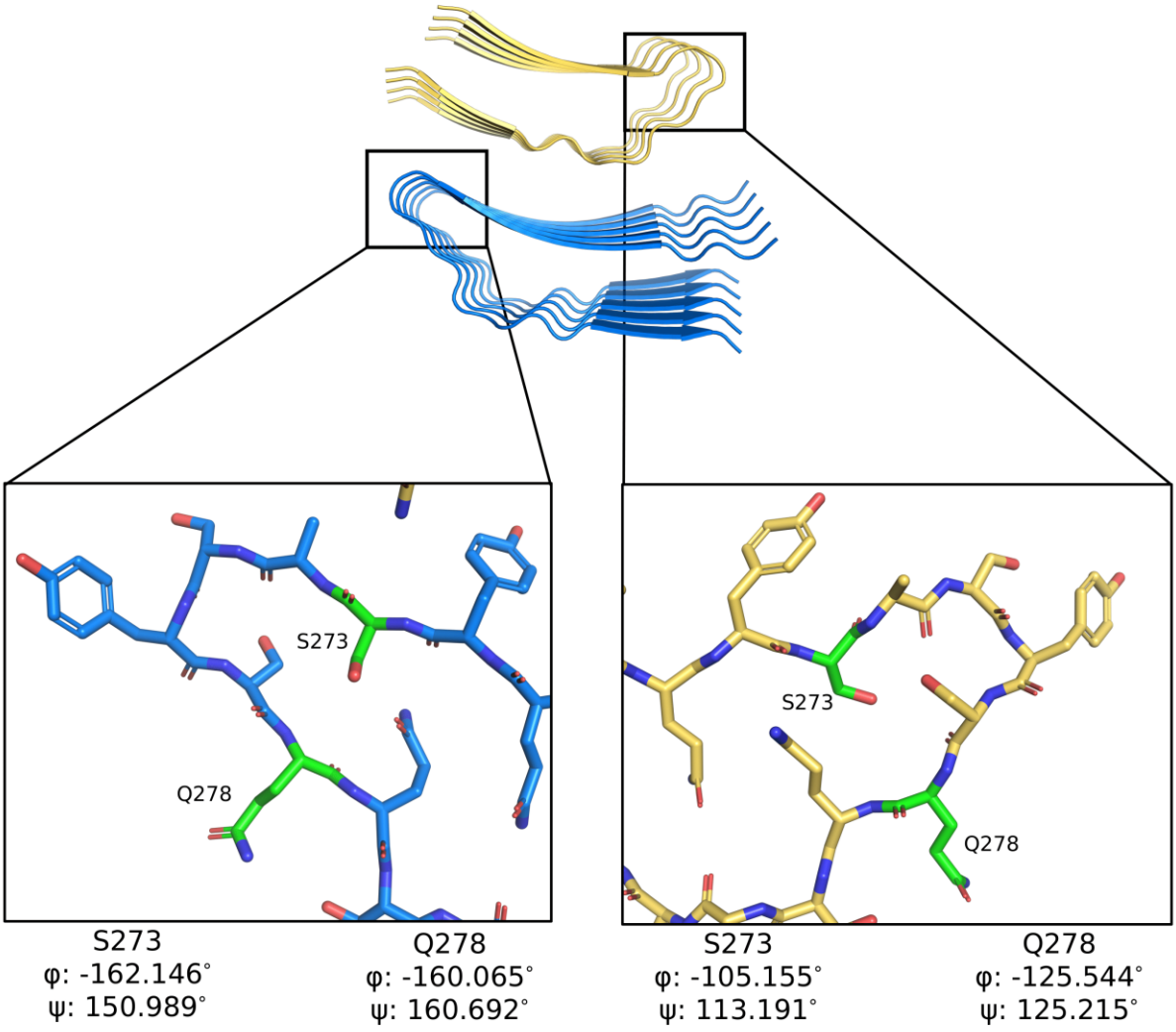


Figure S3-5 Assymetric extended β-sheet conformation between mC-TFG-LCD-G269V protofilaments. The wide protofilament (blue; inset: left) has two residues (S273 and Q278), colored green, with |φ| and |ψ| angles with values which permit an extended β-sheet

conformation (> 150°). The narrow protofilament (yellow; inset: right) has no residues with appropriate φ and ψ angles to permit this structure. Residues in the narrow protofilament corresponding to the residues in the wide protofilament with an extended β-sheet conformation are also colored green. An extended β-sheet conformation makes C5-hydrogen bonds possible. φ and ψ angle values for residues S273 and Q278 for each protofilament are listed below each inset of the corresponding protofilament.

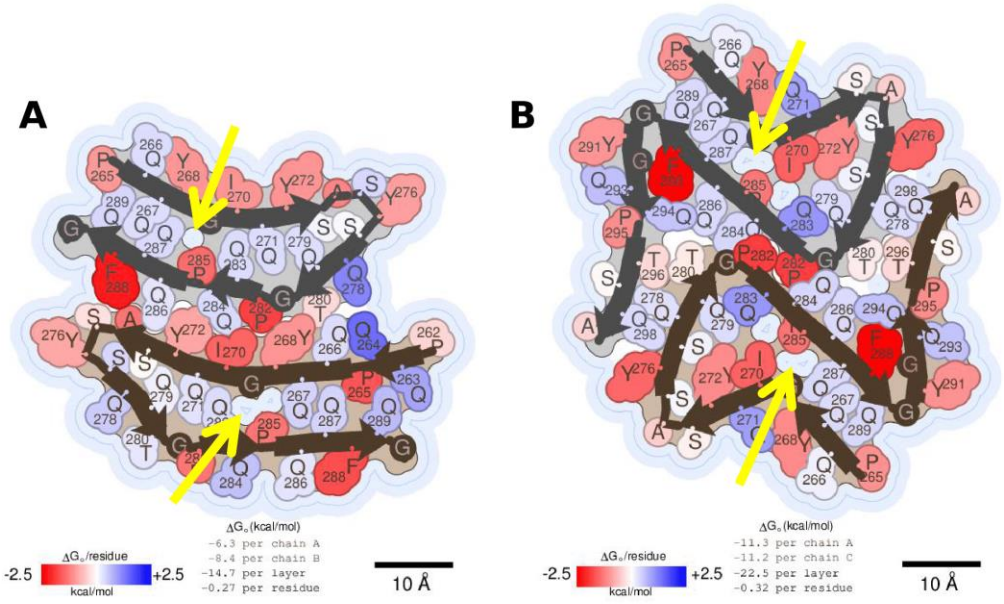


Figure S3-6 Hypothetical structures of mC-TFG-LCD WT amyloid fibrils. The mutant residues in the TFG LCD fibril structure models were replaced with the WT residues (in the mC-TFG-LCD-G269V model, the mutant valine residue is now the WT glycine; in the mC-TFG-LCD-P285L model, the mutant leucine residue is not the WT proline). A,B) Solvation energy maps of the mC-TFG-LCD-G269V fibril core with the WT sequence (A) and the mC-TFG-LCD-P285L fibril core with the WT sequence (B). Residues are colored according to their stabilization

energies. Deeper blue (positive) is unfavorable for amyloid assembly and deeper red (negative) is favorable. The thin dark blue line represents the solvent-accessible surface. Energy values are listed below the structure illustration. In (A) chain A is the wide protofilament and chain B is the narrow protofilament in the illustration. Yellow arrows point to solvent accessible channels which do not exist in the mutant structures.

References

1. Rosenberg, G. M., Murray, K. A., Salwinski, L., Hughes, M. P., Abskharon, R., and Eisenberg, D. S. (2022) Bioinformatic identification of previously unrecognized amyloidogenic proteins. *J Biol Chem*. **298**, 101920
2. Johnson, A., Bhattacharya, N., Hanna, M., Pennington, J. G., Schuh, A. L., Wang, L., Otegui, M. S., Stagg, S. M., and Audhya, A. (2015) TFG clusters COPII-coated transport carriers and promotes early secretory pathway organization. *EMBO J*. **34**, 811–827
3. Hanna, M. G., Block, S., Frankel, E. B., Hou, F., Johnson, A., Yuan, L., Knight, G., Moresco, J. J., Yates, J. R., Ashton, R., Schekman, R., Tong, Y., and Audhya, A. (2017) TFG facilitates outer coat disassembly on COPII transport carriers to promote tethering and fusion with ER-Golgi intermediate compartments. *Proc Natl Acad Sci U S A*. **114**, E7707–E7716
4. Tsai, P.-C., Huang, Y.-H., Guo, Y.-C., Wu, H.-T., Lin, K.-P., Tsai, Y.-S., Liao, Y.-C., Liu, Y.-T., Liu, T.-T., Kao, L.-S., Yet, S.-F., Fann, M.-J., Soong, B.-W., and Lee, Y.-C. (2014) A novel TFG mutation causes Charcot-Marie-Tooth disease type 2 and impairs TFG function. *Neurology*. **83**, 903–912
5. Ishiura, H., Sako, W., Yoshida, M., Kawarai, T., Tanabe, O., Goto, J., Takahashi, Y., Date, H., Mitsui, J., Ahsan, B., Ichikawa, Y., Iwata, A., Yoshino, H., Izumi, Y., Fujita, K., Maeda, K., Goto, S., Koizumi, H., Morigaki, R., Ikemura, M., Yamauchi, N., Murayama, S.,

Nicholson, G. A., Ito, H., Sobue, G., Nakagawa, M., Kaji, R., and Tsuji, S. (2012) The TRK-fused gene is mutated in hereditary motor and sensory neuropathy with proximal dominant involvement. *Am J Hum Genet*. **91**, 320–329

6. Reilly, M. M. (2005) Axonal Charcot-Marie-Tooth disease: the fog is slowly lifting! *Neurology*. **65**, 186–187
7. Takashima, H., Nakagawa, M., Nakahara, K., Suehara, M., Matsuzaki, T., Higuchi, I., Higa, H., Arimura, K., Iwamasa, T., Izumo, S., and Osame, M. (1997) A new type of hereditary motor and sensory neuropathy linked to chromosome 3. *Ann Neurol*. **41**, 771–780
8. Eisenberg, D. S., and Sawaya, M. R. (2017) Structural Studies of Amyloid Proteins at the Molecular Level. *Annu Rev Biochem*. **86**, 69–95
9. Sawaya, M. R., Sambashivan, S., Nelson, R., Ivanova, M. I., Sievers, S. A., Apostol, M. I., Thompson, M. J., Balbirnie, M., Wiltzius, J. J. W., McFarlane, H. T., Madsen, A. Ø., Riekel, C., and Eisenberg, D. (2007) Atomic structures of amyloid cross-beta spines reveal varied steric zippers. *Nature*. **447**, 453–457
10. Lu, J., Cao, Q., Hughes, M. P., Sawaya, M. R., Boyer, D. R., Cascio, D., and Eisenberg, D. S. (2020) CryoEM structure of the low-complexity domain of hnRNPA2 and its conversion to pathogenic amyloid. *Nat Commun*. **11**, 4090
11. Sawaya, M. R., Hughes, M. P., Rodriguez, J. A., Riek, R., and Eisenberg, D. S. (2021) The expanding amyloid family: Structure, stability, function, and pathogenesis. *Cell*. **184**, 4857–4873
12. Murray, D. T., Kato, M., Lin, Y., Thurber, K. R., Hung, I., McKnight, S. L., and Tycko, R. (2017) Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. *Cell*. **171**, 615-627.e16

13. Wasmer, C., Lange, A., Van Melckebeke, H., Siemer, A. B., Riek, R., and Meier, B. H. (2008) Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core. *Science*. **319**, 1523–1526

14. Zhang, W., Falcon, B., Murzin, A. G., Fan, J., Crowther, R. A., Goedert, M., and Scheres, S. H. (2019) Heparin-induced tau filaments are polymorphic and differ from those in Alzheimer's and Pick's diseases. *Elife*. **8**, e43584

15. Hervas, R., Rau, M. J., Park, Y., Zhang, W., Murzin, A. G., Fitzpatrick, J. A. J., Scheres, S. H. W., and Si, K. (2020) Cryo-EM structure of a neuronal functional amyloid implicated in memory persistence in Drosophila. *Science*. **367**, 1230–1234

16. Cao, Q., Boyer, D. R., Sawaya, M. R., Ge, P., and Eisenberg, D. S. (2019) Cryo-EM structures of four polymorphic TDP-43 amyloid cores. *Nat Struct Mol Biol*. **26**, 619–627

17. Scherzinger, E., Lurz, R., Turmaine, M., Mangiarini, L., Hollenbach, B., Hasenbank, R., Bates, G. P., Davies, S. W., Lehrach, H., and Wanker, E. E. (1997) Huntingtin-encoded polyglutamine expansions form amyloid-like protein aggregates in vitro and in vivo. *Cell*. **90**, 549–558

18. Zhou, X., Sumrow, L., Tashiro, K., Sutherland, L., Liu, D., Qin, T., Kato, M., Liszczak, G., and McKnight, S. L. (2022) Mutations linked to neurological disease enhance self-association of low-complexity protein sequences. *Science*. **377**, eabn5582

19. Newberry, R. W., and Raines, R. T. (2016) A prevalent intraresidue hydrogen bond stabilizes proteins. *Nat Chem Biol*. **12**, 1084–1088

20. Murray, K. A., Evans, D., Hughes, M. P., Sawaya, M. R., Hu, C. J., Houk, K. N., and Eisenberg, D. (2022) Extended β-Strands Contribute to Reversible Amyloid Formation. *ACS Nano*. **16**, 2154–2163

21. Levitt, M. (1978) Conformational preferences of amino acids in globular proteins. *Biochemistry*. **17**, 4277–4285

22. Scharner, J., Brown, C. A., Bower, M., Iannaccone, S. T., Khatri, I. A., Escolar, D., Gordon, E., Felice, K., Crowe, C. A., Grosmann, C., Meriggioli, M. N., Asamoah, A., Gordon, O., Gnocchi, V. F., Ellis, J. A., Mendell, J. R., and Zammit, P. S. (2011) Novel LMNA mutations in patients with Emery-Dreifuss muscular dystrophy and functional characterization of four LMNA mutations. *Hum Mutat*. **32**, 152–167

23. Keith, J. L., Swinkin, E., Gao, A., Alminawi, S., Zhang, M., McGoldrick, P., McKeever, P., Robertson, J., Rogaeva, E., and Zinman, L. (2020) Neuropathologic description of CHCHD10 mutated amyotrophic lateral sclerosis. *Neurol Genet*. **6**, e394

24. Murray, K. A., Hughes, M. P., Hu, C. J., Sawaya, M. R., Salwinski, L., Pan, H., French, S. W., Seidler, P. M., and Eisenberg, D. S. (2022) Identifying amyloid-related diseases by mapping mutations in low-complexity protein domains to pathologies. *Nat Struct Mol Biol*. **29**, 529–536

25. Kobayashi, H., and Hashimoto, K. (1983) Amyloidogenesis in organ-limited cutaneous amyloidosis: an antigenic identity between epidermal keratin and skin amyloid. *J Invest Dermatol*. **80**, 66–72

26. Huilgol, S. C., Ramnarain, N., Carrington, P., Leigh, I. M., and Black, M. M. (1998) Cytokeratins in primary cutaneous amyloidosis. *Australas J Dermatol*. **39**, 81–85

27. Chang, Y. T., Liu, H. N., Wang, W. J., Lee, D. D., and Tsai, S. F. (2004) A study of cytokeratin profiles in localized cutaneous amyloids. *Arch Dermatol Res*. **296**, 83–88

28. Wasif, N., Naqvi, S. K. ul-Hassan, Basit, S., Ali, N., Ansar, M., and Ahmad, W. (2011) Novel mutations in the keratin-74 (KRT74) gene underlie autosomal dominant woolly hair/hypotrichosis in Pakistani families. *Hum Genet*. **129**, 419–424

29. Caubet, C., Bousset, L., Clemmensen, O., Sourigues, Y., Bygum, A., Chavanas, S., Coudane, F., Hsu, C.-Y., Betz, R. C., Melki, R., Simon, M., and Serre, G. (2010) A new

amyloidosis caused by fibrillar aggregates of mutated corneodesmosin. *FASEB J*. **24**, 3416–3426

30. Hughes, M. P., Sawaya, M. R., Boyer, D. R., Goldschmidt, L., Rodriguez, J. A., Cascio, D., Chong, L., Gonen, T., and Eisenberg, D. S. (2018) Atomic structures of low-complexity protein segments reveal kinked β sheets that assemble networks. *Science*. **359**, 698–701
31. Mompeán, M., Li, W., Li, J., Laage, S., Siemer, A. B., Bozkurt, G., Wu, H., and McDermott, A. E. (2018) The Structure of the Necrosome RIPK1-RIPK3 Core, a Human Hetero-Amyloid Signaling Complex. *Cell*. **173**, 1244-1253.e10
32. Fowler, D. M., Koulov, A. V., Alory-Jost, C., Marks, M. S., Balch, W. E., and Kelly, J. W. (2006) Functional amyloid formation within mammalian tissue. *PLoS Biol*. **4**, e6
33. Cereghetti, G., Saad, S., Dechant, R., and Peter, M. (2018) Reversible, functional amyloids: towards an understanding of their regulation in yeast and humans. *Cell Cycle*. **17**, 1545–1558
34. Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A. P., Kim, H. J., Mittag, T., and Taylor, J. P. (2015) Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell*. **163**, 123–133
35. Kato, M., Han, T. W., Xie, S., Shi, K., Du, X., Wu, L. C., Mirzaei, H., Goldsmith, E. J., Longgood, J., Pei, J., Grishin, N. V., Frantz, D. E., Schneider, J. W., Chen, S., Li, L., Sawaya, M. R., Eisenberg, D., Tycko, R., and McKnight, S. L. (2012) Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell*. **149**, 753–767
36. Starck, C. S., and Sutherland-Smith, A. J. (2010) Cytotoxic aggregation and amyloid formation by the myostatin precursor protein. *PLoS One*. **5**, e9170

37. Pepys, M. B., Hawkins, P. N., Booth, D. R., Vigushin, D. M., Tennent, G. A., Soutar, A. K., Totty, N., Nguyen, O., Blake, C. C., and Terry, C. J. (1993) Human lysozyme gene mutations cause hereditary systemic amyloidosis. *Nature*. **362**, 553–557

38. Siddiqi, M. K., Malik, S., Majid, N., Alam, P., and Khan, R. H. (2019) Cytotoxic species in amyloid-associated diseases: Oligomers or mature fibrils. *Adv Protein Chem Struct Biol*. **118**, 333–369

39. Irvine, G. B., El-Agnaf, O. M., Shankar, G. M., and Walsh, D. M. (2008) Protein aggregation in the brain: the molecular basis for Alzheimer's and Parkinson's diseases. *Mol Med*. **14**, 451–464

40. Yamamotoya, T., Hasei, S., Akasaka, Y., Ohata, Y., Nakatsu, Y., Kanna, M., Fujishiro, M., Sakoda, H., Ono, H., Kushiyama, A., Misawa, H., and Asano, T. (2022) Involvement of neuronal and muscular Trk-fused gene (TFG) defects in the development of neurodegenerative diseases. *Sci Rep*. **12**, 1966

41. Wootton, J. C., and Federhen, S. (1993) Statistics of local complexity in amino acid sequences and sequence databases. *Computers & Chemistry*. **17**, 149–163

42. Mastronarde, D. N. (2005) Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol*. **152**, 36–51

43. Grant, T., and Grigorieff, N. (2015) Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. *Elife*. **4**, e06980

44. Rohou, A., and Grigorieff, N. (2015) CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol*. **192**, 216–221

45. Tang, G., Peng, L., Baldwin, P. R., Mann, D. S., Jiang, W., Rees, I., and Ludtke, S. J. (2007) EMAN2: an extensible image processing suite for electron microscopy. *J Struct Biol*. **157**, 38–46

46. Wagner, T., Merino, F., Stabrin, M., Moriya, T., Antoni, C., Apelbaum, A., Hagel, P., Sitsel, O., Raisch, T., Prumbaum, D., Quentin, D., Roderer, D., Tacke, S., Siebolds, B., Schubert, E., Shaikh, T. R., Lill, P., Gatsogiannis, C., and Raunser, S. (2019) SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Commun Biol*. **2**, 218

47. He, S., and Scheres, S. H. W. (2017) Helical reconstruction in RELION. *J Struct Biol*. **198**, 163–176

48. Scheres, S. H. W. (2012) RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol*. **180**, 519–530

49. Terwilliger, T. C., Sobolev, O. V., Afonine, P. V., and Adams, P. D. (2018) Automated map sharpening by maximization of detail and connectivity. *Acta Crystallogr D Struct Biol*. **74**, 545–559

50. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*. **66**, 486–501

51. Afonine, P. V., Poon, B. K., Read, R. J., Sobolev, O. V., Terwilliger, T. C., Urzhumtsev, A., and Adams, P. D. (2018) Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol*. **74**, 531–544

52. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*. **66**, 213–221

53. Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*. **66**, 12–21

54. Eisenberg, D., and McLachlan, A. D. (1986) Solvation energy in protein folding and binding. *Nature*. **319**, 199–203

Chapter 4: Genetic causes of amyloid fibrils and their structures

Gregory M. Rosenberg, Kevin A. Murray, Michael R. Sawaya, Yi Xiao Jiang, David S. Eisenberg

Abstract

The conversion of proteins into the insoluble fibrillar aggregates known as amyloids is a component of the etiology of a wide variety of diseases. This particular kind of protein misfolding can occur sporadically but can also be induced or accelerated by genetic mutations. With the growing number of high-resolution structures of amyloid fibers, our knowledge of the structural details that define amyloid fibers is greater than ever. In this review, we catalog 58 proteins which have been demonstrated to form amyloid fibers *in vivo* or *in vitro* and which have pathogenic mutations which have been shown to induce such fiber formation. We also examine the mechanisms by which genetic changes influence amyloid formation and the resulting amyloid fiber structures as well as the relationship between those structures and the observed disease outcomes. We also examine how mutation-driven amyloid fiber formation contributes to disease in the context of inheritance patterns of the associated diseases and whether or not the genes are tolerant to loss-of-function, i.e. haplosufficient. While our base of knowledge surrounding amyloid fibers and their structures has grown considerably over recent years, there is still much work to be done in understanding the role mutations play in the diverse landscape of amyloid proteins.

Introduction

An amyloid disease was probably first described by Nicolao Fontano in 1639 in *Responsionum & Curationum Medicinalium*, in which he reported on an autopsy and noted that the patient had a large spleen filled with white stones(1). In hindsight this may have been a case of sago spleen, which indicates a systemic amyloidosis, either primary or secondary (light chain amyloidosis or amyloid A amyloidosis)(2, 3). Our understanding of the nature and diversity of amyloids is still

evolving even today. Amyloids are fibrillar aggregates consisting of stacked layers of protein monomers. Amyloid can be identified *in vitro* and *in vivo* through its affinity to certain dyes such as Congo red, thioflavin S, and thioflavin T(4, 5). Also, amyloid fibers, no matter which protein forms the aggregate, have a characteristic structural motif of β-strands positioned across from each other in an orientation perpendicular to the fiber axis, called a “cross-β motif”, which can be visualized through x-ray diffraction(6, 7). Dozens of amyloid proteins have been identified in a wide range of pathologies with and without familial mutations. Structural details of the amyloid fibers, including their cross-β motif, have been deduced over time through a wide variety of techniques like solid state NMR, x-ray fiber and powder diffraction, cryo-electron microscopy (cryo-EM), peptide single crystal x-ray diffraction, and more(8, 9). Relatively recently, cryo-EM was used to determine the high resolution structure of amyloid fibers of microtubule-associated protein tau extracted from the brain of an Alzheimer’s disease patient(10) and since then over 250 structures of amyloid fibers (recombinant, synthetic, and patient-derived; functional, pathogenic, and neither) have been determined mainly by cryo-EM(9). The Nomenclature Committee of the International Society of Amyloidosis meets every other year to update definitions, nomenclature, and their lists of amyloid proteins which fall under these updated criteria(11). The field of amyloid research is truly massive and has experienced many breakthroughs.

Even so, there is much more to learn and discover about amyloid proteins. The topic of this review is one such area of research which is still in relative infancy: the influence of genetic mutations on amyloid formation and structure. This is not to say that there is nothing known about the relationship between genetics and amyloidosis; familial amyloid polyneuropathy (FAP; caused by transthyretin mutations) was first described in 1952(12). Rather, the goal of this review is to demonstrate where there are gaps in our knowledge of the crossroads of structural biology and

molecular genetics. In this review, we will catalog and categorize every known pathogenic amyloid protein and summarize what we know about the effects mutations (if they have any) have on their behavior and molecular structure before and after conversion to an amyloid fiber. We will also discuss how inheritance patterns of pathogenic amyloidogenic mutations, haplosufficiency/haploinsufficiency, and molecular structures of mutant amyloid fibers contribute to our understanding of amyloid disease etiology. Through discussing these topics, it will become clear that, while a lot is known, the study of mutations in amyloid proteins is a field of research which is very open to new discoveries in terms of basic science and clinical application.

Classes of pathogenic amyloids

Proteins which this review will consider amyloid proteins are any proteins which have evidence *in vivo* and/or *in vitro* of amyloid fibril formation and which are associated with human disease. *In vivo* evidence is usually the protein making up the congophilic material (called “amyloid”) within human tissue, confirmed through proteomic analysis of said tissue. This tissue can also sometimes be visualized using electron microscopy to see the characteristic unbranched amyloid fibers. *In vitro* evidence is more broad, including any protein which has been shown to form amyloid fibers in the lab by ThT fluorescence, x-ray fiber diffraction, electron microscopy, etc. It is widely accepted that nearly any protein is capable of amyloid formation under proper conditions(13), so it must be made clear whether the origin of an amyloid aggregate is a living organism or a test tube(11). The proteins included in our list are all implicated in disease, so whether their amyloid-forming capabilities have been demonstrated *in vivo* or not, the capability to form amyloid fibers is relevant to note. So, for the purposes of this review, we will not distinguish between amyloid proteins with evidence of *in vivo* or *in vitro* fiber formation except in the supplemental amyloid protein profiles. Also, aligning with this philosophy, we have excluded PIK3R1 from our list. This protein’s SH3 domain has been used as a model amyloid *in vitro* and

has been structurally characterized in its native and amyloid state, but is not associated with any human disease(14, 15). We have also excluded many “functional amyloids" which have been shown to form amyloids or amyloid-like structures *in vivo* as part of their native function, but do not cause disease through amyloid formation(16–18), although some functional amyloids are capable of forming pathogenic amyloids and are thus included. One further consideration is the difference between a mutation and a polymorphism. A polymorphism is a genetic variant that may be associated with disease but is also observed frequently in healthy individuals. So, for the purposes of this review, when the term “mutation” is used it refers to a genetic variant only observed in individuals with a disease, as distinct from a polymorphism. As such, amyloid proteins only associated with polymorphisms will not be considered to have amyloidogenic mutations. With these considerations, we have compiled a comprehensive list of 58 human proteins associated with disease which are capable of forming amyloid fibers (**Table S4-1**).

We can further subdivide this list of proteins into groups relevant to the scope of this review (**Figure 4-1**).

Sporadic amyloids

The first group to delineate is proteins which cause no known hereditary forms of amyloidosis or have any documented hereditary mutations which influence their amyloid aggregation. In other words, proteins for which only the wild-type form is known to form amyloids. This group will be referred to as the “sporadic” group. This term is borrowed from its use in the context of cancer, by which a “sporadic cancer” is a cancer that arises in absence of family history or known genetic variants. 19 of the proteins in our list belong to this group: apolipoprotein A IV, atrial natriuretic factor, calcitonin, cathepsin K, EGF-containing fibulin-like extracellular matrix protein 1, galectin-

7, glucagon, lactadherin, lactotransferrin, leukocyte cell-derived chemotaxin-2, odontogenic ameloblast-associated protein, parathyroid hormone, prolactin, pulmonary surfactant-associated protein C, S100-A8/A9, semenogelin 1, somatostatin, transcription elongation regulator 1 (a.k.a. CA150), and transmembrane protein 106B.

Proto-amyloids

The next group is a class of proteins for which we have coined the term “proto-amyloids”. These are proteins which are only known to form amyloids if and when they have a mutation and the wild-type form does not or cannot form amyloids. In other words, these proteins only cause hereditary amyloid diseases and do not form amyloids without hereditary mutations. The name of this class of proteins is reminiscent of the term “proto-oncogene” which refers to a gene involved in normal cell growth that when mutated becomes an “oncogene” which causes the growth of cancer cells. In an analogous way, the proteins in this group are functioning non-amyloidogenic proteins only until they are mutated, in which case they become amyloids. We currently know of 9 proto-amyloids: integral membrane protein 2B (which is the precursor of ABriPP and ADanPP), apolipoprotein A II, corneodesmosin, fibrinogen α chain, gelsolin, huntingtin, lysozyme, transforming growth factor-beta-induced protein ig-h3, and ubiquilin-2.

Ambimorph amyloids

The next group contains proteins which form amyloid fibers in their wild-type forms and also have documented amyloidogenic mutations associated with disease. Since these proteins form amyloids in both native and mutant forms, we will term this group “ambimorphs” (“both forms”). This group contains the plurality of proteins in the list with 24 members. The proteins in this group cause many of the common amyloid diseases most people are familiar with: Alzheimer’s disease,

Parkinson's disease, ALS, and more. The majority of pathogenic amyloid fibers for which we have solved structures are also from proteins in this group. The ambimorph proteins are amyloid-β, apolipoprotein A I, apolipoprotein C II, apolipoprotein C III, cellular tumor antigen p53, cystatin C, desmin, heterogeneous nuclear ribonucleoprotein A1, heterogeneous nuclear ribonucleoprotein A2, islet amyloid polypeptide, keratin-5, keratin-8, major prion protein, melanocyte protein PMEL microtubule-associated protein tau, polyadenylate-binding protein 2, protein TFG, RNA-binding protein FUS, serum amyloid A, superoxide dismutase, TAR DNA-binding protein 43, transthyretin, α-synuclein, and β2-microglobulin.

Special cases

Lastly, 6 "special cases": antibody light chain, antibody heavy chain, C9orf72, and iatrogenic amyloids. The proteins in this group do not conform to the mutant/wild-type paradigm that exists for the rest of the proteins in the list. The antibody amyloids have completely unique protein sequences from person to person and thus cannot be considered in the context of mutations in conserved protein sequences, although there are some conserved motifs in the most amyloidogenic segments of these proteins(19). These may also be referred to as "personalized" amyloids. Similarly, amyloidogenic dipeptide repeat proteins are produced from a translated intron of the gene C9orf72(20–23) so there is no wild-type version of the protein to compare. Iatrogenic amyloids are proteins that form amyloids when they are used for medical treatment and thus usually at abnormally high concentrations caused by non-physiological forces. These include insulin(24, 25), Liraglutide (recombinant glucagon-like peptide 1 with a K125R alteration)(26), and Anakinra (recombinant Interleukin-1 receptor antagonist protein)(27). There is another peptide drug which is an iatrogenic amyloid, enfuvirtide(28–30), but unlike the ones included in our list enfuvirtide is not an analog of a human protein, so it was not included. One protein that may be considered an iatrogenic amyloid is β2-microglobulin which forms amyloids in dialysis patients

due to being abnormally concentrated because dialysis machines cannot clear this protein from blood plasma(31). However, unlike the other iatrogenic amyloids, an hereditary form of β2-microglobulin amyloidosis exists(32), so it will not be considered as part of this group.

A description of each protein and its evidence for amyloidogenicity can be found in chapter 5.

Structural effects of amyloidogenic mutations

Out of the 58 proteins in our list, there are resolved amyloid fiber structures for 19 of them (excluding the peptide crystal structure of keratin-8). This proportion is actually quite impressive, given how recently cryo-electron microscopy technology started to be used for solving the structures of amyloid fibers, especially patient-derived fibers(10). Out of these 19 proteins, 14 are from the ambimorph group, 4 are from the sporadic group, and 1 is a special case protein (immunoglobulin light chain); there are no fiber structures for any of the proto-amyloids. For the purposes of this review, we will highlight the proteins in the ambimorph group for which there are structures for both the wild-type fiber and the mutant fiber in order to understand which, if any, structural changes of the amyloid form can be accounted for by the mutation and how those changes contribute to the induction of the disease state. There are 7 proteins for which we have this data: amyloid-β, heterogeneous nuclear ribonucleoprotein A2, islet amyloid polypeptide, major prion protein, TAR DNA-binding protein 43, transthyretin, and α-synuclein. Also, the structures which most reliably relate to the disease state are those derived from patient tissue, so if available those are the structures which will be of focus.

Amyloid-β

Amyloid-β is an amyloidogenic peptide which plays a role in Alzheimer's disease, cerebral amyloid angiopathy, and Down syndrome(33–36). It is generated via the secretase-mediated cleavage of amyloid-β precursor protein encoded by the APP gene. The familial mutations in this gene mainly cluster around the sites recognized by the secretases and accordingly they mainly affect the production of the amyloidogenic peptide itself, but the mutant structures that exist show that at least some of these mutations also affect the stability of the resultant amyloid fiber, potentially accelerating their formation. Three mutant forms of amyloid-β fibers have been structurally determined: D23N (pdb IDs: 2lnq, 2mpz)(37, 38), E22Δ (pdb ID: 2mvx)(39), and E22G (pdb IDs: 8bfz, 8bg0)(40) (**Figure 4-2**) (The numbering is based on the amyloid-β sequence, not the amyloid-β precursor protein sequence. Numbering for the amyloid-β precursor protein would be D694N, E693Δ, and E693G). These structures will be compared to wild-type fiber structures which were extracted from Alzheimer's disease patient brains (pdb IDs: 7q4b, 7q4m(41) and pdb IDs: 8azs, 8azt(42)) (**Figure 4-2**). It should be noted that all but one of the E22G mutant fiber structures are made from amyloid-β 1-40, while wild-type structures they are compared to are composed of amyloid-β 1-42. There is a wild-type, patient extracted structure of amyloid-β 1-40 (pdb ID: 6shs)(43), but the fold of its core is completely distinct from all other structures referenced here, so it is difficult to compare to.

The D23N mutation replaces a charged residue with an uncharged one adjacent to another charged residue E22. This change seems to allow for extension of a β-strand from K15 to V24 since residue 23 is no longer charged and is not as disfavored from being buried away from solvent. This allows a single protofilament to form a longer hydrophobic steric zipper than the wild-type is capable of, since in the wild-type structure both E22 and D23 face the same direction and break the β-strand conformation.

The E22Δ mutation similarly removes a charged residue that is adjacent to another one (D23). This seems to compact the turn in the protofilament that E22 would have been a part of and makes some interactions available for the N-terminal of the peptide to take part in, so the entire length of the peptide is resolved in the mutant structure, unlike the wild-type structure. The contribution of more interlayer hydrogen bonding by these extra core residues makes this fiber core more stable, and this more stable structure is made possible by the deletion mutation.

The E22G mutation also replaces a charged residue with an uncharged one. The fold in the double protofilament structure is almost identical to the wild-type fold of the type II filament and the four-protofilament structure does not look immediately more stable than the wild-type structure, especially with the mutant fiber core's large, solvent-accessible central pocket which is very uncharacteristic of amyloid structures. However, the mutation actually allows more flexibility in the adjacent main chain groups (Ramachandran angles) which leads to more interlayer hydrogen bonding in the mutant fiber.

Heterogeneous nuclear ribonucleoprotein A2

Heterogeneous nuclear ribonucleoprotein A2 is ostensibly a functional amyloid. The wild-type protein is able to form reversible amyloids and this behavior is thought to be part of its native function of stress granule formation(44, 45). Its fiber formation is mediated by a low-complexity domain, also called a prion-like domain, and the disease-causing mutation D290V is right in the center of this domain (This numbering is based on the main isoform of the protein which lacks 12 N-terminal residues compared to the longest construct. The numbering of the longest construct

would be D302V). The D290V mutation is found in some patients with multisystem proteinopathy (MSP), a disease with features of inclusion body myopathy associated with Paget's disease of the bone, fronto-temporal dementia, and amyotrophic lateral sclerosis. Biochemical data shows that the mutant protein is more prone to forming fibers than the wild-type and has increased recruitment to stress granules(46), but the reason for these changes is revealed by the structures.

The wild-type fiber structure (pdb ID: 6wqk) consists of a single protofilament with multiple LARKS motifs, creating a fiber with greatly reduced stability compared to one consisting of steric zipper interactions(45). The mutant structure **(unpublished)** reveals that the mutation abolishes some intra-chain hydrogen bonds made by D290 in the wild-type structure and thus encourages more β-strand secondary structure in adjacent residues which influences an increase in β-strand content in the protein overall. Also, the PY-nuclear localization signal peptide is resolved in both the wild-type and mutant structures and in the wild-type it is exposed to solvent, but in the mutant it is buried within the fiber core. This structural change in the mutant protein may explain the increased recruitment to stress-granules since, once formed, it may be more difficult to import the protein into the nucleus, allowing the protein to accumulate in the cytoplasm and aggregate further.

Islet amyloid polypeptide

Islet amyloid polypeptide, also called amylin, is a peptide hormone secreted by β-cells of the pancreatic islets of Langerhans which, like other peptide hormones, is resultant from the processing of a larger prohormone. Its physiological function is not entirely clear, but its main function seems to be related to regulation of glucose metabolism by modifying the activity of insulin(47). Aggregation of this protein has been linked to type II diabetes and it is the main

component of islet amyloid deposits in type II diabetes patients(47–49). A mutation in the IAPP gene, S20G (the numbering for the unprocessed prohormone is S53G), is associated with early onset type 2 non-insulin dependent diabetes(50). Biochemical assays on this mutation show that the mutant protein is more amyloidogenic than the wild-type(51–53) and it has been demonstrated that a possible mechanism is a decrease in the entropy cost of fiber formation by the mutation increasing local flexibility to encourage certain long-range interactions within the peptide that "preorganize" it toward fiber formation(54).

All extant structures of islet amyloid polypeptide fibers are from constructs which are synthetic, recombinant, or synthetic but seeded by patient-extracted fibers. For the purposes of this review, we will focus on a comparison of a wild-type (pdb ID: 6zrf) and two S20G structures (pdb IDs: 6zrq,6zrr) from synthetic peptides reported together in the one publication(55) (**Figure 4-3**). The protofilament common to both mutant structures has a bend at position 20 which turns the backbone in the opposite direction of the backbone of the wild-type protofilament, causing some of the side chains which are buried in the wild-type structure to become solvent-facing and vice versa. This structural change also alters which residues participate in the protofilament interface. The consequences of these structural changes are not entirely clear, but the authors report that the wild-type fiber structure is the most stable out of their three structures, even compared to the mutant structure with 3 protofilaments. The observation that the mutant protein aggregates more rapidly, but forms a less stable fiber is consistent with the idea that the mutation drops the entropic cost of the monomer to transition to a fiber by altering the preferred fold of the monomer rather than driving the protein toward a more stable fiber structure.

Major Prion Protein

Major prion protein is the infectious agent of human prion diseases as well as the subject of mutations which cause inherited human prion diseases. Part of the protein experiences a transition from α-helical structure to β-sheet(56), caused by templating on an already misfolded copy of the protein (infection) or destabilization induced by mutations. The mechanism of sporadic cases of prion disease (sporadic Creutzfeldt-Jakob disease, in particular) are not clear, but may be due to somatic mutations or spontaneous conversion of the protein into the amyloidogenic form(57), but there are risk factors such as homozygosity at codon 129 (rs1799990)(58, 59). There are over 60 known mutations in the PRNP gene (which codes for major prion protein), but ~85% of hereditary prion diseases are caused by five of them: E200K, V210I, V180I, D178N, and P102L(60). Three mutations have their associated amyloid structures solved, but they are not any of the five listed above. The mutations we have solved structures for are E196K (pdb ID: 7dwv)(61), Y145* (pdb ID: 7rl4)(62), and F198S (pdb IDs: 7umq,7un5)(63) (**Figure 4-4**). Only the F198S structure is from patient-derived fibers. It is difficult to directly compare all these structures to the same wild-type structure since they each have different portions of the protein sequence resolved (and the Y145X structure is of a truncated protein, of course). So, to compare the structures which are the most similar in which part of the protein is in the resolved fiber core, the F198S structure and Y145X structure will be compared to the wild-type structure published by Glynn et. al. (pdb ID: 6uur)(64) and the E196K structure will be compared to the wild-type structure published by Wang et. al. (pdb ID: 6lni)(65) (**Figure 4-4**). Another difficulty is that only the E196K structure has the actual mutation site resolved in the structure.

Residue 198 is not actually resolved in the F198S structure, so its influence is difficult to discern on the overall structure, but there is another sequence difference between it and the Glynn wild-type structure: the polymorphism at residue 129. Patients with the F198S mutation who are homozygous for valine rather than methionine at position 129 generally have earlier disease onset

than those heterozygous at that position, and in the patient from which the amyloid fibers were extracted F198S was inherited in cis with valine 129(63). The F198S structure is sterically incompatible with a methionine at position 129 because the valine is packed too tightly. Even with fibers extracted from M/V heterozygous individuals, the electron density map was more consistent with valine rather than methionine at position 129, indicating that M129 prion protein was less likely to be incorporated into the fibers. However, the protein is clearly able to form fibers with a methionine in this position as demonstrated by the wild-type structure. So, residue 129 is probably exerting the most influence over which fiber structure forms. But the familial mutation is at position 198, so what effect is that having on the structure? In fact, the proteins making up the aggregates in Gerstmann–Sträussler–Scheinker disease are proteolytically cleaved and do not even retain residue 198(66). The group reporting the structure performed mass spectrometry on the brain-extracted protein and found fragments corresponding to the full length of the protein without N- and C-terminal signal peptides, but the majority of observed peptides span approximately from G80 to V160(63). The most consistent explanation is that the F198S mutation is affecting the native structure such that it is aberrantly processed and cleaved into amyloidogenic fragments. The valine at position 129 acts as a disease modifier by preferring an amyloid fold which is overall more stable than the wild-type M129 fold, but unable to incorporate monomers with a methionine at position 129. Homozygosity at position 129 more consistently results in the production of fibers with this more stable fold.

Unlike the F1298S structure, the Y145* structure has a methionine at position 129, so can be more directly compared to the Glynn wild-type structure. However, the mutation is a truncation, and while the wild-type structure was solved from a recombinant protein spanning residues 94-178, the structure only resolves residues 106-145, and Y145 is facing solvent and not necessarily as important for the formation of that particular structure as the electrostatic interactions formed

by D144. In the same vein, the Y145* structure, solved from a recombinant protein spanning residues 23-144, does not resolve the truncated C-terminal, starting at residue N108 and stopping at residue F141. Since the wild-type and mutant structures resolve very similar portions of the protein despite being formed from quite different constructs, it is difficult to explain why the structures are so vastly different. The structural differences may simply be a result of the different methods of generating the fibers in each case: the mutant fibers were grown under quiescent conditions at 400μM concentration for a week, sonicated, then used to seed a diluted sample (40μM) under quiescent conditions, and sonicated again; the wild-type fibers were grown at ~100μM with acoustic resonance mixing for 1-3 days. The buffers in which the fibers were grown were also quite different from each other with the wild-type fibers being grown in a slightly denaturing buffer at pH 4, while the mutant fibers were grown in non-denaturing buffer at pH 6.5. These differing growth conditions may do more to explain the differences in observed structures than the different starting protein segments, given that the portion of the protein which ended up in the rigid fiber core was relatively consistent between the two constructs.

The E196K structure and the wild-type structure spanning residues 170-229 were solved and reported by the same group, so their direct comparison is easier and more justified. The mutant structure also contains the mutant residue, so its effects are more observable. The mutation flips the charge of residue 196 from negative to positive, which disrupts an important interaction between the protofilaments in the wild-type structure: an electrostatic interaction between E196 and K194. Due to this, an overall rearrangement of the fiber core occurs to form new electrostatic interactions: K196 now interacts with E200 and K194 now interacts with E207. This change induces other differences including a hydrophilic cavity in the mutant structure instead of a hydrophobic one in the wild-type, additional hydrophobic interactions between the two protofilaments in the mutant structure, and an increase in β-sheet content in the mutant. However,

chemical denaturation and thermostability assays indicated that the mutant fibers were actually less stable than the wild-type fibers. It has been suggested that the effect of the mutation is destabilization of the monomer through disruption of important salt bridges, which accelerates fibrillization(67). The structural differences of the amyloid fibers are, therefore, more indicative of strain differences between individuals with prion diseases and the structure is just one piece of the puzzle of the difference in progression of disease in individuals with this mutation rather than others.

TAR DNA-binding protein 43

TAR DNA-binding protein 43, like the heterogeneous ribonucleoproteins, has a functional aggregated state(68–71). Though not necessarily a functional amyloid, the protein is able to form reversible, potentially amyloid-like, aggregates(72). Disruption of its normal self-interaction within the nucleus seems to be a key mechanism of familial mutations(73) and pathological inclusion of TAR DNA-binding protein 43 are found mislocalized in the cytoplasm(68, 74–78). Pathological inclusions of this protein are comprised of a C-terminal fragment consisting of residues spanning positions 252-414(76). There are more than 60 known mutations in this protein associated with amyotrophic lateral sclerosis (ALS) or frontotemporal lobar dementia with TDP-43 inclusions (FTLD-TDP), but the only mutation for which we have the structure of an amyloid fiber is A315E (pdb ID: 6n3c)(72) (**Figure 4-5**) and this structure was solved from fibers generated from a short fragment of the protein spanning residues 286-331 and the structure itself resolves residues 288-319. This makes comparison to wild-type structures difficult because many of them are solved from fibers made from a different segment of the protein (residues 311-360)(72) or much larger constructs such as the entire low-complexity domain (residues 267-414)(79). Fortuitously, the wild-type structure with the most comparable sequence identity was solved from fibers extracted from two individuals diagnosed with ALS with FTLD (both patients' fibers revealed the same

structure) (pdb ID: 7py2)(80) (**Figure 4-5**). However, the comparison is still complicated by the fact that this patient-extracted structure still resolves a significantly larger fragment of the protein than the mutant structure, residues 282-360, and some of the differences in the structures may be accountable to the differing length of protein fragments forming the fibers rather than the mutation itself.

With these considerations in mind, we can still identify structural changes that can likely be attributed to the mutation. The mutation itself, A315E, is unusual for amyloidogenic mutations because it is a transition from a nonpolar residue to a charged residue, while most amyloidogenic transitions go in the other direction. This new charged residue is complementary to the nearby R293, which is solvent-facing in the wild-type structure. In the mutant structure, E315 and R293 form an electrostatic interaction which requires a significant structural conversion relative to the wild-type structure and creates several new side chain interactions such as L299 participating in hydrophobic interactions with M307, F313, and M311, whereas in the wild-type structure L299 is far away from those residues. Another, perhaps more obvious, difference is that the wild-type structure consists of a single protofilament while the mutant structure consists of four. It is unclear, however, if the high number of protofilaments in the mutant structure would form if a longer construct was used to form the fibers. Regardless, the mutant structure with its multitude of protofilaments creates a more stable fiber than the wild-type in regard to solvation energy.

It is worth noting that E315 may also interact with a different nearby arginine residue R361, in which case a completely different structure may form. However, this cannot be confirmed without structural characterization of either a longer construct of the protein or of fibers extracted from a patient with the same mutation.

Transthyretin

Transthyretin is a thyroid hormone distributor protein which is secreted into the blood by the liver and into the cerebro-spinal fluid (CSF) by the epithelial cells of the choroid plexus(81). Transthyretin functions as a tetramer and binds to thyroid hormones and retinol-binding protein, as well as certain drugs and pollutants(81). The protein circulates in the blood and CSF and when misfolded forms extracellular amyloid deposits in multiple organs, most commonly resulting in cardiomyopathy and/or polyneuropathy(82, 83). There are over 120 documented pathogenic mutations affecting transthyretin, but the most common one is V30M(82, 84) (this numbering excludes the 20 amino acid signal peptide which is cleaved from the mature protein, so an alternative numbering is V50M. The numbering based on the mature protein will be used going forward) and this is the mutant for which we have fiber structures (pdb IDs: 6sdz,7ob4)(85, 86) (**Figure 4-6**) and both are patient extracted. We also have a fiber structure of the wild-type sequence (pdb ID: 8ade)(87) (**Figure 4-6**) which is also patient extracted. The wild-type structure and one of the mutant structures were both extracted from the respective patients' hearts, and so can be directly compared without qualification.

The wild-type and mutant structures are nearly identical, with only slight, inconsequential differences in the packing of some side chains like I107 and K76. Even the mutant structure from fibers extracted from a patient's eye (7ob4) have an extremely similar fold, with major differences being a double protofilament rather than a single one and some altered glutamate-lysine interactions which slightly change the shape of the backbone and open up a hydrophilic cleft in the core; none of these differences are attributable to the mutation which does not alter the orientations of any of the residues near it. The differences could simply be patient-specific fiber

conformations. As Steinebrei et. al. (who published the wild-type structure) points out, these results make it reasonable to conclude that structural differences are not the mechanism of the differential clinical symptoms in sporadic or hereditary transthyretin amyloidoses. Instead, V30M most likely acts by destabilizing the native tetramer(88) to encourage amyloid formation.

α-synuclein

α-synuclein is a protein whose function is not entirely clear, but localizes to presynaptic terminals and interacts with lipid membranes, i.e. vesicles, and is able to adopt an α-helical secondary structure when associated with membranes despite being disordered in solution(89). This protein is the main component of Lewy bodies and Lewy neurites which are the hallmarks of Parkinson's disease (PD) and dementia with Lewy bodies (DLB)(90) and also aggregates in multiple system atrophy (MSA) as well as other "synucleinopathies". Mutations in this protein cause early onset of these diseases and we have amyloid structural data for five mutations (only one mutant structure is from patient-extracted fibers: T22TMAAAEKT) as well as many wild-type structures including some extracted from patients with MSA or PD and DLB. The mutations for which we have structures are H50Q (pdb IDs: 6peo,6pes)(91), A53T (pdb ID: 6lrq)(92), E46K (pdb IDs: 6l4s,6ufr)(93, 94), G51D(pdb ID: 7e0f)(95), and T22TMAAAEKT (pdb IDs: 8bqv,8bqw)(96) (**Figure 4-7**). When comparing structures, we must keep in mind that different diseases caused by α-synuclein harbor distinct amyloid folds, so differences in mutant structures should be viewed in light of this fact. This is complicated by the fact that all but one of the mutations listed above (T22TMAAAEKT being the exception) are associated with early-onset PD or DLB, but the protofilament folds of the H50Q and A53T mutant fibers seem to be more similar to the MSA brain fold (pdb IDs: 6xyq, 6xyp,6xyo)(97) than the PD and DLB brain fold (pdb ID: 8a9l)(98) (**Figure 4-7**), these being the patient-extracted wild-type structures mentioned above. Also, the patient-extracted fiber structures for both MSA and PD/DLB, despite distinct folds, show extra non-

peptide density coordinated mainly by lysine residues (K43, K45, and H50 of both protofilaments in MSA; K32, K34, K43, K45, and Y39 of the single protofilament in PD), but it is not known what cofactor creates this density. The presence of this cofactor may be important for the formation of these specific patient structures, and all recombinant structures must be presumably qualified by the exclusion of this unknown cofactor.

The H50Q mutant α-synuclein forms a fiber with a very similar fold to recombinant wild-type structures (pdb IDs: 6cu7,6cu8)(99) (**Figure 4-7**) and the patient-extracted MSA fold (also wild-type), but has a distinct protofilament interface or, in one isoform, only a single protofilament. The mutation may contribute to this altered interface by abolishing a potential electrostatic interaction between H50 and E57, as seen in the 6cu7 structure, and instead Q50 forms an intramolecular hydrogen bond with K45(91). Q50 may also induce a steric clash with E57. However, based on this structure, the stabilities of the mutant and wild-type structures are comparable to each other. Although, denaturation resistance assays establish a higher stability of H50Q fibers than wild-type(100), so the structure of the fiber core may not be the only contributor to overall stability. It is unclear why this mutation favors a fold more similar to the MSA patient fold than the PD patient fold.

The A53T mutant forms a fiber which is extremely similar to those formed by the H50Q mutant, although superstructural features, like overall fibril twist, differ(92). In some recombinant wild-type structures, e.g. 6cu7, A53 is at the center of the protofilament interface and the A53T mutation directly disrupts this interface through a steric clash, favoring the interface also seen formed by the H50Q mutant protein. Again, this interface appears less stable than the wild-type interface, making the increased aggregation propensity of the mutant protein(92, 101, 102) difficult to

explain by the structure of the resulting fiber alone. And, again, it is unclear why this mutation favors a fold more similar to the MSA patient fold than the PD patient fold.

Two distinct structures have been solved for fibers with the E46K mutation. These two structures even have apparently different biochemical properties, with the 6l4s fibers being apparently less stable than wild-type fibers(94) while the 6ufr fibers are apparently more stable than wild-type fibers(93). However, both structures represent a departure from the fold observed in the H50Q mutant structure, the A53T mutant structure, some recombinant wild-type structures, and the patient-extracted MSA structure. The E46K mutation disrupts a salt-bridge formed by E46 and K80 that is present in the aforementioned structures. In the E46K structures, the removal of this electrostatic interaction results in a complete rearrangement of the core. However, the PD and DLB patient-extracted structure does not contain the E46-K80 interaction, despite not having this mutation, and has a distinct fold to both E46K structures. So, the particular arrangement of the fiber core seen in these recombinant mutant structures cannot be entirely attributable to the mutation, especially since the mutated residue does not seem to participate in any structure-specific interactions. The differences between the two mutant structures may be attributable to differences in the preparation of fibers, i.e. different buffers, different protein concentrations, and the 6l4s fibers were grown in the presence of preformed fiber seeds. And, again, neither structure was obtained from fibers grown in the presence of whatever cofactor creates the density in the 8a9l and 6xyq/6xyp/6xyo structures. The coordination of lysines around this cofactor may preclude the K45-E57 interactions seen in both E46K structures. In all, the mutation explains certain structural differences from the human-extracted MSA structure as well as recombinant wild-type structures bearing an E46-K80 interaction, but cannot explain the differences from the wild-type PD and DLB structure.

The G51D mutant structure is extremely similar to the 6l4s E46K structure, both fibers even bearing a unique right-handed twist(95), although D51 abolishes a turn present in the E46K structure which allows the interaction between K45 and E57. Also, like the 6l4s fibers, the G51D fibers were biochemically shown to be less stable than wild-type fibers, but also more cytotoxic and more seeding-competent(95). A decrease in stability has been hypothesized to increase propagation of seeds for other mutations as well(92, 94). Like the A53T mutation, the G51D mutation creates a steric clash that would directly disrupt the protofilament interface of some recombinant wild-type fibers, but it is not entirely clear why these two mutations induce such distinct structures from each other. The G51D mutation may also be somewhat incompatible with the patient-extracted PD structure, since it could form a steric clash with the undefined peptide density interacting with the H50-V55 region, but it is unclear how important this interface is to the overall structure of that fiber core. Also, none of the lysines coordinating the mystery density in the patient-derived structure are resolved in the G51D structure, so it is unclear if they could be part of the ordered core if the cofactor were present or if D51 somehow precludes this interaction.

The T22TMAAAEKT mutation is a 7 amino acid insertion after T22 which was found in a patient with extremely early onset of disease (13 years old) and rapid progression (2 years from manifestation of symptoms to death)(96, 103). This was the only case of a disease now known as juvenile-onset synucleinopathy, so named distinctly from DLB due to the discovery of the unique fold of α-synuclein fibers extracted from the patient's brain. The fiber consists of a single protofilament with a fold very similar to the patient-extracted MSA fiber, containing the salt bridge formed from E46 and K80. The difference is that the protofilament interface is supplanted with an interaction with the N-terminal region of the protein harboring the insertion mutation. Like the other

patient-extracted structures, this structure also contains a network of lysines coordinating an undefined density. The peptide density which forms the interface with H50-E57 is actually too ambiguous to definitively model in the mutant sequence, but the extra length of the mutant sequence is better able to fill in the other island of ambiguous peptide density forming an interface with K60-T64. It is also possible that the fiber contains both mutant and wild-type protein, leading to the ambiguity in the cryo-em data. The mutation probably contributes to this unique fold by making the intra-protofilament interfaces revealed in this structure easier to form by lengthening the protein with a repeated sequence. This mutation clearly induces a particularly aggressive synucleinopathy, but the structure does not immediately allow us to intuit why.

Biochemical mechanisms of amyloidogenic mutations

The way that genetic mutations can contribute to the formation of amyloids is varied and usually multifaceted for each individual protein and even for individual mutations. The mechanisms we have identified here also tend to tie into one another and thus overlap somewhat, so that mutations which are amyloidogenic through a given mechanism may also work through another mechanism entailed by the first. Also, proteins with multiple amyloidogenic mutations may become amyloidogenic through multiple distinct mechanisms depending on which mutation it has. For this review we've focused solely on mutations which directly affect either the amino acid sequence of the amyloidogenic protein or the amount of the amyloidogenic protein being expressed, although there are amyloidogenic proteins whose aggregation is ostensibly influenced by mutations in other genes, such as PSEN1 and PSEN2 in Alzheimer's disease(104–107). We have delineated six classes of mutational effects which contribute to the conversion of proteins into amyloid aggregates (**Figure 4-8**). The mechanisms by which mutations lead to amyloidogenesis of the proteins in our list can be found in **Table S4-1** and we have assigned

mechanisms to 33 of the 58 proteins in our list: the proto-amyloids and ambimorphs since these are the proteins whose amyloid formation is known to be mediated by mutations.

i) Native structure destabilization

Mutations which change the amino acid sequence can destabilize a protein's native fold or interrupt its normal interactions with copies of itself, both of which can lead to amyloidogenic portions of the protein which would normally be buried becoming "exposed" and available to nucleate the growth of a fiber. The mutation does not even have to be in the amyloidogenic portion of the protein, such as in the case of L170P in apolipoprotein AI: a mutation outside of the N-terminal fragment found in amyloid deposits which has been shown to destabilize the structure of the protein in solution, and this may induce misfolding which leads to proteolytic degradation of the protein which frees the amyloidogenic fragment(108, 109). Monomer/oligomer destabilization is a fairly common mechanism of amyloid formation, with 16 of the 33 proteins having this mechanism tied to their amyloid formation. It is also an aspect of how amyloids are formed in the lab, although this usually is not achieved through mutations, but rather environmental manipulations like fluctuating the pH or adding cofactors like heparin in the case of tau(13, 110). Some examples of proteins with mutations of this mechanism are transthyretin, whose mutations destabilize its native tetramer and result in a fiber core fold almost identical to the wild-type protein(85–88), and lysozyme, whose amyloidogenic mutations were shown by molecular dynamics simulations to increase flexibility of the protein and allow it to sample more conformations(111).

ii) Fibril stabilization

This applies to mutations which induce structural changes which promote fibrillization including better shape complementarity for a steric zipper, improved ability to form interstrand hydrogen bonds, better charge compensation in a protofilament fold, etc. These effects can be observed directly in fiber structures or indirectly through comparing the kinetics of amyloid fiber formation between wild-type and mutant protein. This often goes hand-in-hand with monomer destabilization since for globular proteins the monomer form is usually far removed from the structure of an amyloid fiber. In order for the transition to a fiber to be kinetically favored, even if it is more thermodynamically stable than the monomer, the monomer structure must first be disrupted. This category also encompasses stop codon mutations which generate amyloidogenic segments essentially *ab initio*, such as in the cases of the ABriPP/ADanPP and apolipoprotein A II(112–118). It is easier to determine if this mechanism is at play if a fiber structure exists with the mutant residue in a stabilizing role, as in the case of G269V and P285L in protein TFG **(unpublished)** and E22G in amyloid-β(40). The fibril stabilization mechanism applies to mutations in 20 of the 33 proteins in our list, with 8 of them also having monomer destabilization as a mechanism. The proteins which do not have these two mechanisms working in tandem are peptides like amyloid-β or proteins with intrinsically disordered or low-complexity regions like heterogeneous nuclear ribonucleoprotein A2, in which cases there is no stable monomer structure to begin with. Mutations which stabilize amyloid fibers tend to be changes to hydrophobic residues which prefer β-strand secondary structure, especially if the wild-type residue is polar or charged, such as D25V in apolipoprotein C III(119), but there are exceptions to this trend such as A120D in desmin(120).

iii) Increased seeding

Amyloid formation starts with the nucleation of a single fiber, which is an ostensibly rare event, but prion-like seeding causes the growth of many more fibers relatively quickly after the first is able to form. Seeding capacity influences the spread of amyloid fibers throughout the tissue they inhabit(121–123), which is why blocking seeding is often a metric which is considered for amyloid inhibitor compounds(124–126). The seeding capacity can be influenced by the efficiency of production of seeds or the efficacy of the seeds themselves in reducing the lag phase of fiber nucleation. Whether mutations affect seeding or not is difficult to know, however, and requires biochemical experiments to specifically test it. Thus, only 3 of the 33 proteins taken into account here have this mechanism assigned to their mutations: desmin(127), cellular tumor antigen p53(128), and α-synuclein(92, 94, 129). It is likely that many of the known amyloidogenic mutations affect seeding capacity, especially if the mutation also causes one or both of the two effects described above since they would already lower the energetic barrier to nucleation (so monomer destabilization and/or fiber stabilization might actually entail increased seeding capacity), but this is difficult to infer or computationally predict, unlike the other two effects previously mentioned. Paradoxically, decreased fiber stability may also increase seeding, as seems to be the case for α-synuclein, since a propensity to break apart into smaller oligomers (which may just be short fibers) might be beneficial to the propagation of seeds.

iv) Altered processing

This category refers to changes in how the protein's mRNA, its gene's DNA, or the protein itself is treated by factors extrinsic to the protein itself. Essentially, this mechanism applies to mutations which alter how cellular homeostasis machinery interacts with the protein at the DNA, transcript, and/or protein level. This can be overexpression, changes in which isoforms of the protein are produced, and/or differences in proteolytic processing. This is the only category which includes

noncoding mutations, such as intronic mutations that affect splicing as in the case of microtubule-associated protein tau(130–135), promoter mutations which affect expression levels as in the case of serum amyloid A(136) and islet amyloid polypeptide(137), or gene duplications which also affect expression levels as in the case of α-synuclein(138, 139). Overexpression, caused by promoter mutations or gene duplications, can induce amyloid formation by allowing the protein to achieve a critical concentration to form an amyloid without alteration to the amino acid sequence. This is also the mechanism by which the iatrogenic amyloids become amyloidogenic, although they are not overexpressed, but directly injected at a high local concentration. Microtubule-associated protein tau is a good case study for how altered isoform ratios can lead to amyloid formation. There are documented mutations which affect the production of tau protein isoforms with four tandem repeats in relation to isoforms with only three, many of which do not affect the amino acid sequence(130–135). However, it is not entirely clear why this change leads to amyloid aggregation. It is theorized that there is limited availability for each isoforms of tau to bind to microtubules(131, 135) and while not bound to a microtubule tau exists in a disordered state. So, when the precise regulation of tau isoform ratios is disrupted, excess unbound tau can build up in the cytoplasm and accumulate post-translational modifications like phosphorylation and acetylation leading to its amyloid aggregation(131). Changes in proteolytic processing usually leads to amyloid formation by generating an amyloidogenic peptide from a protein that may not be able to form an amyloid while the full-length form is intact. A good example of this mechanism is in the case of the amyloid-β precursor protein: mutations in the gene encoding this protein, APP, can affect how the protein is cleaved by secretase complexes (increased cleavage by β-secretase and/or γ-secretase instead of the benign α-secretase) and lead to the release of excessive amounts of the amyloid-β peptide which is highly amyloidogenic(33, 106, 140, 141). In fact, presenilin 1 and presenilin 2 (proteins besides amyloid-β which harbor familial Alzheimer's disease mutations) are part of the γ-secretase complex(105, 107, 142). Amyloidogenic peptides

like ABriPP and ADanPP also fall into this category because, although the amyloidogenic peptide is produced by the same furin-mediated cleavage process that the normal protein undergoes, the process seems to be enhanced by the mutations for reasons that are not entirely clear(143). 12 out of 33 proteins have been identified as having amyloidogenic mutations of this category.

v) Subcellular mislocalization

Some proteins move between different subcellular compartments as a part of their function. The most relevant examples would be RNA-binding proteins like RNA-binding protein FUS and TAR DNA-binding protein 43, both of which primarily reside in the nucleus but move to the cytoplasm and form phase-separated stress granules in response to cellular stress such as heat shock, oxidative stress, or chemical stress(68–71, 144–148). Mutations in these and other proteins can interfere with nuclear localization, trapping the proteins in the cytoplasm where they remain at high concentrations inside stress granules long enough to begin to irreversibly aggregate into amyloids(46, 68, 74, 75, 149–151). 4 of the 33 proteins in the list have mutations of this category. Only one of these proteins is not a nucleic acid-binding protein: desmin. Although, for desmin, the mislocalization may result from changes in its ability to form an intermediate filament network and aggregation(120, 152). Indeed, for all of these proteins it is not entirely clear if mislocalization precedes amyloid aggregation or if it is a result of amyloid aggregation.

vi) Decreased binding to native partner(s)

Mutations of this class disrupt the ability of the protein to bind to its normal binding partners. The binding partners of proteins in our list with this class of mutations are lipids, metal ions, and other proteins. This category does not include mutations which disrupt binding to copies of the identical

protein, since that is considered disruption of native structure (the first category). Still, this category is closely related to the first category (native structure destabilization) since, in most cases, the native structure would need to be disrupted in order to abolish its normal binding. Also, decreased binding affinity to native substrates can induce structure destabilization. In fact, all mutations of this category in our list also destabilize the native structure except for mutations in microtubule-associated protein tau, whose microtubule-binding regions are natively disordered when not bound to microtubules(153), ubiquilin-2, whose mutations occur in a disordered region of the protein(154), and α-synuclein, which is intrinsically disordered when not bound to lipid membranes(155, 156). 9 of the 33 proteins in the list have mutations in this category. 4 of those proteins are apolipoproteins, which are amyloidogenic when unbound to their lipid binding partners, but are resistant to aggregation in the bound wild-type state(109, 115, 116, 119, 157–161). The others are gelsolin, whose binding to calcium ions and actin filaments is disrupted by mutations(162); superoxide dismutase, whose binding to zinc ions is disrupted by mutations(163); microtubule-associated protein tau, whose mutations disrupt binding to microtubules(135); ubiquilin-2, whose mutations disrupt binding to ubiquitinated protein substrates(164); and α-synuclein, whose mutations disrupt binding to lipid membranes(165, 166).

One protein in the list has an amyloidogenic mutation, but the mechanism is not clear from the literature: keratin-5. Keratin-5 is one of the main components of amyloid deposits in primary localized cutaneous amyloidosis, but there are no mutations in this protein associated with this disease, although familial forms associated with mutations in other proteins do exist(167). Mutations in keratin-5 are rather associated with another disease not normally associated with amyloid: epidermolysis bullosa simplex (EBS). Keratin-5 has a mutation (V324A) associated with a case of Weber-Cockayne type of epidermolysis bullosa simplex (EBS-WC) with amyloid deposits, which, again, is an anomalous symptom for EBS even in cases with other mutations in

KRT5(168). Due to the unique amyloid phenotype associated with this keratin-5 mutant, it is likely that the mutation is driving the amyloid formation, but no structural or biochemical data is available in the literature and there is no speculation of this mutation's amyloidogenic influence.

Inheritance patterns and haplosufficiency

One unifying feature of amyloidogenic mutations is their dominant inheritance patterns. The vast majority of mutations affecting the proteins in our list are inherited in an autosomal dominant fashion, although there are exceptions such as H517Q in RNA-binding protein FUS and D90A in superoxide dismutase. This is of note because another common feature of the proteins in our list is their haplosufficiency. A haplosufficient gene is one which can tolerate a loss of function of at least one of its copies; a haploinsufficient gene is one for which a loss of function of only one copy is sufficient to cause disease. Only 7 of 54 proteins' genes (immunoglobulin light and heavy chain and iatrogenic amyloids Liraglutide, insulin, and Anakinra were excluded from analysis and S100A8 and S100A9 were scored separately from each other) were scored as haploinsufficient based on the metrics available from the Genome Aggregation Database (gnomAD) (Figure 9). These proteins are EGF-containing fibulin-like extracellular matrix protein 1, heterogeneous nuclear ribonucleoprotein A1, heterogeneous nuclear ribonucleoprotein A2, huntingtin, RNA-binding protein FUS, TAR DNA-binding protein 43, and transcription elongation regulator 1 (a.k.a. CA150). α-synuclein also comes very close to the threshold, but cannot quite be scored as haploinsufficient. Interestingly, this list encompasses the four RNA-binding proteins in our list which form functional aggregates in vivo. Out of the other proteins, CA150 and huntingtin are both associated with Huntington's disease and EGF-containing fibulin-like extracellular matrix protein 1 seems to be a unique case.

Since most of the haplosufficient proteins in our list are also subject to mainly dominantly inherited mutations, it is unlikely that these mutations are straightforward loss-of-function mutations. This is not to say that this means all amyloid fibers are cytotoxic, just that the aggregation of mutant protein into amyloid fibers is unlikely to solely affect the mutant version of the protein. These mutations must either induce a toxic gain-of-function or cause a dominant negative effect, and amyloid fiber formation can conform to either or both of these mechanisms. A toxic gain-of-function means that the protein's activity is either increased or altered in such a way that is toxic to cells. Amyloid fibers have been shown to be directly toxic to cells in many cases(22, 47, 91, 95, 127, 147, 169–172), and so can be considered a toxic gain-of-function. It is unlikely that amyloid fibers retain, let alone enhance, the original function of the protein due to being in a misfolded state. A dominant negative mutation is one which interferes with the function of the protein produced by the unaltered copy of the gene. This usually manifests in proteins which must polymerize to carry out their function, since addition of mutated substituents can interfere with the assembly and function of protein polymers. Mutant amyloid fibers may be able to incorporate wild-type protein into growing fibers(173–175) which constitutes a dominant negative effect. Both of these effects can manifest at once in mutant/wild-type hybrid fibers that are also toxic to cells, but the dominant negative effect of amyloid formation may be sufficient to cause disease even if the fibers themselves are inert.

Another consideration for haplosufficient amyloid proteins is that there could, in theory, be a case where a proto-amyloid forms non-toxic, inert amyloids which knock out the mutant protein's function but do not incorporate the wild-type protein and thus does not have a toxic gain-of-function or dominant negative effect. In this case, an individual heterozygous for an amyloidogenic mutation in a haplosufficient proto-amyloid would have no disease and that protein's

amyloidogenic behavior could go unnoticed. It is not clear if any human proteins actually conform to this scenario in reality, however.

In haploinsufficient genes, the pathogenic effect of amyloid formation due to dominant mutations can be more ambiguous. Since the sequestration of mutant protein into amyloid fibers essentially knocks out its function, even if the fibers are inert and only contain mutant protein, this would be sufficient to cause disease. Or even if the misfolded proteins are able to be degraded by proteasomes, the consistent misfolding of the mutant protein would significantly hamper its function and lead to disease.

In any case, amyloid formation is a consistent explanation for disease progression since amyloid diseases are frequently late onset. Amyloid formation is a stochastic, probabilistic process which requires a rare nucleation event to occur and also propagate enough to become resistant to cellular proteostasis machinery, and so late onset is expected for this kind of process. One major disease-modifying effect of mutations is the acceleration of disease onset, which can be attributable to mutations making nucleation events less rare through the various mechanisms already discussed. These mechanisms are also consistent with the observation that the vast majority of amyloid mutations are dominantly inherited even though the majority of amyloid proteins are haplosufficient.

Why study mutant structures of amyloid proteins?

The mechanisms which lead to amyloid formation are diverse. One benefit of studying the mutant versions of amyloid fibers found in disease is that it can give us strong clues as to what

mechanisms are relevant to the formation of the amyloid fibers of particular proteins based on what effect the mutations have. For example, the transthyretin structures reveal that the influence of the V30M mutation on the native structure is more relevant than its effect on the fiber structure since the resulting fiber structure is nearly identical to the wild-type. Destabilization of the native tetramer, rather than the stability of the amyloid fiber, is driving this mutation's association with disease, and so sporadic disease is likely related to destabilization of transthyretin's native structure through other mechanisms. In another case, the TAR DNA-binding protein 43 structures reveal that the A315E mutation induces a significant structural change compared to wild-type fibers which allows for a structure with more stability. In this case, stabilization of the amyloid state, which makes its formation more thermodynamically favorable, seems to be the relevant factor for the association of the mutation with disease. This means that, in sporadic cases, factors leading to disease are likely also allowing for stabilization and persistence of aggregated TAR DNA-binding protein 43. Elucidating the factors affecting transthyretin and TAR DNA-binding protein 43 in these ways in sporadic cases should be goals of future research, and these factors would be prime therapeutic targets.

Another practical application of these structures is the ability to design structure-based aggregation inhibitors. Inhibitors of this nature have already been designed for microtubule associated protein tau(124, 125, 176), amyloid-β(124, 177), and α-synuclein(124, 126). The effectiveness of structure-based inhibitors may be restricted to forms of disease in which the same fiber structures form, but familial forms of amyloid diseases tend to exhibit different amyloid structures than their sporadic counterparts. If these kinds of inhibitors become mainstream treatment options, structural information on familial forms of amyloid diseases will be necessary in order to ensure effective inhibitors exist for those forms of disease. In this vein, it is of note that no fiber structures have been solved thus far for any of the proto-amyloids. The diseases caused

by this class of proteins may be most amenable to structure-based therapeutics since there may be less variety in amyloid structures in those diseases due to the lack of sporadic occurrences. Thus, our body of knowledge on mutant amyloid fibers has plenty of room to grow.

Also, solving mutant structures is an important pursuit in the context of basic science. Amyloid fibers already subvert traditional biological assumptions about how protein sequence relates to structure and function, and mutant fibers further complicate the picture by adding more structural diversity to the amyloid fibers themselves. Gathering knowledge on the range of structural diversity for amyloid proteins is worthwhile for its practical applications, but also purely for increasing our body of knowledge regarding these proteins. Solving mutant structures helps to give us a more complete picture of the complex biological systems that are amyloid proteins.

Tables and Figures

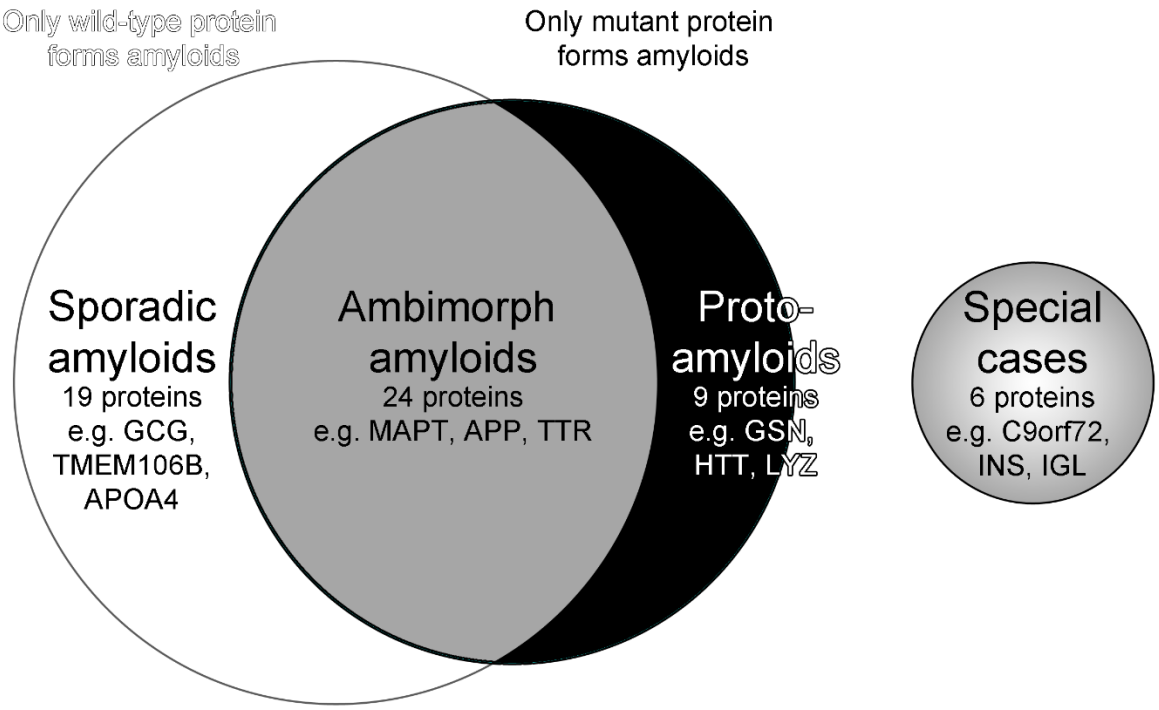


Figure 4-1: Classification of pathogenic amyloid proteins. A Venn diagram representation of how proteins capable of forming pathogenic amyloids can be classified based on whether their wild-type or mutant forms exhibit amyloidogenicity. Sporadic amyloid proteins do not have any known naturally occurring amyloidogenic mutations and their wild-type forms are able to form amyloid fibrils. Ambimorph amyloid proteins are found to form amyloid fibrils in their wild-type form as well as with disease mutations. Proto-amyloid proteins are only known to form amyloid fibrils when containing a mutation and their wild-type forms do not or cannot form amyloid fibrils. Special cases are proteins which do not necessarily have wild-type and mutant forms in the same sense as the rest of the proteins, including immunoglobulins, C9orf72 which is an aberrantly expressed intronic region, and iatrogenic amyloids.

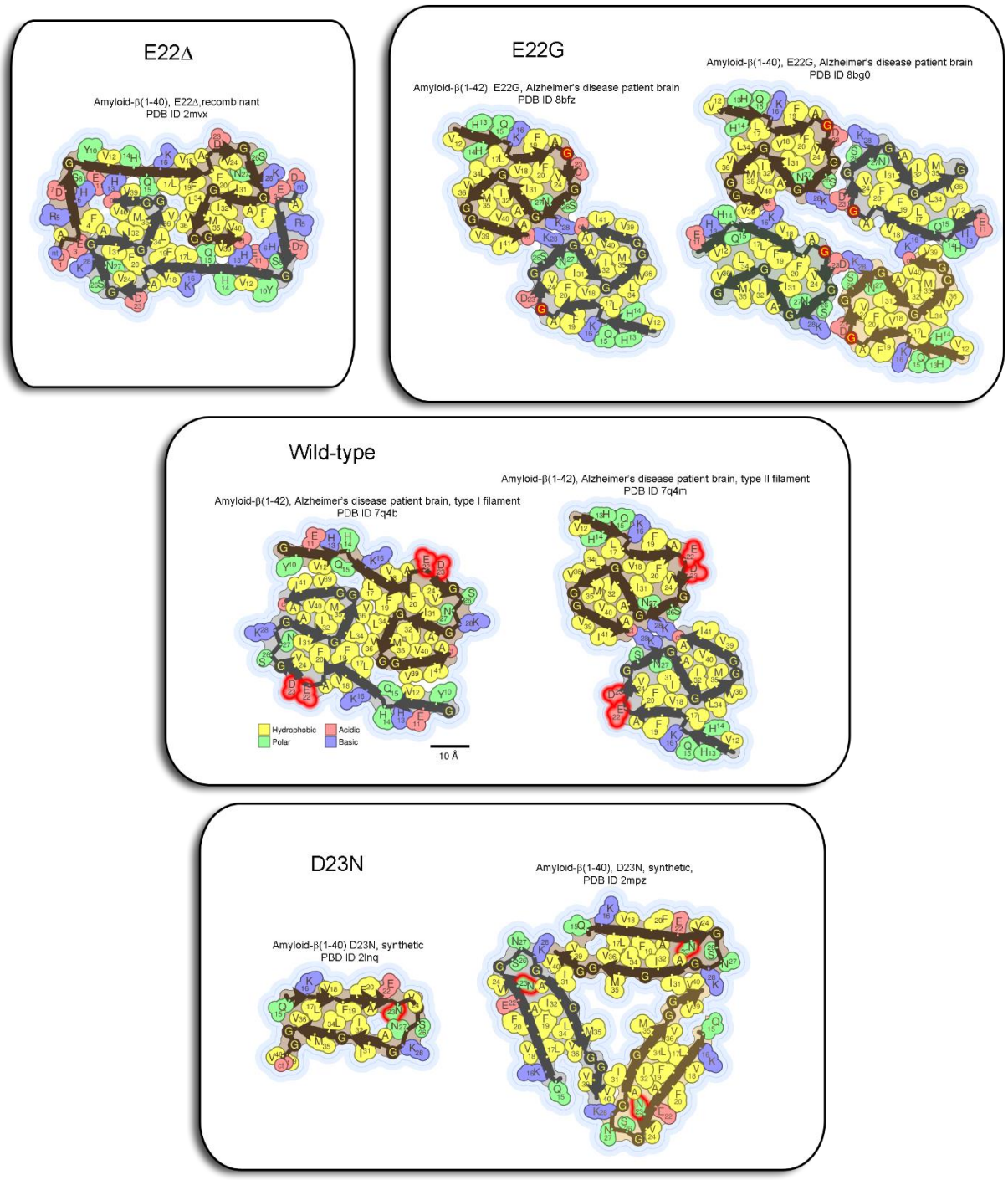


Figure 4-2: Fiber structures of amyloid-β. Structures of amyloid fiber cores of wild-type, E22Δ, E22G, and D23N amyloid-β. Structures are grouped in boxes based on which mutation they have and wild-type structures are in the centermost box. In the wild-type structures, residues which are

highlighted with a red outline correspond to the mutated residues in the mutant structures. In the mutant structures, the residue highlighted with a red outline corresponds to the mutant residue in that structure.

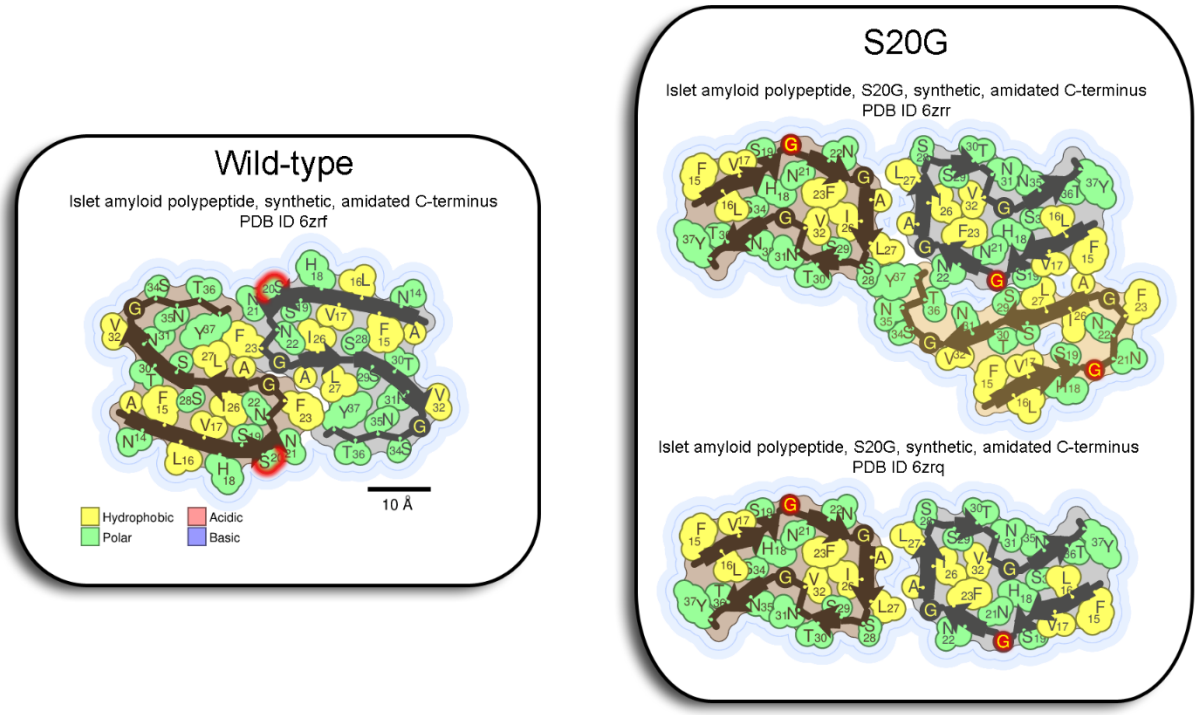
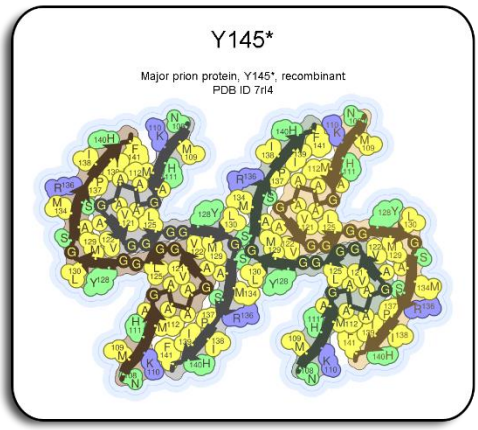
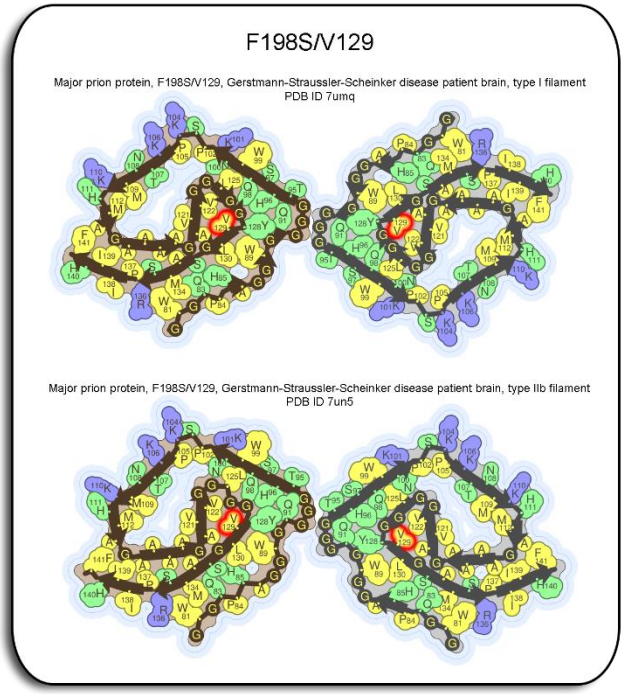
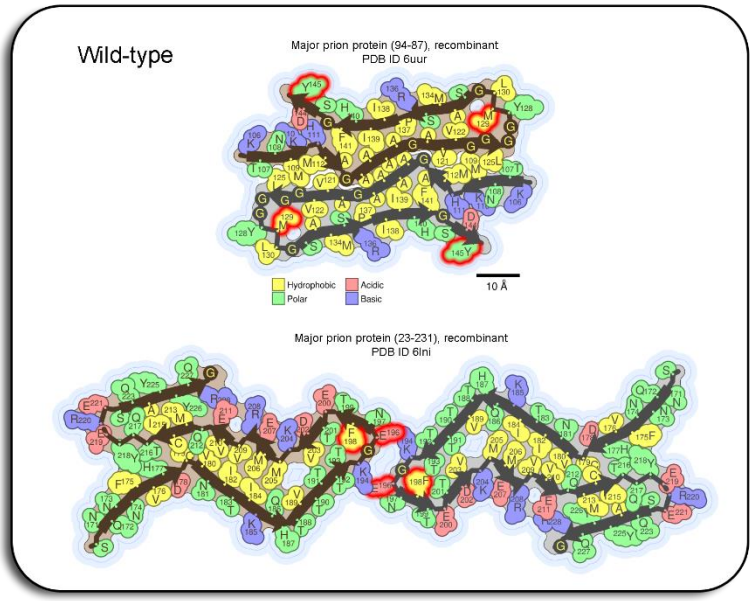


Figure 4-3: Fiber structures of islet amyloid polypeptide. Structures of amyloid fiber cores of wild-type and S20G islet amyloid polypeptide. The wild-type structure is in the left box and the mutant structures are in the right box. In the wild-type structure, residues which are highlighted with a red outline correspond to the mutated residues in the mutant structures. In the mutant structures, the residue highlighted with a red outline corresponds to the mutant residue in that structure.


Y145*
Major prion protein, Y145*, recombinant
PDB ID 7rl4


F198S/V129
Major prion protein, F198S/V129, Gerstmann-Straussler-Scheinker disease patient brain, type I filament
PDB ID 7umq
Major prion protein, F198S/V129, Gerstmann-Straussler-Scheinker disease patient brain, type IIb filament
PDB ID 7un5


Wild-type
Major prion protein (94-87), recombinant
PDB ID 6uur
Hydrophobic
Polar
Acidic
Basic
10 Å
Major prion protein (23-231), recombinant
PDB ID 6lni

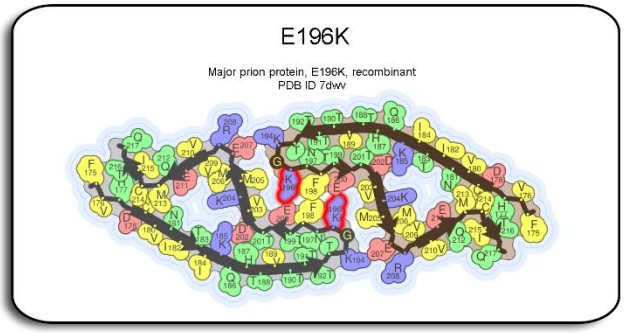

E196K
Major prion protein, E196K, recombinant
PDB ID 7dwv

Figure 4-4: Fiber structures of major prion protein. Structures of amyloid fiber cores of wild-type, Y145*, F198S/V129, and E196K major prion protein. Structures are grouped in boxes based on which mutation they have and wild-type structures are in the centermost box. In the wild-type structures, residues which are highlighted with a red outline correspond to the mutated residues in the mutant structures. In the mutant structures, the residue highlighted with a red outline corresponds to the mutant residue in that structure. In this case, for the PDB ID 6uur wild-type structure and the F198S/V129 mutant structures, residue 129 is highlighted despite being a polymorphism because it is relevant to the structural differences.

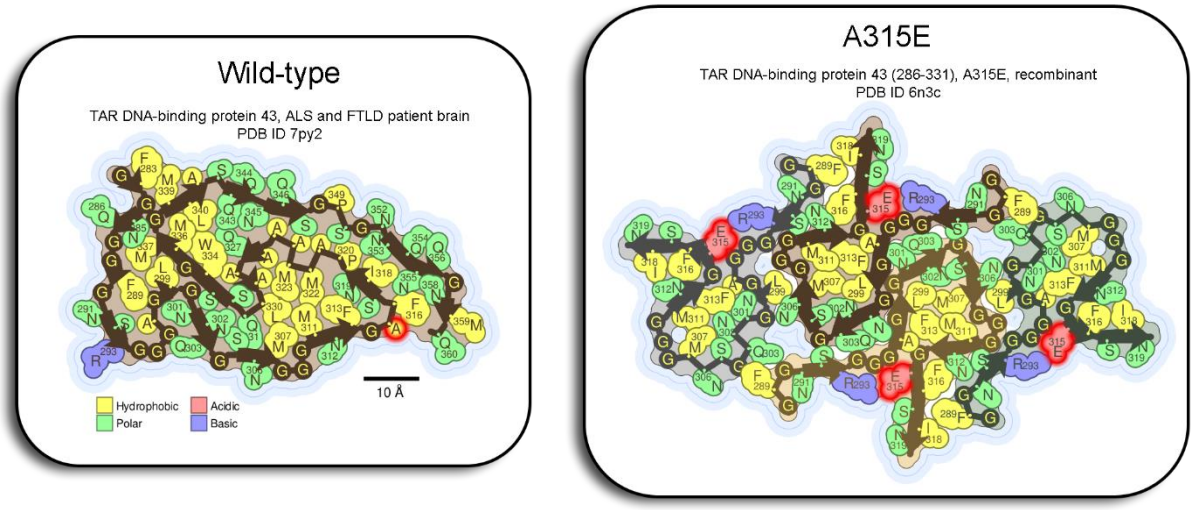


Figure 4-5: Fiber structures of TAR DNA-binding protein 43. Structures of amyloid fiber cores of wild-type and A315E TAR DNA-binding protein 43. The wild-type structure is in the left box and the mutant structure is in the right box. In the wild-type structure, the residue which is highlighted with a red outline corresponds to the mutated residue in the mutant structure. In the mutant structure, the residue highlighted with a red outline corresponds to the mutant residue in that structure.

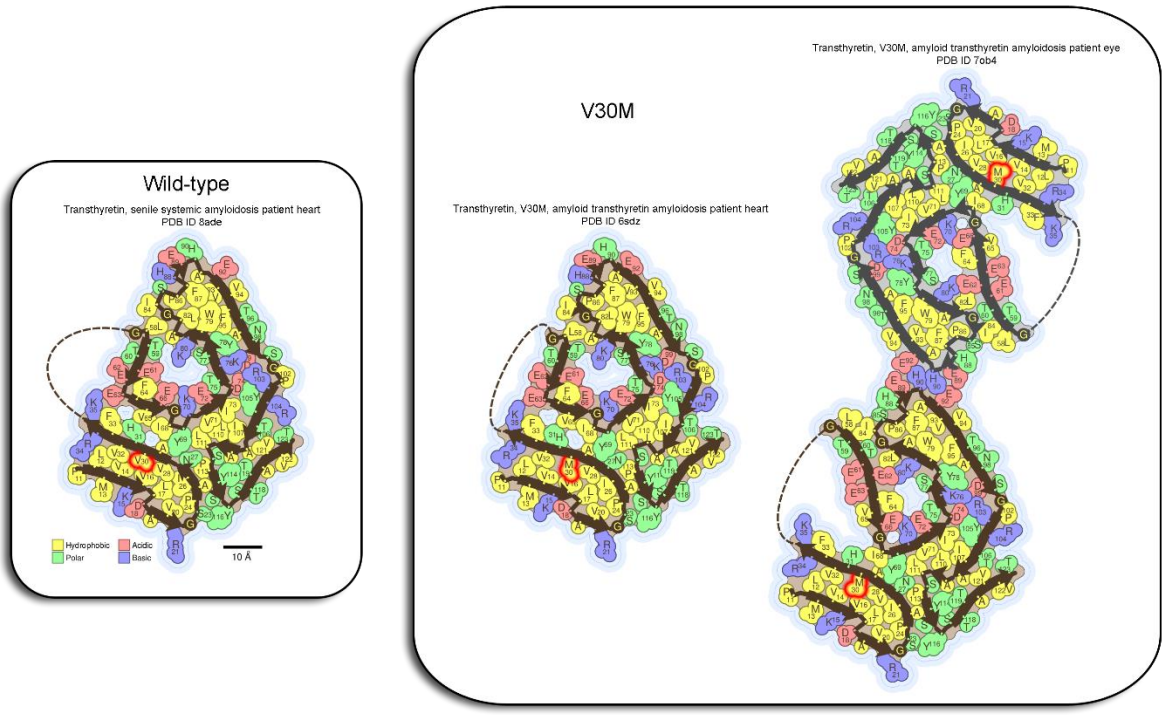


Figure 4-6: Fiber structures of transthyretin. Structures of amyloid fiber cores of wild-type and V30M transthyretin. The wild-type structure is in the left box and the mutant structures are in the right box. In the wild-type structure, the residue which is highlighted with a red outline corresponds to the mutated residue in the mutant structures. In the mutant structures, the residue highlighted with a red outline corresponds to the mutant residue in that structure.

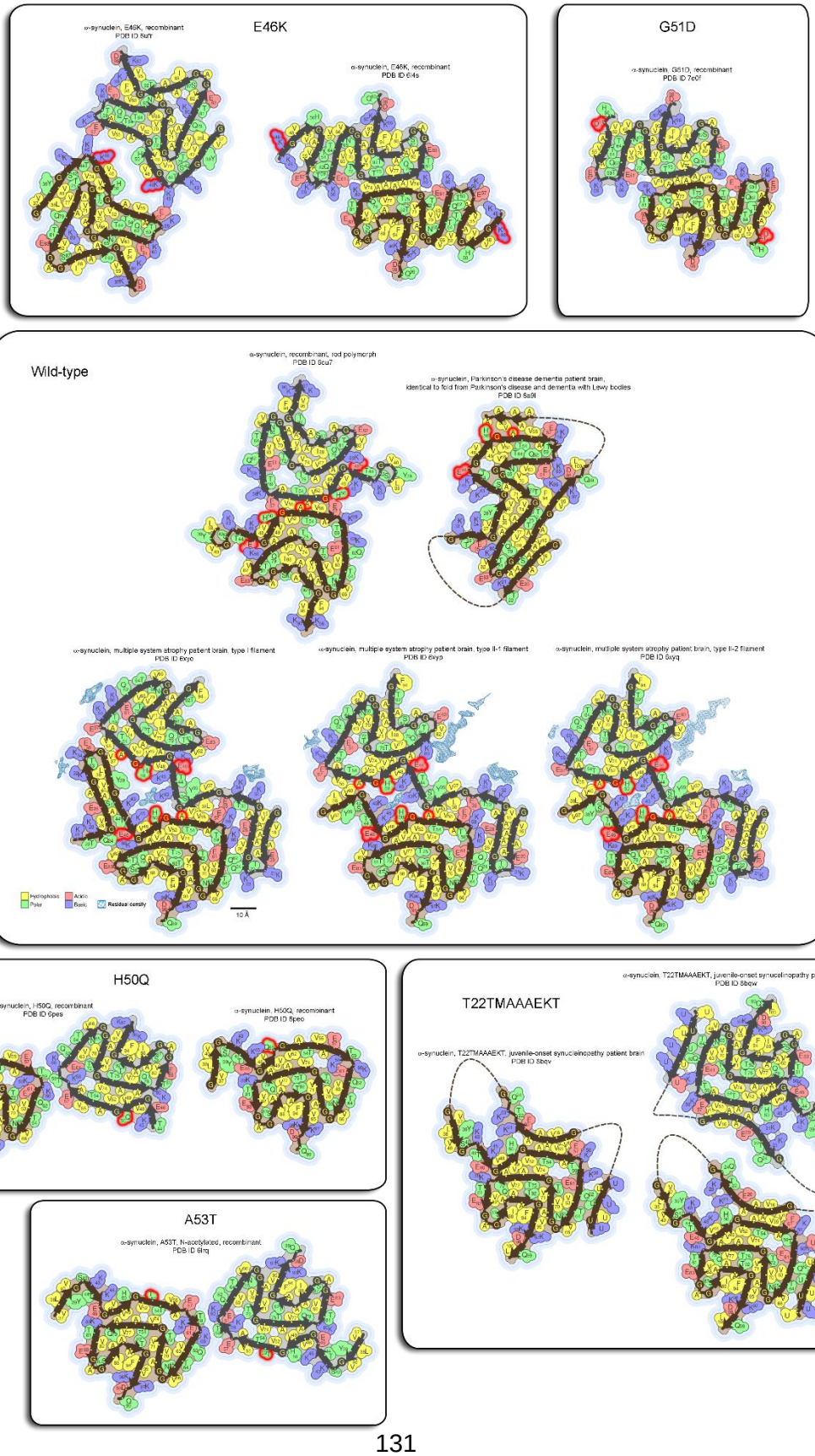

E46K
α-synuclein, E46K, recombinant
PDB ID 6uiy
α-synuclein, E46K, recombinant
PDB ID 6l4s
G51D
α-synuclein, G51D, recombinant
PDB ID 7e0f
Wild-type
α-synuclein, recombinant, rod polymorph
PDB ID 6cu7
α-synuclein, Parkinson's disease dementia patient brain, identical to fold from Parkinson's disease and dementia with Lewy bodies
PDB ID 6a9l
α-synuclein, multiple system atrophy patient brain, type I filament
PDB ID 6xyc
α-synuclein, multiple system atrophy patient brain, type II-1 filament
PDB ID 6xyp
α-synuclein, multiple system atrophy patient brain, type II-2 filament
PDB ID 6xyq
Hydrophobic
Polar
Acidic
Basic
Residual density
10 Å
H50Q
α-synuclein, H50Q, recombinant
PDB ID 6pes
α-synuclein, H50Q, recombinant
PDB ID 6peo
A53T
α-synuclein, A53T, N-acetylated, recombinant
PDB ID 6lrq
T22TMAAAEKT
α-synuclein, T22TMAAAEKT, juvenile-onset synucleinopathy patient brain
PDB ID 6bev
α-synuclein, T22TMAAAEKT, juvenile-onset synucleinopathy patient brain
PDB ID 6bew

Figure 4-7: Fiber structures of α-synuclein. Structures of amyloid fiber cores of wild-type, E46K, G51D, H50Q, A53T, and T22TMAAAEKT α-synuclein. Structures are grouped in boxes based on which mutation they have and wild-type structures are in the centermost box. In the wild-type structures, residues which are highlighted with a red outline correspond to the mutated residues in the mutant structures. In the mutant structures, the residue highlighted with a red outline corresponds to the mutant residue in that structure.

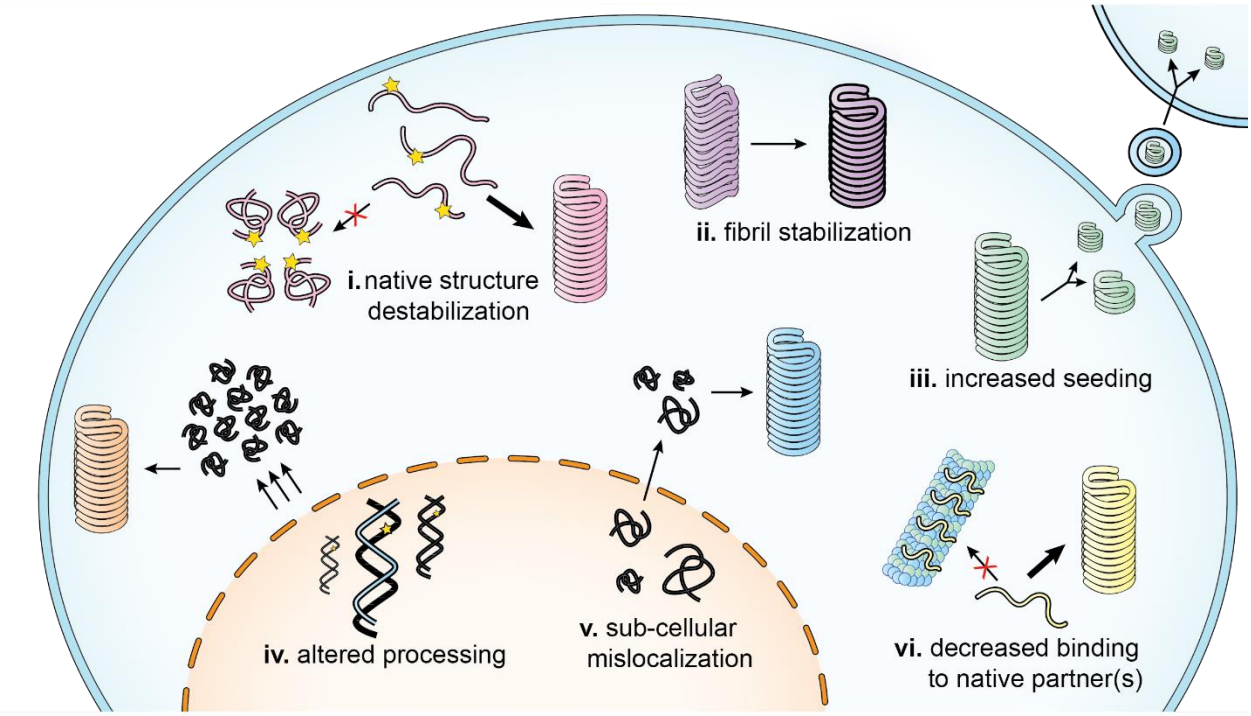


Figure 4-8: Biochemical mechanisms of amyloidogenic mutations. An illustration representing all six classes of mechanisms of amyloidogenic mutations. These mechanisms are i) native structure destabilization, ii) fibril stabilization, iii) increased seeding, iv) altered processing, v) subcellular mislocalization, and vi) decreased binding to native partner(s).

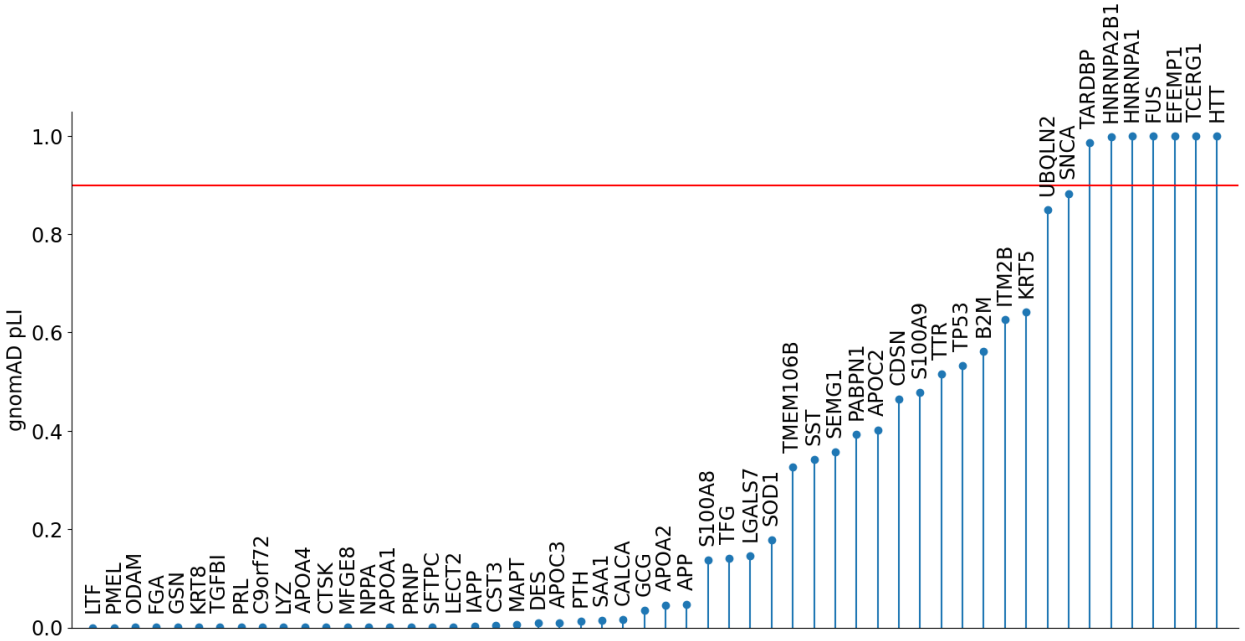


Figure 4-9: Most amyloid protein-coding genes are haplosufficient. A stemplot showing the probability of loss-of-function intolerance (pLI) from the genome aggregation database (gnomAD) for each of the amyloid proteins in our list (excluding the immunoglobulins and iatrogenic amyloids). The y-axis is the gnomAD pLI and the threshold for calling a gene haploinsufficient is a pLI of 0.9, represented by the horizontal red line. Each point is labeled with the gene name of the corresponding amyloid protein. 7 out of 54 amyloid protein-coding genes are haploinsufficient: TARDBP, HNRNPA2B1, HNRNPA1, FUS, EFEMP1, TCERG1, and HTT.

Table S4-1 Pathogenic amyloid proteins. All amyloid proteins which were considered in this review are listed in this table along with other relevant information including which amyloid diseases they cause, their amyloidogenic mutations, and the biochemical effects of their mutations (based on the numbering system used in the text and Figure 8)

References

1. Kyle, R. A. (2001) Amyloidosis: a convoluted story. *Br J Haematol*. **114**, 529–538
2. Ohyama, T., Shimokama, T., Yoshikawa, Y., and Watanabe, T. (1990) Splenic amyloidosis: correlations between chemical types of amyloid protein and morphological features. *Mod Pathol*. **3**, 419–422
3. Buck, F. S., and Koss, M. N. (1991) Hepatic amyloidosis: morphologic differences between systemic AL and AA types. *Hum Pathol*. **22**, 904–907
4. LeVine, H. (1999) Quantification of beta-sheet amyloid fibril structures with thioflavin T. *Methods Enzymol*. **309**, 274–284
5. Nilsson, M. R. (2004) Techniques to study amyloid fibril formation in vitro. *Methods*. **34**, 151–160
6. Astbury, W. T., Dickinson, S., and Bailey, K. (1935) The X-ray interpretation of denaturation and the structure of the seed globulins. *Biochem J*. **29**, 2351-2360.1
7. Sunde, M., Serpell, L. C., Bartlam, M., Fraser, P. E., Pepys, M. B., and Blake, C. C. (1997) Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J Mol Biol*. **273**, 729–739
8. Eisenberg, D., and Jucker, M. (2012) The amyloid state of proteins in human diseases. *Cell*. **148**, 1188–1203
9. Sawaya, M. R., Hughes, M. P., Rodriguez, J. A., Riek, R., and Eisenberg, D. S. (2021) The expanding amyloid family: Structure, stability, function, and pathogenesis. *Cell*. **184**, 4857–4873
10. Fitzpatrick, A. W. P., Falcon, B., He, S., Murzin, A. G., Murshudov, G., Garringer, H. J., Crowther, R. A., Ghetti, B., Goedert, M., and Scheres, S. H. W. (2017) Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature*. **547**, 185–190
11. Buxbaum, J. N., Dispenzieri, A., Eisenberg, D. S., Fändrich, M., Merlini, G., Saraiva, M. J. M., Sekijima, Y., and Westermark, P. (2022) Amyloid nomenclature 2022: update, novel

proteins, and recommendations by the International Society of Amyloidosis (ISA) Nomenclature Committee. *Amyloid*. **29**, 213–219

12. Andrade, C. (1952) A peculiar form of peripheral neuropathy; familiar atypical generalized amyloidosis with special involvement of the peripheral nerves. *Brain*. **75**, 408–427
13. Fändrich, M., Fletcher, M. A., and Dobson, C. M. (2001) Amyloid fibrils from muscle myoglobin. *Nature*. **410**, 165–166
14. Guijarro, J. I., Sunde, M., Jones, J. A., Campbell, I. D., and Dobson, C. M. (1998) Amyloid fibril formation by an SH3 domain. *Proc Natl Acad Sci U S A*. **95**, 4224–4228
15. Röder, C., Vettore, N., Mangels, L. N., Gremer, L., Ravelli, R. B. G., Willbold, D., Hoyer, W., Buell, A. K., and Schröder, G. F. (2019) Atomic structure of PI3-kinase SH3 amyloid fibrils by cryo-electron microscopy. *Nat Commun*. **10**, 3754
16. Rubel, M. S., Fedotov, S. A., Grizel, A. V., Sopova, J. V., Malikova, O. A., Chernoff, Y. O., and Rubel, A. A. (2020) Functional Mammalian Amyloids and Amyloid-Like Proteins. *Life (Basel)*. **10**, 156
17. Matiiv, A. B., Trubitsina, N. P., Matveenko, A. G., Barbitoff, Y. A., Zhouravleva, G. A., and Bondarev, S. A. (2020) Amyloid and Amyloid-Like Aggregates: Diversity and the Term Crisis. *Biochemistry (Mosc)*. **85**, 1011–1034
18. Brown, A., and Török, M. (2021) Functional amyloids in the human body. *Bioorg Med Chem Lett*. **40**, 127914
19. Brumshtein, B., Esswein, S. R., Sawaya, M. R., Rosenberg, G., Ly, A. T., Landau, M., and Eisenberg, D. S. (2018) Identification of two principal amyloid-driving segments in variable domains of Ig light chains in systemic light-chain amyloidosis. *J Biol Chem*. **293**, 19659–19671
20. Mori, K., Arzberger, T., Grässer, F. A., Gijselinck, I., May, S., Rentzsch, K., Weng, S.-M., Schludi, M. H., van der Zee, J., Cruts, M., Van Broeckhoven, C., Kremmer, E.,

Kretzschmar, H. A., Haass, C., and Edbauer, D. (2013) Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathol*. **126**, 881–893

21. Flores, B. N., Dulchavsky, M. E., Krans, A., Sawaya, M. R., Paulson, H. L., Todd, P. K., Barmada, S. J., and Ivanova, M. I. (2016) Distinct C9orf72-Associated Dipeptide Repeat Structures Correlate with Neuronal Toxicity. *PLoS One*. **11**, e0165084
22. Chang, Y.-J., Jeng, U.-S., Chiang, Y.-L., Hwang, I.-S., and Chen, Y.-R. (2016) The Glycine-Alanine Dipeptide Repeat from C9orf72 Hexanucleotide Expansions Forms Toxic Amyloids Possessing Cell-to-Cell Transmission Properties. *J Biol Chem*. **291**, 4903–4911
23. Brasseur, L., Coens, A., Waeytens, J., Melki, R., and Bousset, L. (2020) Dipeptide repeat derived from C9orf72 hexanucleotide expansions forms amyloids or natively unfolded structures in vitro. *Biochem Biophys Res Commun*. **526**, 410–416
24. Dische, F. E., Wernstedt, C., Westermark, G. T., Westermark, P., Pepys, M. B., Rennie, J. A., Gilbey, S. G., and Watkins, P. J. (1988) Insulin as an amyloid-fibril protein at sites of repeated insulin injections in a diabetic patient. *Diabetologia*. **31**, 158–161
25. Nagase, T., Katsura, Y., Iwaki, Y., Nemoto, K., Sekine, H., Miwa, K., Oh-I, T., Kou, K., Iwaya, K., Noritake, M., and Matsuoka, T. (2009) The insulin ball. *Lancet*. **373**, 184
26. Martins, C. O., Lezcano, C., Yi, S. S., Landau, H. J., Chapman, J. R., and Dogan, A. (2018) Novel iatrogenic amyloidosis caused by peptide drug liraglutide: a clinical mimic of AL amyloidosis. *Haematologica*. **103**, e610–e612
27. Alehashemi, S., Dasari, S., de Jesus, A. A., Cowen, E. W., Lee, C.-C. R., Goldbach-Mansky, R., and McPhail, E. D. (2022) Anakinra-Associated Amyloidosis. *JAMA Dermatol*. **158**, 1454–1457
28. D'Souza, A., Theis, J. D., Vrana, J. A., and Dogan, A. (2014) Pharmaceutical amyloidosis associated with subcutaneous insulin and enfuvirtide administration. *Amyloid*. **21**, 71–75

29. Naujokas, A., Vidal, C. I., Mercer, S. E., Harp, J., Kurtin, P. J., Fox, L. P., and Thompson, M. M. (2012) A novel form of amyloid deposited at the site of enfuvirtide injection. *J Cutan Pathol*. **39**, 220–221; quiz 219

30. Morilla, M. E., Kocher, J., and Harmaty, M. (2009) Localized amyloidosis at the site of enfuvirtide injection. *Ann Intern Med*. **151**, 515–516

31. Gejyo, F., Yamada, T., Odani, S., Nakagawa, Y., Arakawa, M., Kunitomo, T., Kataoka, H., Suzuki, M., Hirasawa, Y., and Shirahama, T. (1985) A new form of amyloid protein associated with chronic hemodialysis was identified as beta 2-microglobulin. *Biochem Biophys Res Commun*. **129**, 701–706

32. Valleix, S., Gillmore, J. D., Bridoux, F., Mangione, P. P., Dogan, A., Nedelec, B., Boimard, M., Touchard, G., Goujon, J.-M., Lacombe, C., Lozeron, P., Adams, D., Lacroix, C., Maisonobe, T., Planté-Bordeneuve, V., Vrana, J. A., Theis, J. D., Giorgetti, S., Porcari, R., Ricagno, S., Bolognesi, M., Stoppini, M., Delpech, M., Pepys, M. B., Hawkins, P. N., and Bellotti, V. (2012) Hereditary systemic amyloidosis due to Asp76Asn variant β2-microglobulin. *N Engl J Med*. **366**, 2276–2283

33. Nilsberth, C., Westlind-Danielsson, A., Eckman, C. B., Condron, M. M., Axelman, K., Forsell, C., Stenh, C., Luthman, J., Teplow, D. B., Younkin, S. G., Näslund, J., and Lannfelt, L. (2001) The “Arctic” APP mutation (E693G) causes Alzheimer’s disease by enhanced Abeta protofibril formation. *Nat Neurosci*. **4**, 887–893

34. Masters, C. L., Simms, G., Weinman, N. A., Multhaup, G., McDonald, B. L., and Beyreuther, K. (1985) Amyloid plaque core protein in Alzheimer disease and Down syndrome. *Proc Natl Acad Sci U S A*. **82**, 4245–4249

35. Melchor, J. P., McVoy, L., and Van Nostrand, W. E. (2000) Charge alterations of E22 enhance the pathogenic properties of the amyloid beta-protein. *J Neurochem*. **74**, 2209–2212

36. Van Nostrand, W. E., Melchor, J. P., Cho, H. S., Greenberg, S. M., and Rebeck, G. W. (2001) Pathogenic effects of D23N Iowa mutant amyloid beta -protein. *J Biol Chem*. **276**, 32860–32866

37. Qiang, W., Yau, W.-M., Luo, Y., Mattson, M. P., and Tycko, R. (2012) Antiparallel β-sheet architecture in Iowa-mutant β-amyloid fibrils. *Proc Natl Acad Sci U S A*. **109**, 4443–4448

38. Sgourakis, N. G., Yau, W.-M., and Qiang, W. (2015) Modeling an in-register, parallel "iowa" aβ fibril structure using solid-state NMR data from labeled samples with rosetta. *Structure*. **23**, 216–227

39. Schütz, A. K., Vagt, T., Huber, M., Ovchinnikova, O. Y., Cadalbert, R., Wall, J., Güntert, P., Böckmann, A., Glockshuber, R., and Meier, B. H. (2015) Atomic-resolution three-dimensional structure of amyloid β fibrils bearing the Osaka mutation. *Angew Chem Int Ed Engl*. **54**, 331–335

40. Yang, Y., Zhang, W., Murzin, A. G., Schweighauser, M., Huang, M., Lövestam, S., Peak-Chew, S. Y., Saito, T., Saido, T. C., Macdonald, J., Lavenir, I., Ghetti, B., Graff, C., Kumar, A., Nordberg, A., Goedert, M., and Scheres, S. H. W. (2023) Cryo-EM structures of amyloid-β filaments with the Arctic mutation (E22G) from human and mouse brains. *Acta Neuropathol*. **145**, 325–333

41. Yang, Y., Arseni, D., Zhang, W., Huang, M., Lövestam, S., Schweighauser, M., Kotecha, A., Murzin, A. G., Peak-Chew, S. Y., Macdonald, J., Lavenir, I., Garringer, H. J., Gelpi, E., Newell, K. L., Kovacs, G. G., Vidal, R., Ghetti, B., Ryskeldi-Falcon, B., Scheres, S. H. W., and Goedert, M. (2022) Cryo-EM structures of amyloid-β 42 filaments from human brains. *Science*. **375**, 167–172

42. Stern, A. M., Yang, Y., Jin, S., Yamashita, K., Meunier, A. L., Liu, W., Cai, Y., Ericsson, M., Liu, L., Goedert, M., Scheres, S. H. W., and Selkoe, D. J. (2023) Abundant Aβ fibrils in ultracentrifugal supernatants of aqueous extracts from Alzheimer's disease brains.

Neuron. 10.1016/j.neuron.2023.04.007

43. Kollmer, M., Close, W., Funk, L., Rasmussen, J., Bsoul, A., Schierhorn, A., Schmidt, M., Sigurdson, C. J., Jucker, M., and Fändrich, M. (2019) Cryo-EM structure and polymorphism of Aβ amyloid fibrils purified from Alzheimer's brain tissue. *Nat Commun*. **10**, 4760

44. Kato, M., Han, T. W., Xie, S., Shi, K., Du, X., Wu, L. C., Mirzaei, H., Goldsmith, E. J., Longgood, J., Pei, J., Grishin, N. V., Frantz, D. E., Schneider, J. W., Chen, S., Li, L., Sawaya, M. R., Eisenberg, D., Tycko, R., and McKnight, S. L. (2012) Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell*. **149**, 753–767

45. Lu, J., Cao, Q., Hughes, M. P., Sawaya, M. R., Boyer, D. R., Cascio, D., and Eisenberg, D. S. (2020) CryoEM structure of the low-complexity domain of hnRNPA2 and its conversion to pathogenic amyloid. *Nat Commun*. **11**, 4090

46. Kim, H. J., Kim, N. C., Wang, Y.-D., Scarborough, E. A., Moore, J., Diaz, Z., MacLea, K. S., Freibaum, B., Li, S., Molliex, A., Kanagaraj, A. P., Carter, R., Boylan, K. B., Wojtas, A. M., Rademakers, R., Pinkus, J. L., Greenberg, S. A., Trojanowski, J. Q., Traynor, B. J., Smith, B. N., Topp, S., Gkazi, A.-S., Miller, J., Shaw, C. E., Kottlors, M., Kirschner, J., Pestronk, A., Li, Y. R., Ford, A. F., Gitler, A. D., Benatar, M., King, O. D., Kimonis, V. E., Ross, E. D., Weihl, C. C., Shorter, J., and Taylor, J. P. (2013) Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature*. **495**, 467–473

47. Westermark, P., Andersson, A., and Westermark, G. T. (2011) Islet amyloid polypeptide, islet amyloid, and diabetes mellitus. *Physiol Rev*. **91**, 795–826

48. Westermark, P., Wernstedt, C., Wilander, E., and Sletten, K. (1986) A novel peptide in the calcitonin gene related peptide family as an amyloid fibril protein in the endocrine

pancreas. *Biochem Biophys Res Commun*. **140**, 827–831

49. Westermark, P., Wernstedt, C., Wilander, E., Hayden, D. W., O'Brien, T. D., and Johnson, K. H. (1987) Amyloid fibrils in human insulinoma and islets of Langerhans of the diabetic cat are derived from a neuropeptide-like protein also present in normal islet cells. *Proc Natl Acad Sci U S A*. **84**, 3881–3885
50. Sakagashira, S., Sanke, T., Hanabusa, T., Shimomura, H., Ohagi, S., Kumagaye, K. Y., Nakajima, K., and Nanjo, K. (1996) Missense mutation of amylin gene (S20G) in Japanese NIDDM patients. *Diabetes*. **45**, 1279–1281
51. Ma, Z., Westermark, G. T., Sakagashira, S., Sanke, T., Gustavsson, A., Sakamoto, H., Engström, U., Nanjo, K., and Westermark, P. (2001) Enhanced in vitro production of amyloid-like fibrils from mutant (S20G) islet amyloid polypeptide. *Amyloid*. **8**, 242–249
52. Sakagashira, S., Hiddinga, H. J., Tateishi, K., Sanke, T., Hanabusa, T., Nanjo, K., and Eberhardt, N. L. (2000) S20G mutant amylin exhibits increased in vitro amyloidogenicity and increased intracellular cytotoxicity compared to wild-type amylin. *Am J Pathol*. **157**, 2101–2109
53. Meier, D. T., Entrup, L., Templin, A. T., Hogan, M. F., Mellati, M., Zraika, S., Hull, R. L., and Kahn, S. E. (2016) The S20G substitution in hIAPP is more amyloidogenic and cytotoxic than wild-type hIAPP in mouse islets. *Diabetologia*. **59**, 2166–2171
54. Xu, W., Jiang, P., and Mu, Y. (2009) Conformation preorganization: effects of S20G mutation on the structure of human islet amyloid polypeptide segment. *J Phys Chem B*. **113**, 7308–7314
55. Gallardo, R., Iadanza, M. G., Xu, Y., Heath, G. R., Foster, R., Radford, S. E., and Ranson, N. A. (2020) Fibril structures of diabetes-related amylin variants reveal a basis for surface-templated assembly. *Nat Struct Mol Biol*. **27**, 1048–1056
56. Pan, K. M., Baldwin, M., Nguyen, J., Gasset, M., Serban, A., Groth, D., Mehlhorn, I.,

Huang, Z., Fletterick, R. J., and Cohen, F. E. (1993) Conversion of alpha-helices into beta-sheets features in the formation of the scrapie prion proteins. *Proc Natl Acad Sci U S A*. **90**, 10962–10966

57. Prusiner, S. B. (1998) Prions. *Proc Natl Acad Sci U S A*. **95**, 13363–13383

58. Mead, S., Uphill, J., Beck, J., Poulter, M., Campbell, T., Lowe, J., Adamson, G., Hummerich, H., Klopp, N., Rückert, I.-M., Wichmann, H.-E., Azazi, D., Plagnol, V., Pako, W. H., Whitfield, J., Alpers, M. P., Whittaker, J., Balding, D. J., Zerr, I., Kretzschmar, H., and Collinge, J. (2012) Genome-wide association study in multiple human prion diseases suggests genetic risk factors additional to PRNP. *Hum Mol Genet*. **21**, 1897–1906

59. Palmer, M. S., Dryden, A. J., Hughes, J. T., and Collinge, J. (1991) Homozygous prion protein genotype predisposes to sporadic Creutzfeldt-Jakob disease. *Nature*. **352**, 340–342

60. Minikel, E. V., Vallabh, S. M., Lek, M., Estrada, K., Samocha, K. E., Sathirapongsasuti, J. F., McLean, C. Y., Tung, J. Y., Yu, L. P. C., Gambetti, P., Blevins, J., Zhang, S., Cohen, Y., Chen, W., Yamada, M., Hamaguchi, T., Sanjo, N., Mizusawa, H., Nakamura, Y., Kitamoto, T., Collins, S. J., Boyd, A., Will, R. G., Knight, R., Ponto, C., Zerr, I., Kraus, T. F. J., Eigenbrod, S., Giese, A., Calero, M., de Pedro-Cuesta, J., Haïk, S., Laplanche, J.-L., Bouaziz-Amar, E., Brandel, J.-P., Capellari, S., Parchi, P., Poleggi, A., Ladogana, A., O'Donnell-Luria, A. H., Karczewski, K. J., Marshall, J. L., Boehnke, M., Laakso, M., Mohlke, K. L., Kähler, A., Chambert, K., McCarroll, S., Sullivan, P. F., Hultman, C. M., Purcell, S. M., Sklar, P., van der Lee, S. J., Rozemuller, A., Jansen, C., Hofman, A., Kraaij, R., van Rooij, J. G. J., Ikram, M. A., Uitterlinden, A. G., van Duijn, C. M., Exome Aggregation Consortium (ExAC), Daly, M. J., and MacArthur, D. G. (2016) Quantifying prion disease penetrance using large population control cohorts. *Sci Transl Med*. **8**, 322ra9

61. Wang, L.-Q., Zhao, K., Yuan, H.-Y., Li, X.-N., Dang, H.-B., Ma, Y., Wang, Q., Wang, C., Sun, Y., Chen, J., Li, D., Zhang, D., Yin, P., Liu, C., and Liang, Y. Genetic prion disease–related mutation E196K displays a novel amyloid fibril structure revealed by cryo-EM. *Sci Adv*. **7**, eabg9676

62. Li, Q., Jaroniec, C. P., and Surewicz, W. K. (2022) Cryo-EM structure of disease-related prion fibrils provides insights into seeding barriers. *Nat Struct Mol Biol*. **29**, 962–965

63. Hallinan, G. I., Ozcan, K. A., Hoq, M. R., Cracco, L., Vago, F. S., Bharath, S. R., Li, D., Jacobsen, M., Doud, E. H., Mosley, A. L., Fernandez, A., Garringer, H. J., Jiang, W., Ghetti, B., and Vidal, R. (2022) Cryo-EM structures of prion protein filaments from Gerstmann-Sträussler-Scheinker disease. *Acta Neuropathol*. **144**, 509–520

64. Glynn, C., Sawaya, M. R., Ge, P., Gallagher-Jones, M., Short, C. W., Bowman, R., Apostol, M., Zhou, Z. H., Eisenberg, D., and Rodriguez, J. A. (2020) Cryo-EM structure of a human prion fibril with a hydrophobic, protease-resistant core. *Nat Struct Mol Biol*. **27**, 417–423

65. Wang, L.-Q., Zhao, K., Yuan, H.-Y., Wang, Q., Guan, Z., Tao, J., Li, X.-N., Sun, Y., Yi, C.-W., Chen, J., Li, D., Zhang, D., Yin, P., Liu, C., and Liang, Y. (2020) Cryo-EM structure of an amyloid fibril formed by full-length human prion protein. *Nat Struct Mol Biol*. **27**, 598–602

66. Tagliavini, F., Prelli, F., Ghiso, J., Bugiani, O., Serban, D., Prusiner, S. B., Farlow, M. R., Ghetti, B., and Frangione, B. (1991) Amyloid protein of Gerstmann-Sträussler-Scheinker disease (Indiana kindred) is an 11 kd fragment of prion protein with an N-terminal glycine at codon 58. *EMBO J*. **10**, 513–519

67. Hadži, S., Ondračka, A., Jerala, R., and Hafner-Bratkovič, I. (2015) Pathological mutations H187R and E196K facilitate subdomain separation and prion protein conversion by destabilization of the native structure. *FASEB J*. **29**, 882–893

68. Lagier-Tourenne, C., Polymenidou, M., and Cleveland, D. W. (2010) TDP-43 and FUS/TLS: emerging roles in RNA processing and neurodegeneration. *Hum Mol Genet*. **19**, R46-64

69. Colombrita, C., Zennaro, E., Fallini, C., Weber, M., Sommacal, A., Buratti, E., Silani, V., and Ratti, A. (2009) TDP-43 is recruited to stress granules in conditions of oxidative insult. *J Neurochem*. **111**, 1051–1061

70. Moisse, K., Volkening, K., Leystra-Lantz, C., Welch, I., Hill, T., and Strong, M. J. (2009) Divergent patterns of cytosolic TDP-43 and neuronal progranulin expression following axotomy: implications for TDP-43 in the physiological response to neuronal injury. *Brain Res*. **1249**, 202–211

71. Nishimoto, Y., Ito, D., Yagi, T., Nihei, Y., Tsunoda, Y., and Suzuki, N. (2010) Characterization of alternative isoforms and inclusion body of the TAR DNA-binding protein-43. *J Biol Chem*. **285**, 608–619

72. Cao, Q., Boyer, D. R., Sawaya, M. R., Ge, P., and Eisenberg, D. S. (2019) Cryo-EM structures of four polymorphic TDP-43 amyloid cores. *Nat Struct Mol Biol*. **26**, 619–627

73. Conicella, A. E., Zerze, G. H., Mittal, J., and Fawzi, N. L. (2016) ALS Mutations Disrupt Phase Separation Mediated by α-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. *Structure*. **24**, 1537–1549

74. Johnson, B. S., Snead, D., Lee, J. J., McCaffery, J. M., Shorter, J., and Gitler, A. D. (2009) TDP-43 is intrinsically aggregation-prone, and amyotrophic lateral sclerosis-linked mutations accelerate aggregation and increase toxicity. *J Biol Chem*. **284**, 20329–20339

75. Jiang, L.-L., Zhao, J., Yin, X.-F., He, W.-T., Yang, H., Che, M.-X., and Hu, H.-Y. (2016) Two mutations G335D and Q343R within the amyloidogenic core region of TDP-43 influence its aggregation and inclusion formation. *Sci Rep*. **6**, 23928

76. Neumann, M., Sampathu, D. M., Kwong, L. K., Truax, A. C., Micsenyi, M. C., Chou, T. T.,

Bruce, J., Schuck, T., Grossman, M., Clark, C. M., McCluskey, L. F., Miller, B. L., Masliah, E., Mackenzie, I. R., Feldman, H., Feiden, W., Kretzschmar, H. A., Trojanowski, J. Q., and Lee, V. M.-Y. (2006) Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science*. **314**, 130–133

77. Kwong, L. K., Uryu, K., Trojanowski, J. Q., and Lee, V. M.-Y. (2008) TDP-43 proteinopathies: neurodegenerative protein misfolding diseases without amyloidosis. *Neurosignals*. **16**, 41–51

78. Jiang, L.-L., Che, M.-X., Zhao, J., Zhou, C.-J., Xie, M.-Y., Li, H.-Y., He, J.-H., and Hu, H.-Y. (2013) Structural transformation of the amyloidogenic core region of TDP-43 protein initiates its aggregation and cytoplasmic inclusion. *J Biol Chem*. **288**, 19614–19624

79. Li, Q., Babinchak, W. M., and Surewicz, W. K. (2021) Cryo-EM structure of amyloid fibrils formed by the entire low complexity domain of TDP-43. *Nat Commun*. **12**, 1620

80. Arseni, D., Hasegawa, M., Murzin, A. G., Kametani, F., Arai, M., Yoshida, M., and Ryskeldi-Falcon, B. (2022) Structure of pathological TDP-43 filaments from ALS with FTLD. *Nature*. **601**, 139–143

81. Richardson, S. J. (2007) Cell and molecular biology of transthyretin and thyroid hormones. *Int Rev Cytol*. **258**, 137–193

82. Adams, D., Koike, H., Slama, M., and Coelho, T. (2019) Hereditary transthyretin amyloidosis: a model of medical progress for a fatal disease. *Nat Rev Neurol*. **15**, 387–404

83. Ruberg, F. L., Grogan, M., Hanna, M., Kelly, J. W., and Maurer, M. S. (2019) Transthyretin Amyloid Cardiomyopathy: JACC State-of-the-Art Review. *J Am Coll Cardiol*. **73**, 2872–2891

84. Planté-Bordeneuve, V., and Said, G. (2011) Familial amyloid polyneuropathy. *Lancet Neurol*. **10**, 1086–1097

85. Schmidt, M., Wiese, S., Adak, V., Engler, J., Agarwal, S., Fritz, G., Westermark, P., Zacharias, M., and Fändrich, M. (2019) Cryo-EM structure of a transthyretin-derived amyloid fibril from a patient with hereditary ATTR amyloidosis. *Nat Commun*. **10**, 5008

86. Iakovleva, I., Hall, M., Oelker, M., Sandblad, L., Anan, I., and Sauer-Eriksson, A. E. (2021) Structural basis for transthyretin amyloid formation in vitreous body of the eye. *Nat Commun*. **12**, 7141

87. Steinebrei, M., Gottwald, J., Baur, J., Röcken, C., Hegenbart, U., Schönland, S., and Schmidt, M. (2022) Cryo-EM structure of an ATTRwt amyloid fibril from systemic non-hereditary transthyretin amyloidosis. *Nat Commun*. **13**, 6398

88. Sekijima, Y., Wiseman, R. L., Matteson, J., Hammarström, P., Miller, S. R., Sawkar, A. R., Balch, W. E., and Kelly, J. W. (2005) The biological and chemical basis for tissue-selective amyloid disease. *Cell*. **121**, 73–85

89. Burré, J. (2015) The Synaptic Function of α-Synuclein. *J Parkinsons Dis*. **5**, 699–713

90. Spillantini, M. G., Schmidt, M. L., Lee, V. M., Trojanowski, J. Q., Jakes, R., and Goedert, M. (1997) Alpha-synuclein in Lewy bodies. *Nature*. **388**, 839–840

91. Boyer, D. R., Li, B., Sun, C., Fan, W., Sawaya, M. R., Jiang, L., and Eisenberg, D. S. (2019) Structures of fibrils formed by α-synuclein hereditary disease mutant H50Q reveal new polymorphs. *Nat Struct Mol Biol*. **26**, 1044–1052

92. Sun, Y., Hou, S., Zhao, K., Long, H., Liu, Z., Gao, J., Zhang, Y., Su, X.-D., Li, D., and Liu, C. (2020) Cryo-EM structure of full-length α-synuclein amyloid fibril with Parkinson's disease familial A53T mutation. *Cell Res*. **30**, 360–362

93. Boyer, D. R., Li, B., Sun, C., Fan, W., Zhou, K., Hughes, M. P., Sawaya, M. R., Jiang, L., and Eisenberg, D. S. (2020) The α-synuclein hereditary mutation E46K unlocks a more stable, pathogenic fibril structure. *Proc Natl Acad Sci U S A*. **117**, 3592–3602

94. Zhao, K., Li, Y., Liu, Z., Long, H., Zhao, C., Luo, F., Sun, Y., Tao, Y., Su, X.-D., Li, D., Li,

X., and Liu, C. (2020) Parkinson's disease associated mutation E46K of α-synuclein triggers the formation of a distinct fibril structure. *Nat Commun*. **11**, 2643

95. Sun, Y., Long, H., Xia, W., Wang, K., Zhang, X., Sun, B., Cao, Q., Zhang, Y., Dai, B., Li, D., and Liu, C. (2021) The hereditary mutation G51D unlocks a distinct fibril strain transmissible to wild-type α-synuclein. *Nat Commun*. **12**, 6252

96. Yang, Y., Garringer, H. J., Shi, Y., Lövestam, S., Peak-Chew, S., Zhang, X., Kotecha, A., Bacioglu, M., Koto, A., Takao, M., Spillantini, M. G., Ghetti, B., Vidal, R., Murzin, A. G., Scheres, S. H. W., and Goedert, M. (2023) New SNCA mutation and structures of α-synuclein filaments from juvenile-onset synucleinopathy. *Acta Neuropathol*. **145**, 561–572

97. Schweighauser, M., Shi, Y., Tarutani, A., Kametani, F., Murzin, A. G., Ghetti, B., Matsubara, T., Tomita, T., Ando, T., Hasegawa, K., Murayama, S., Yoshida, M., Hasegawa, M., Scheres, S. H. W., and Goedert, M. (2020) Structures of α-Synuclein Filaments from Multiple System Atrophy. *Nature*. **585**, 464–469

98. Yang, Y., Shi, Y., Schweighauser, M., Zhang, X., Kotecha, A., Murzin, A. G., Garringer, H. J., Cullinane, P. W., Saito, Y., Foroud, T., Warner, T. T., Hasegawa, K., Vidal, R., Murayama, S., Revesz, T., Ghetti, B., Hasegawa, M., Lashley, T., Scheres, S. H. W., and Goedert, M. (2022) Structures of α-synuclein filaments from human brains with Lewy pathology. *Nature*. **610**, 791–795

99. Li, B., Ge, P., Murray, K. A., Sheth, P., Zhang, M., Nair, G., Sawaya, M. R., Shin, W. S., Boyer, D. R., Ye, S., Eisenberg, D. S., Zhou, Z. H., and Jiang, L. (2018) Cryo-EM of full-length α-synuclein reveals fibril polymorphs with a common structural kernel. *Nat Commun*. **9**, 3609

100. Porcari, R., Proukakis, C., Waudby, C. A., Bolognesi, B., Mangione, P. P., Paton, J. F. S., Mullin, S., Cabrita, L. D., Penco, A., Relini, A., Verona, G., Vendruscolo, M., Stoppini, M., Tartaglia, G. G., Camilloni, C., Christodoulou, J., Schapira, A. H. V., and Bellotti, V.

(2015) The H50Q mutation induces a 10-fold decrease in the solubility of α-synuclein. *J Biol Chem*. **290**, 2395–2404

101. Narhi, L., Wood, S. J., Steavenson, S., Jiang, Y., Wu, G. M., Anafi, D., Kaufman, S. A., Martin, F., Sitney, K., Denis, P., Louis, J. C., Wypych, J., Biere, A. L., and Citron, M. (1999) Both familial Parkinson's disease mutations accelerate alpha-synuclein aggregation. *J Biol Chem*. **274**, 9843–9846

102. Conway, K. A., Harper, J. D., and Lansbury, P. T. (1998) Accelerated in vitro fibril formation by a mutant alpha-synuclein linked to early-onset Parkinson disease. *Nat Med*. **4**, 1318–1320

103. Takao, M., Ghetti, B., Yoshida, H., Piccardo, P., Narain, Y., Murrell, J. R., Vidal, R., Glazier, B. S., Jakes, R., Tsutsui, M., Spillantini, M. G., Crowther, R. A., Goedert, M., and Koto, A. (2004) Early-onset dementia with Lewy bodies. *Brain Pathol*. **14**, 137–147

104. Campion, D., Dumanchin, C., Hannequin, D., Dubois, B., Belliard, S., Puel, M., Thomas-Anterion, C., Michon, A., Martin, C., Charbonnier, F., Raux, G., Camuzat, A., Penet, C., Mesnage, V., Martinez, M., Clerget-Darpoux, F., Brice, A., and Frebourg, T. (1999) Early-onset autosomal dominant Alzheimer disease: prevalence, genetic heterogeneity, and mutation spectrum. *Am J Hum Genet*. **65**, 664–670

105. De Strooper, B., Iwatsubo, T., and Wolfe, M. S. (2012) Presenilins and γ-Secretase: Structure, Function, and Role in Alzheimer Disease. *Cold Spring Harb Perspect Med*. **2**, a006304

106. Selkoe, D. J. (1999) Translating cell biology into therapeutic advances in Alzheimer's disease. *Nature*. **399**, A23-31

107. Bentahir, M., Nyabi, O., Verhamme, J., Tolia, A., Horré, K., Wiltfang, J., Esselmann, H., and De Strooper, B. (2006) Presenilin clinical mutations can affect gamma-secretase activity by different mechanisms. *J Neurochem*. **96**, 732–742

108. Arciello, A., Piccoli, R., and Monti, D. M. (2016) Apolipoprotein A-I: the dual face of a protein. *FEBS Lett*. **590**, 4171–4179

109. Das, M., Wilson, C. J., Mei, X., Wales, T. E., Engen, J. R., and Gursky, O. (2016) Structural Stability and Local Dynamics in Disease-Causing Mutants of Human Apolipoprotein A-I: What Makes the Protein Amyloidogenic? *J Mol Biol*. **428**, 449–462

110. Zhang, W., Falcon, B., Murzin, A. G., Fan, J., Crowther, R. A., Goedert, M., and Scheres, S. H. (2019) Heparin-induced tau filaments are polymorphic and differ from those in Alzheimer's and Pick's diseases. *Elife*. **8**, e43584

111. Nasr, S. H., Dasari, S., Mills, J. R., Theis, J. D., Zimmermann, M. T., Fonseca, R., Vrana, J. A., Lester, S. J., McLaughlin, B. M., Gillespie, R., Highsmith, W. E., Lee, J. J., Dispenzieri, A., and Kurtin, P. J. (2017) Hereditary Lysozyme Amyloidosis Variant p.Leu102Ser Associates with Unique Phenotype. *J Am Soc Nephrol*. **28**, 431–438

112. Vidal, R., Revesz, T., Rostagno, A., Kim, E., Holton, J. L., Bek, T., Bojsen-Møller, M., Braendgaard, H., Plant, G., Ghiso, J., and Frangione, B. (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

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

114. De Gracia, R., Fernández, E. J., Riñón, C., Selgas, R., and Garcia-Bustos, J. (2006) Hereditary renal amyloidosis associated with a novel mutation in the apolipoprotein AII gene. *QJM*. **99**, 274

115. Yazaki, M., Liepnieks, J. J., Barats, M. S., Cohen, A. H., and Benson, M. D. (2003) Hereditary systemic amyloidosis associated with a new apolipoprotein AII stop codon mutation Stop78Arg. *Kidney Int*. **64**, 11–16

116. Benson, M. D., Liepnieks, J. J., Yazaki, M., Yamashita, T., Hamidi Asl, K., Guenther, B., and Kluve-Beckerman, B. (2001) A new human hereditary amyloidosis: the result of a stop-codon mutation in the apolipoprotein AII gene. *Genomics*. **72**, 272–277

117. Yazaki, M., Liepnieks, J. J., Yamashita, T., Guenther, B., Skinner, M., and Benson, M. D. (2001) Renal amyloidosis caused by a novel stop-codon mutation in the apolipoprotein A-II gene. *Kidney Int*. **60**, 1658–1665

118. Prokaeva, T., Akar, H., Spencer, B., Havasi, A., Cui, H., O'Hara, C. J., Gursky, O., Leszyk, J., Steffen, M., Browning, S., Rosenberg, A., and Connors, L. H. (2017) Hereditary Renal Amyloidosis Associated With a Novel Apolipoprotein A-II Variant. *Kidney Int Rep*. **2**, 1223–1232

119. Valleix, S., Verona, G., Jourde-Chiche, N., Nédelec, B., Mangione, P. P., Bridoux, F., Mangé, A., Dogan, A., Goujon, J.-M., Lhomme, M., Dauteuille, C., Chabert, M., Porcari, R., Waudby, C. A., Relini, A., Talmud, P. J., Kovrov, O., Olivecrona, G., Stoppini, M., Christodoulou, J., Hawkins, P. N., Grateau, G., Delpech, M., Kontush, A., Gillmore, J. D., Kalopissis, A. D., and Bellotti, V. (2016) D25V apolipoprotein C-III variant causes dominant hereditary systemic amyloidosis and confers cardiovascular protective lipoprotein profile. *Nat Commun*. **7**, 10353

120. Brodehl, A., Dieding, M., Klauke, B., Dec, E., Madaan, S., Huang, T., Gargus, J., Fatima, A., Saric, T., Cakar, H., Walhorn, V., Tönsing, K., Skrzipczyk, T., Cebulla, R., Gerdes, D., Schulz, U., Gummert, J., Svendsen, J. H., Olesen, M. S., Anselmetti, D., Christensen, A. H., Kimonis, V., and Milting, H. (2013) The novel desmin mutant p.A120D impairs filament formation, prevents intercalated disk localization, and causes sudden cardiac death. *Circ Cardiovasc Genet*. **6**, 615–623

121. Morfino, P., Aimo, A., Panichella, G., Rapezzi, C., and Emdin, M. (2022) Amyloid seeding as a disease mechanism and treatment target in transthyretin cardiac amyloidosis. *Heart*

Fail Rev. **27**, 2187–2200

122. Jucker, M., and Walker, L. C. (2013) Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. *Nature*. **501**, 45–51

123. Liberski, P. P. (2014) Prion, prionoids and infectious amyloid. *Parkinsonism Relat Disord*. **20 Suppl 1**, S80-84

124. Murray, K. A., Hu, C. J., Griner, S. L., Pan, H., Bowler, J. T., Abskharon, R., Rosenberg, G. M., Cheng, X., Seidler, P. M., and Eisenberg, D. S. (2022) De novo designed protein inhibitors of amyloid aggregation and seeding. *Proc Natl Acad Sci U S A*. **119**, e2206240119

125. Seidler, P. M., Boyer, D. R., Murray, K. A., Yang, T. P., Bentzel, M., Sawaya, M. R., Rosenberg, G., Cascio, D., Williams, C. K., Newell, K. L., Ghetti, B., DeTure, M. A., Dickson, D. W., Vinters, H. V., and Eisenberg, D. S. (2019) Structure-based inhibitors halt prion-like seeding by Alzheimer's disease-and tauopathy-derived brain tissue samples. *J Biol Chem*. **294**, 16451–16464

126. Sangwan, S., Sahay, S., Murray, K. A., Morgan, S., Guenther, E. L., Jiang, L., Williams, C. K., Vinters, H. V., Goedert, M., and Eisenberg, D. S. (2020) Inhibition of synucleinopathic seeding by rationally designed inhibitors. *Elife*. **9**, e46775

127. Kedia, N., Arhzaouy, K., Pittman, S. K., Sun, Y., Batchelor, M., Weihl, C. C., and Bieschke, J. (2019) Desmin forms toxic, seeding-competent amyloid aggregates that persist in muscle fibers. *Proc Natl Acad Sci U S A*. **116**, 16835–16840

128. Ano Bom, A. P. D., Rangel, L. P., Costa, D. C. F., de Oliveira, G. A. P., Sanches, D., Braga, C. A., Gava, L. M., Ramos, C. H. I., Cepeda, A. O. T., Stumbo, A. C., De Moura Gallo, C. V., Cordeiro, Y., and Silva, J. L. (2012) Mutant p53 aggregates into prion-like amyloid oligomers and fibrils: implications for cancer. *J Biol Chem*. **287**, 28152–28162

129. Rutherford, N. J., Dhillon, J.-K. S., Riffe, C. J., Howard, J. K., Brooks, M., and Giasson, B.

I. (2017) Comparison of the in vivo induction and transmission of α-synuclein pathology by mutant α-synuclein fibril seeds in transgenic mice. *Hum Mol Genet*. **26**, 4906–4915

130. Spina, S., Farlow, M. R., Unverzagt, F. W., Kareken, D. A., Murrell, J. R., Fraser, G., Epperson, F., Crowther, R. A., Spillantini, M. G., Goedert, M., and Ghetti, B. (2008) The tauopathy associated with mutation +3 in intron 10 of Tau: characterization of the MSTD family. *Brain*. **131**, 72–89

131. Spillantini, M. G., Murrell, J. R., Goedert, M., Farlow, M. R., Klug, A., and Ghetti, B. (1998) Mutation in the tau gene in familial multiple system tauopathy with presenile dementia. *Proc Natl Acad Sci U S A*. **95**, 7737–7741

132. Varani, L., Hasegawa, M., Spillantini, M. G., Smith, M. J., Murrell, J. R., Ghetti, B., Klug, A., Goedert, M., and Varani, G. (1999) Structure of tau exon 10 splicing regulatory element RNA and destabilization by mutations of frontotemporal dementia and parkinsonism linked to chromosome 17. *Proc Natl Acad Sci U S A*. **96**, 8229–8234

133. Hasegawa, M., Smith, M. J., Iijima, M., Tabira, T., and Goedert, M. (1999) FTDP-17 mutations N279K and S305N in tau produce increased splicing of exon 10. *FEBS Lett*. **443**, 93–96

134. Hutton, M., Lendon, C. L., Rizzu, P., Baker, M., Froelich, S., Houlden, H., Pickering-Brown, S., Chakraverty, S., Isaacs, A., Grover, A., Hackett, J., Adamson, J., Lincoln, S., Dickson, D., Davies, P., Petersen, R. C., Stevens, M., de Graaff, E., Wauters, E., van Baren, J., Hillebrand, M., Joosse, M., Kwon, J. M., Nowotny, P., Che, L. K., Norton, J., Morris, J. C., Reed, L. A., Trojanowski, J., Basun, H., Lannfelt, L., Neystat, M., Fahn, S., Dark, F., Tannenberg, T., Dodd, P. R., Hayward, N., Kwok, J. B., Schofield, P. R., Andreadis, A., Snowden, J., Craufurd, D., Neary, D., Owen, F., Oostra, B. A., Hardy, J., Goate, A., van Swieten, J., Mann, D., Lynch, T., and Heutink, P. (1998) Association of missense and 5’-splice-site mutations in tau with the inherited dementia FTDP-17.

Nature. **393**, 702–705

135. Goedert, M. (2005) Tau gene mutations and their effects. *Mov Disord*. **20 Suppl 12**, S45-52

136. Sikora, J., Kmochová, T., Mušálková, D., Pohludka, M., Přikryl, P., Hartmannová, H., Hodaňová, K., Trešlová, H., Nosková, L., Mrázová, L., Stránecký, V., Lunová, M., Jirsa, M., Honsová, E., Dasari, S., McPhail, E. D., Leung, N., Živná, M., Bleyer, A. J., Rychlík, I., Ryšavá, R., and Kmoch, S. (2022) A mutation in the SAA1 promoter causes hereditary amyloid A amyloidosis. *Kidney Int*. **101**, 349–359

137. Poa, N. R., Cooper, G. J. S., and Edgar, P. F. (2003) Amylin gene promoter mutations predispose to Type 2 diabetes in New Zealand Maori. *Diabetologia*. **46**, 574–578

138. Chartier-Harlin, M.-C., Kachergus, J., Roumier, C., Mouroux, V., Douay, X., Lincoln, S., Levecque, C., Larvor, L., Andrieux, J., Hulihan, M., Waucquier, N., Defebvre, L., Amouyel, P., Farrer, M., and Destée, A. (2004) Alpha-synuclein locus duplication as a cause of familial Parkinson's disease. *Lancet*. **364**, 1167–1169

139. Singleton, A. B., Farrer, M., Johnson, J., Singleton, A., Hague, S., Kachergus, J., Hulihan, M., Peuralinna, T., Dutra, A., Nussbaum, R., Lincoln, S., Crawley, A., Hanson, M., Maraganore, D., Adler, C., Cookson, M. R., Muenter, M., Baptista, M., Miller, D., Blancato, J., Hardy, J., and Gwinn-Hardy, K. (2003) alpha-Synuclein locus triplication causes Parkinson's disease. *Science*. **302**, 841

140. Haass, C., Hung, A. Y., Selkoe, D. J., and Teplow, D. B. (1994) Mutations associated with a locus for familial Alzheimer's disease result in alternative processing of amyloid beta-protein precursor. *J Biol Chem*. **269**, 17741–17748

141. Watson, D. J., Selkoe, D. J., and Teplow, D. B. (1999) Effects of the amyloid precursor protein Glu693-->Gln "Dutch" mutation on the production and stability of amyloid beta-protein. *Biochem J*. **340**, 703–709

142. De Strooper, B. (2003) Aph-1, Pen-2, and Nicastrin with Presenilin generate an active gamma-Secretase complex. *Neuron*. **38**, 9–12

143. Kim, S. H., Wang, R., Gordon, D. J., Bass, J., Steiner, D. F., Lynn, D. G., Thinakaran, G., Meredith, S. C., and Sisodia, S. S. (1999) Furin mediates enhanced production of fibrillogenic ABri peptides in familial British dementia. *Nat Neurosci*. **2**, 984–988

144. Ayala, Y. M., Zago, P., D'Ambrogio, A., Xu, Y.-F., Petrucelli, L., Buratti, E., and Baralle, F. E. (2008) Structural determinants of the cellular localization and shuttling of TDP-43. *J Cell Sci*. **121**, 3778–3785

145. Zinszner, H., Sok, J., Immanuel, D., Yin, Y., and Ron, D. (1997) TLS (FUS) binds RNA in vivo and engages in nucleo-cytoplasmic shuttling. *J Cell Sci*. **110 (Pt 15)**, 1741–1750

146. Nomura, T., Watanabe, S., Kaneko, K., Yamanaka, K., Nukina, N., and Furukawa, Y. (2014) Intranuclear aggregation of mutant FUS/TLS as a molecular pathomechanism of amyotrophic lateral sclerosis. *J Biol Chem*. **289**, 1192–1202

147. Sun, Y., Zhang, S., Hu, J., Tao, Y., Xia, W., Gu, J., Li, Y., Cao, Q., Li, D., and Liu, C. (2022) Molecular structure of an amyloid fibril formed by FUS low-complexity domain. *iScience*. **25**, 103701

148. Murray, D. T., Kato, M., Lin, Y., Thurber, K. R., Hung, I., McKnight, S. L., and Tycko, R. (2017) Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. *Cell*. **171**, 615-627.e16

149. Kwiatkowski, T. J., Bosco, D. A., Leclerc, A. L., Tamrazian, E., Vanderburg, C. R., Russ, C., Davis, A., Gilchrist, J., Kasarskis, E. J., Munsat, T., Valdmanis, P., Rouleau, G. A., Hosler, B. A., Cortelli, P., de Jong, P. J., Yoshinaga, Y., Haines, J. L., Pericak-Vance, M. A., Yan, J., Ticozzi, N., Siddique, T., McKenna-Yasek, D., Sapp, P. C., Horvitz, H. R., Landers, J. E., and Brown, R. H. (2009) Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science*. **323**, 1205–1208

150. Sun, Z., Diaz, Z., Fang, X., Hart, M. P., Chesi, A., Shorter, J., and Gitler, A. D. (2011) Molecular determinants and genetic modifiers of aggregation and toxicity for the ALS disease protein FUS/TLS. *PLoS Biol*. **9**, e1000614

151. Patel, A., Lee, H. O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M. Y., Stoynov, S., Mahamid, J., Saha, S., Franzmann, T. M., Pozniakovski, A., Poser, I., Maghelli, N., Royer, L. A., Weigert, M., Myers, E. W., Grill, S., Drechsel, D., Hyman, A. A., and Alberti, S. (2015) A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell*. **162**, 1066–1077

152. Schröder, R., Goudeau, B., Simon, M. C., Fischer, D., Eggermann, T., Clemen, C. S., Li, Z., Reimann, J., Xue, Z., Rudnik-Schöneborn, S., Zerres, K., van der Ven, P. F. M., Fürst, D. O., Kunz, W. S., and Vicart, P. (2003) On noxious desmin: functional effects of a novel heterozygous desmin insertion mutation on the extrasarcomeric desmin cytoskeleton and mitochondria. *Hum Mol Genet*. **12**, 657–669

153. Jeganathan, S., von Bergen, M., Mandelkow, E.-M., and Mandelkow, E. (2008) The natively unfolded character of tau and its aggregation to Alzheimer-like paired helical filaments. *Biochemistry*. **47**, 10526–10539

154. Dao, T. P., Kolaitis, R.-M., Kim, H. J., O'Donovan, K., Martyniak, B., Colicino, E., Hehnly, H., Taylor, J. P., and Castañeda, C. A. (2018) Ubiquitin Modulates Liquid-Liquid Phase Separation of UBQLN2 via Disruption of Multivalent Interactions. *Mol Cell*. **69**, 965-978.e6

155. Weinreb, P. H., Zhen, W., Poon, A. W., Conway, K. A., and Lansbury, P. T. (1996) NACP, a protein implicated in Alzheimer's disease and learning, is natively unfolded. *Biochemistry*. **35**, 13709–13715

156. Uversky, V. N. (2003) A protein-chameleon: conformational plasticity of alpha-synuclein, a disordered protein involved in neurodegenerative disorders. *J Biomol Struct Dyn*. **21**, 211–234

157. Sethi, S., Dasari, S., Plaisier, E., Ronco, P., Nasr, S. H., Brocheriou, I., Theis, J. D., Vrana, J. A., Zimmermann, M. T., Quint, P. S., McPhail, E. D., and Kurtin, P. J. (2018) Apolipoprotein CII Amyloidosis Associated With p.Lys41Thr Mutation. *Kidney Int Rep*. **3**, 1193–1201

158. Nasr, S. H., Dasari, S., Hasadsri, L., Theis, J. D., Vrana, J. A., Gertz, M. A., Muppa, P., Zimmermann, M. T., Grogg, K. L., Dispenzieri, A., Sethi, S., Highsmith, W. E., Merlini, G., Leung, N., and Kurtin, P. J. (2017) Novel Type of Renal Amyloidosis Derived from Apolipoprotein-CII. *J Am Soc Nephrol*. **28**, 439–445

159. 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

160. Lagerstedt, J. O., Cavigiolio, G., Roberts, L. M., Hong, H.-S., Jin, L.-W., Fitzgerald, P. G., Oda, M. N., and Voss, J. C. (2007) Mapping the structural transition in an amyloidogenic apolipoprotein A-I. *Biochemistry*. **46**, 9693–9699

161. Teoh, C. L., Pham, C. L. L., Todorova, N., Hung, A., Lincoln, C. N., Lees, E., Lam, Y. H., Binger, K. J., Thomson, N. H., Radford, S. E., Smith, T. A., Müller, S. A., Engel, A., Griffin, M. D. W., Yarovsky, I., Gooley, P. R., and Howlett, G. J. (2011) A structural model for apolipoprotein C-II amyloid fibrils: experimental characterization and molecular dynamics simulations. *J Mol Biol*. **405**, 1246–1266

162. Solomon, J. P., Page, L. J., Balch, W. E., and Kelly, J. W. (2012) Gelsolin amyloidosis: genetics, biochemistry, pathology and possible strategies for therapeutic intervention. *Crit Rev Biochem Mol Biol*. **47**, 282–296

163. Tiwari, A., Xu, Z., and Hayward, L. J. (2005) Aberrantly increased hydrophobicity shared by mutants of Cu,Zn-superoxide dismutase in familial amyotrophic lateral sclerosis. *J Biol Chem*. **280**, 29771–29779

164. Sharkey, L. M., Safren, N., Pithadia, A. S., Gerson, J. E., Dulchavsky, M., Fischer, S., Patel, R., Lantis, G., Ashraf, N., Kim, J. H., Meliki, A., Minakawa, E. N., Barmada, S. J., Ivanova, M. I., and Paulson, H. L. (2018) Mutant UBQLN2 promotes toxicity by modulating intrinsic self-assembly. *Proc Natl Acad Sci U S A*. **115**, E10495–E10504

165. Ghosh, D., Sahay, S., Ranjan, P., Salot, S., Mohite, G. M., Singh, P. K., Dwivedi, S., Carvalho, E., Banerjee, R., Kumar, A., and Maji, S. K. (2014) The newly discovered Parkinson's disease associated Finnish mutation (A53E) attenuates α-synuclein aggregation and membrane binding. *Biochemistry*. **53**, 6419–6421

166. Robotta, M., Cattani, J., Martins, J. C., Subramaniam, V., and Drescher, M. (2017) Alpha-Synuclein Disease Mutations Are Structurally Defective and Locally Affect Membrane Binding. *J Am Chem Soc*. **139**, 4254–4257

167. Lin, M.-W., Lee, D.-D., Liu, T.-T., Lin, Y.-F., Chen, S.-Y., Huang, C.-C., Weng, H.-Y., Liu, Y.-F., Tanaka, A., Arita, K., Lai-Cheong, J., Palisson, F., Chang, Y.-T., Wong, C.-K., Matsuura, I., McGrath, J. A., and Tsai, S.-F. (2010) Novel IL31RA gene mutation and ancestral OSMR mutant allele in familial primary cutaneous amyloidosis. *Eur J Hum Genet*. **18**, 26–32

168. Chiang, Y.-Y., Chao, S.-C., Chen, W.-Y., Lee, W.-R., and Wang, K.-H. (2008) Weber-Cockayne type of epidermolysis bullosa simplex associated with a novel mutation in keratin 5 and amyloid deposits. *Br J Dermatol*. **159**, 1370–1372

169. Wang, L.-Q., Ma, Y., Yuan, H.-Y., Zhao, K., Zhang, M.-Y., Wang, Q., Huang, X., Xu, W.-C., Dai, B., Chen, J., Li, D., Zhang, D., Wang, Z., Zou, L., Yin, P., Liu, C., and Liang, Y. (2022) Cryo-EM structure of an amyloid fibril formed by full-length human SOD1 reveals its conformational conversion. *Nat Commun*. **13**, 3491

170. Marin-Argany, M., Lin, Y., Misra, P., Williams, A., Wall, J. S., Howell, K. G., Elsbernd, L. R., McClure, M., and Ramirez-Alvarado, M. (2016) Cell Damage in Light Chain

Amyloidosis: FIBRIL INTERNALIZATION, TOXICITY AND CELL-MEDIATED SEEDING. *J Biol Chem*. **291**, 19813–19825

171. Iwaya, K., Zako, T., Fukunaga, J., Sörgjerd, K. M., Ogata, K., Kogure, K., Kosano, H., Noritake, M., Maeda, M., Ando, Y., Katsura, Y., and Nagase, T. (2019) Toxicity of insulin-derived amyloidosis: a case report. *BMC Endocr Disord*. **19**, 61
172. Pieri, L., Madiona, K., Bousset, L., and Melki, R. (2012) Fibrillar α-Synuclein and Huntingtin Exon 1 Assemblies Are Toxic to the Cells. *Biophys J*. **102**, 2894–2905
173. Sidhu, A., Segers-Nolten, I., and Subramaniam, V. (2016) Conformational Compatibility Is Essential for Heterologous Aggregation of α-Synuclein. *ACS Chem Neurosci*. **7**, 719–727
174. Ihse, E., Suhr, O. B., Hellman, U., and Westermark, P. (2011) Variation in amount of wild-type transthyretin in different fibril and tissue types in ATTR amyloidosis. *J Mol Med (Berl)*. **89**, 171–180
175. Cao, P., Tu, L.-H., Abedini, A., Levsh, O., Akter, R., Patsalo, V., Schmidt, A. M., and Raleigh, D. P. (2012) Sensitivity of amyloid formation by human islet amyloid polypeptide to mutations at residue 20. *J Mol Biol*. **421**, 282–295
176. Seidler, P. M., Murray, K. A., Boyer, D. R., Ge, P., Sawaya, M. R., Hu, C. J., Cheng, X., Abskharon, R., Pan, H., DeTure, M. A., Williams, C. K., Dickson, D. W., Vinters, H. V., and Eisenberg, D. S. (2022) Structure-based discovery of small molecules that disaggregate Alzheimer's disease tissue derived tau fibrils in vitro. *Nat Commun*. **13**, 5451
177. Griner, S. L., Seidler, P., Bowler, J., Murray, K. A., Yang, T. P., Sahay, S., Sawaya, M. R., Cascio, D., Rodriguez, J. A., Philipp, S., Sosna, J., Glabe, C. G., Gonen, T., and Eisenberg, D. S. (2019) Structure-based inhibitors of amyloid beta core suggest a common interface with tau. *Elife*. **8**, e46924

Chapter 5: Amyloid protein profiles

As discussed in chapter 4, not all of the proteins considered in chapter 4 are formally considered amyloids by many in the field. In this chapter, we will name each protein (and its gene name) in our list, present a short description of it, name the disease or diseases its amyloid forms are associated with, describe the evidence of its amyloid nature, and describe the rationale for the amyloid mutation mechanisms assigned to it. The proteins are ordered alphabetically by protein name.

Amyloid-β precursor protein (APP)

Amyloid-β precursor protein is the protein which is cleaved by secretase complexes to produce the amyloid-β peptide. The function of amyloid-β is not entirely understood, but it has been proposed to have a variety of beneficial functions including antimicrobial activity, tumor suppression, blood-brain barrier upkeep, recovery from brain injury, and synaptic function regulation(1). Aggregates of amyloid-β are a hallmark of Alzheimer's disease, cerebral amyloid angiopathy, and Down syndrome(2–5). This peptide was identified as the main component of amyloid deposits in the brains of Alzheimer's patients and people with Down syndrome through mass spectrometry analysis of congophilic materials from patient brains(3, 6, 7). The pathogenic mutations in this protein can affect its cleavage by α-, β-, and γ-secretase(2, 8–13) as well as result in a fiber structure which is more stable than the wild-type(14, 15), the mechanisms of fiber stabilization and altered processing were assigned to this protein's mutations.

Apolipoprotein A I (APOA1)

Apolipoprotein A I is a protein which binds cholesterol and phospholipids and is the principal component of high-density lipoproteins (HDL)(16). This protein was first identified as a component

of amyloid deposits when apolipoprotein A I with a G26R mutation was identified by amino acid sequence analysis of tryptic peptides from congophilic material from a patient's spleen with familial amyloid polyneuropathy type III(17), and wild-type apolipoprotein A I was found to form amyloids when an N-terminal fragment was isolated from congophilic amyloid deposits in atherosclerotic plaques(18). Mutations in this protein have various possible mechanisms associated with amyloid formation, depending on which mutation the protein has, including destabilization of the native structure, an increase in fiber-stabilizing β-sheet secondary structure, altered processing due to increased availability of the cleavage site which produces the amyloidogenic fragment, and decreased binding to its native lipid binding partners(16, 19–21).

Apolipoprotein A II (APOA2)

Apolipoprotein A II is another protein which is a component of HDL(22). This protein was first identified as an amyloid in a case of renal amyloidosis by isolation of congophilic amyloid material from the kidneys and use of Edman degradation sequence analysis(23). When associated with HDL, this protein is aggregation-resistant, but separation from bound lipids makes it very prone to misfolding(24, 25). This protein is a proto-amyloid, so only mutations cause it to be found in an amyloid state. All known mutations in this protein are stop codon mutations which extend the protein by 21 residues(23, 24, 26–28). All amyloidogenic mutations create nearly the same aggregation-prone segment to the C-terminal of the protein, both destabilizing the native structure, stabilizing a fiber form, and detaching the protein from its native binding partners(24, 25).

Apolipoprotein A IV (APOA4)

Apolipoprotein A IV is a lipid-binding protein involved in various physiological functions related to lipid metabolism including being protective against atherosclerosis and inhibiting lipoprotein

oxidation(29)**.** This protein was first identified as amyloidogenic when an N-terminal fragment was identified as a component of amyloid deposits in the heart of a patient with senile systemic amyloidosis (SSA) associated with the aggregation of wild-type transthyretin(30). This protein has no associated amyloidogenic mutations, but the aggregation-prone fragment seems to be an N-terminal signal sequence that is not present in healthy controls(31).

Apolipoprotein C II (APOC2)

Apolipoprotein C II is a component of various triglyceride-rich lipoproteins and functions in the hydrolysis of plasma triglycerides(32). This protein was identified as an amyloid in a case of renal amyloidosis through mass spectroscopy analysis of congophilic amyloid material from the kidney of the patient(33). The wild-type form of this protein had previously been shown to be able to form amyloid fibers *in vitro* (albeit in a lipid-unbound state)(34) so we could not classify it as a proto-amyloid, despite only being found in amyloid deposits in humans when it is mutated. The amyloidogenic mutations are thought to destabilize the native structure of the protein which, in turn, also interferes with its lipid-binding capabilities(33, 35).

Apolipoprotein C III (APOC3)

Apolipoprotein C III functions to raise plasma triglyceride levels by inhibiting the hydrolysis of triglycerides(36). This protein was identified as an amyloid in a French family with severe renal amyloidosis by immunohistochemistry on congophilic amyloid deposits in various tissues(37). The wild-type form of this protein had previously been shown to be able to form amyloid fibers *in vitro*, (albeit in a lipid-unbound state)(38) so we could not classify it as a proto-amyloid, despite only being found in amyloid deposits in humans when it is mutated, and only mutant protein being found in *ex vivo* amyloid samples retrieved from patients(37). The amyloidogenic mutation

disrupts the native structure, inducing more fiber-stabilizing β-sheet secondary structure, and reduces its efficiency at binding lipids(37).

Atrial natriuretic factor (ANF)

Atrial natriuretic factor is a peptide hormone secreted by the heart atria in order to regulate blood volume and pressure through acting on the kidneys to increase sodium excretion(39, 40). It is the main component of the amyloid deposits in isolated atrial amyloidosis and was first identified as an amyloid through electron microscopy-based ultrastructural analysis and immunogold staining of amyloid fibers in a piece of right atrial appendage removed in a coronary bypass surgery, although Congo red staining was negative(41). This protein is a sporadic amyloid, but its amyloid aggregation is associated with increased expression of the peptide; this is hard to disentangle from age-related factors, though(42, 43).

C9orf72 dipeptide repeat protein (C9orf72)

C9orf72 dipeptide repeat (DPR) proteins are generated from RNA transcripts from a C9orf72 gene containing an intronic hexanucleotide repeat expansion of the sequence GGGGCC. This repeat expansion mutation causes ALS/FTD(44, 45). The RNA undergoes aberrant translation potentially via repeat-associated non-ATG translation(46). This process can generate three types of dipeptide repeat proteins: glycine-alanine repeats (GA), glycine-arginine repeats (GR), and glycine-proline repeats (GP)(47); only the GA protein has been shown to form amyloid fibers. The amyloid nature of GA DPR proteins has been demonstrated only in vitro through ThT fluorescence assays, Congo red staining, electron microscopy analysis, atomic force microscopy analysis, and wide angle x-ray scattering of synthetic peptides(48, 49), although longer constructs expressed in bacteria formed fibers that did not bind ThT but still have cross-β secondary structure typical of amyloid as revealed by FTIR measurement(50). Since the GA DPR is an aberrantly translated

protein from a normally noncoding DNA sequence, meaning there is no wild-type version of the protein, we have grouped this protein with the special case proteins, despite resulting from a repeat expansion genetic mutation. For this reason, we did not assign a mutation mechanism to this protein.

Calcitonin (CALCA)

Calcitonin is a peptide hormone secreted by the thyroid gland which functions to prevent hypercalcemia by reducing serum calcium levels(51). Calcitonin amyloid is found in thyroid tissue with medullary thyroid carcinoma (MTC) and can also be found deposited in the kidneys of patients with MTC(52–54). The amyloid material in MTC was first identified as possibly an alternately processed prohormone of calcitonin by Edman degradation sequence analysis of congophilic amyloid material from a patient with MTC(55). This identification was confirmed by immunogold staining(56) and refined by mass spectrometry to show that the amyloid consists of the normal, full-length calcitonin hormone and not an alternately processed prohormone(52). Calcitonin is a sporadic amyloid, since no mutations are associated with its amyloid formation, but since it is associated with cancerous thyroid tissue, calcitonin's amyloid aggregation may be downstream of significant overexpression.

Cathepsin K (CTSK)

Cathepsin K is an extremely potent protease secreted by osteoclasts which degrades collagen during bone resorption(57). It is also expressed by multinucleated giant cells and may have a role in degrading amyloid fibers, ironically(58). This protein was identified by Edman degradation sequence analysis as the amyloid component of a congophilic angiomyolipoma, determined to be a hamartoma, in a woman's kidney, which was removed(59). This tumor and the unpublished results of an *in vitro* study of a synthetic peptide by the same group who reported the tumor are

the only data points for the amyloidogenicity of this protein, and it is unclear if the patient who was the source of the tumor had any genetic variants in their CTSK gene. So, unless it is shown otherwise, cathepsin K will be characterized as a sporadic amyloid with no known amyloidogenic mutations.

Cellular tumor antigen p53 (TP53)

p53 is a tumor suppressor whose loss of function is associated with over 50% of human cancers(60). Amyloid formation of this protein has been shown to transform it into an oncoprotein(61–63) and amyloid deposits consisting of p53 has been shown in various cancer tissues by immunostaining and staining with amyloidophilic dyes(62, 64) . Both wild-type and mutant p53 is able to form amyloid fibers, and mutations encourage amyloid formation by destabilization of the native tetrameric form(65) along with increased aggregation propensity (fiber stabilization) and increased seeding activity(62–64, 66).

Corneodesmosin (CDSN)

Corneodesmosin is a glycoprotein found in the cornified squamous epithelia and functions in cell adhesion in skin and hair follicles(67). Its amyloid formation is associated with hypotrichosis simplex of the scalp (HSS) and it was identified as the constituent of the amyloid deposits in HSS by immunohistochemical staining of congophilic biopsies from HSS patients(68). Corneodesmosin is a proto-amyloid and all the amyloidogenic mutations in this protein are nonsense mutations which truncate the protein(68–70). The full-length protein is almost entirely disordered(68), and the production of a shorter disordered version apparently favors fiber-formation over its native function. Since both the full-length protein and the truncations are already intrinsically disordered, the only mechanism assigned to the mutations is native structure

destabilization, since they do truncate the protein and remove whatever was interrupting their amyloid aggregation.

Cystatin C (CST3)

Cystatin C is a cysteine protease inhibitor found in bodily fluids(71). It may be a functional amyloid, with the wild-type protein contributing to the formation of the epididymal luminal amyloid matrix in mice(72) and the wild-type protein has been shown to form amyloid fibers *in vitro* via ThT assay and electron microscopy(73). When mutated, its amyloid aggregation causes hereditary cystatin C amyloid angiopathy (HCAA), otherwise known as hereditary cerebral hemorrhage with amyloidosis (HCHWA)(71, 74–77). Cystatin C was identified as the amyloid protein responsible for this disease through amino acid sequence analysis of amyloid fibers purified from patient tissue(75). Cystatin C is an ambimorph amyloid for which a single mutation, L94Q, is known to cause its associated disease. This mutation introduces a polar side chain into a hydrophobic pocket of the protein and encourages its misfolding(77). How this affects the resulting amyloid fiber is less clear, so the only assigned amyloidogenic mechanism for this mutation is native structure destabilization.

Desmin (DES)

Desmin is an intermediate filament protein present in muscle fibers which forms an extra-sarcomeric cytoskeleton connecting myofibers to each other and other structures(78). Conversion of desmin into amyloid fibers is hypothesized to be associated with myofibrillar myopathy, cases of which have been shown to develop congophilic lesions which desmin colocalizes with(79, 80). Wild-type and mutant desmin has been shown to form amyloid fibers *in vitro* and mutant desmin has accelerated fiber formation(81). Desmin amyloid was also shown to be toxic to mouse myoblast cells(81). Mutations in desmin destabilize the native protein, as demonstrated by its

reduced solubility and critical concentration for amyloid formation(81) and seem to stabilize the amyloid form since not all destabilizing mutations in this protein cause it to form amyloids. The mutations also have been shown to exhibit increased seeding(81) as well as induce mislocalization of the protein(82, 83), although it is unclear if the mislocalization precedes amyloid formation.

EGF-containing fibulin-like extracellular matrix protein 1 (EFEMP1)

EGF-containing fibulin-like extracellular matrix protein 1, also known as fibulin-3, is a protein which competes with epidermal growth factor for binding to the EGF receptor and promotes tumor growth in adenocarcinoma(84). While it has a role in the progression of cancer, this protein also forms amyloids mainly in the venous walls of the bowels of mainly elderly females, but also in other tissues(85), and was first identified to do so by mass spectrometry and immunohistochemical analysis of congophilic intestinal venous walls obtained at autopsy from a patient(86). This protein is a sporadic amyloid, so has no mutations associated with its amyloid formation, and, in fact, a patient with Doyne honeycomb retinal dystrophy caused by an autosomal dominant mutation in fibulin-3 did not have amyloid deposits of the protein(86). However, higher expression of the protein seems to accompany aging and the amyloid deposits consist of a C-terminal fragment of the protein(86), although it is unclear why this cleavage product is generated.

Fibrinogen α chain (FGA)

Fibrinogen α chain is a glycoprotein which is essential for blood coagulation(87). Its amyloid formation is associated with hereditary renal amyloidosis and it was first identified as the amyloid component by amino acid sequence analysis of amyloid material harvested from the renal transplant of a patient (harvested postmortem), and all sequences corresponded to the C-terminal portion of the protein(88). Fibrinogen α chain is a proto-amyloid and there are 15 known

amyloidogenic mutations in it, all in the C-terminal region, which consist of substitution mutations, indels, and frame-shifts(87). Some of these mutations (namely the frame-shifts) seem to interfere with the normal function of the protein, evidenced by lower circulating plasma levels of it(89) so the mechanism of native structure destabilization was assigned to this protein. However, many other mutations (namely the substitution mutations) do not seem to interfere with the function(90), and so likely do not significantly affect the native structure, so we can infer that these mutations stabilize a fiber form. Also, since the amyloid seems to consist exclusively of a C-terminal fragment, and wild-type C-terminal fragments are not found in the amyloid deposits, we can infer that the normal processing of the protein has been disrupted by the mutations.

Galectin-7 (LGALS7)

Galectin-7 is an epidermal protein with various functions including controlling apoptosis, cell migration, and cell adhesion(91). Its amyloid formation was associated with localized cutaneous amyloidosis by identification of galectin-7 and actin as components of amyloid deposits in skin lesions of patients through mass spectrometry and immunohistochemistry analyses(92). However, this result was contested by a more refined analysis using mass spectrometry and immunohistochemistry analysis of laser microdissected of skin biopsies which detected only keratin proteins (mainly keratin-5) in the congophilic material and galectin-7 only in the surrounding non-congophilic epidermis, while actin was found in both(93). Still, galectin-7 and peptide fragments of the protein were shown to be capable of forming amyloid fibers *in vitro*, though only at very low pH (pH 2.0 and 4.0)(94). The amyloid nature of galectin-7 is somewhat unclear, but if it can form amyloids it would be a sporadic amyloid, as no amyloidogenic mutations have been found in it.

Gelsolin (GSN)

Gelsolin is a calcium-binding protein which modulates the growth of actin filaments(95). Its amyloid formation is associated with hereditary gelsolin amyloidosis, also known as familial amyloidosis of the Finnish type, which presents as lattice corneal dystrophy(95–101). Its amyloid nature was first identified through Edman degradation sequence analysis of congophilic amyloid material from a patient's kidney obtained at autopsy(99) and confirmed by another group very shortly after(100). The amyloid deposits consisted of a central fragment of the protein and the group which identified gelsolin as the amyloid protein first also confirmed that the amyloidogenic protein fragment had a D to N substitution(101). This protein is a proto-amyloid, so only familial mutations have been found to enable its amyloid formation, namely two mutations at a single residue: D214N and D214Y (also numbered D187 for the mature protein). These mutations interrupts gelsolin's calcium binding activity, which also causes it to spend a longer time in an intermediate state between its active and inactive state(95). This intermediate state is more susceptible to furin-mediated cleavage, which produces a fragment which is further cleaved to eventually produce the amyloidogenic fragment(95). For these reasons, we assigned the mutation mechanisms of native structure destabilization, altered proceeding, and decreased binding to native partners. Also, it has been shown the fragments corresponding to the amyloidogenic fragment but with the wild-type sequence do not form amyloid fibers *in vitro* while the mutant fragment does(95, 96, 102, 103), so the mechanism of fiber stabilization was also assigned.

Glucagon (GCG)

Glucagon is a peptide hormone secreted by alpha cells of the pancreas which regulates blood glucose levels by encouraging production of glucose through the breakdown of energy storage molecules like glycogen and triglycerides(104). It was found in an amyloid form in a patient with pancreatic neuroendocrine tumors which were positive for Congo red staining(105). The amyloid was confirmed to consist of glucagon through mass spectrometry analysis with laser

microdissection along with immunohistochemistry. The peptide was also shown to form fibers rapidly at the acidic pH required to solubilize it and an atomic structure of the fibers was solved using solid-state NMR(106). This protein is ostensibly a sporadic amyloid, as no mutations have been associated with its fiber formation, but in the case of the pancreatic tumor glucagon was being produced in high quantities but was not being secreted by the pancreas, and the tumor may have needed to reach a non-physiological critical concentration of nonfunctional glucagon before amyloid deposits began to form.

Glucagon-like peptide 1 (Liraglutide)

Liraglutide is a peptide drug which is a mimic of glucagon-like peptide 1 which is administered through subcutaneous injections for the management of diabetes and acts through stimulating glycogenesis(107). It was found in amyloid deposits of abdominal skin biopsies of an elderly man taking the drug to manage diabetes and was confirmed as the main constituent of the amyloid through mass spectrometry analysis of the samples(107). The dangers of this amyloid buildup were noted to be possible drug resistance due to poor absorption as well as misdiagnosis of AL amyloidosis(107). This is included in our list of amyloid proteins since, although it is a drug, it is a peptide with over 90% sequence homology to the peptide hormone glucagon-like peptide 1. It is also worth noting that Liraglutide has some important differences from the hormone it is based on including being a shorter version of the peptide and having a substitution corresponding to K125R using the numbering of the glucagon prohormone. Because of this, it is not entirely clear if the actual glucagon-like peptide 1 can form amyloids in the same way.

Heterogeneous nuclear ribonucleoprotein A1 (HNRNPA1)

Heterogeneous nuclear ribonucleoprotein A1 is potentially a functional amyloid, as there is evidence that is able to form reversible amyloid fibers and this reversible form of aggregation is

necessary for its function(108). This protein is an RNA-binding protein mainly localizing to the nucleus with various functions in RNA processing including transcription, splicing, translation, nuclear export, and others(109). Its reversible aggregation is related to its ability to form stress granules in the cytoplasm during cell stress(108). Amyloidogenic mutations in this protein are associated with ALS and MSP(110). Although amyloid fibers of this protein have not been isolated from human tissue thus far, there is ample *in vitro* evidence (ThT assays and electron microscopy including a cryo-EM structure of the wild-type amyloid) that the protein is able to form amyloid fibers, and that this activity is enhanced by disease-relevant mutations(110, 111). The amyloidogenic mutations in this protein fall in its low-complexity domain, or prion-like domain, which is intrinsically disordered and already able to form a reversible fiber, so the only mechanism assigned is stabilization of the fiber form. The PY-nuclear localization signal of the protein also appears to be a key driver of its self-association (being the main component of the wild-type amyloid fiber(111)) and is within the low-complexity domain, however it is unclear if the mutations affect the normal activity of the PY-nuclear localization signal, so mislocalization was not assigned as a mutation mechanism.

Heterogeneous nuclear ribonucleoprotein A2 (HNRNPA2B1)

Heterogeneous nuclear ribonucleoprotein A2 is the main isoform of the two spliceoforms of the HNRNPA2B1 gene. This RNA-binding protein which mainly localizes to the nucleus is potentially a functional amyloid and has similar functions related to RNA metabolism as the previous entry and also forms cytoplasmic stress granules under cell stress(112). A mutation in this protein, D290V (also numbered D302V for the longer isoform), is associated with MSP(110). Although amyloids of this protein have not been extracted from human tissue, there is ample *in vitro* evidence (ThT assays and electron microscopy including both wild-type and mutant cryo-EM structures) that both wild-type and mutant protein can form amyloid fibers, and that the disease

mutation enhances fiber formation(110, 113) **(mutant structure unpublished)**. Since the region of the protein which drives fiber formation is a low-complexity domain which is intrinsically disordered and already able to form a reversible fiber, the mutation mechanism was assigned to be fiber stabilization. Also, the cryo-EM structure of the mutant fiber reveals that the mutation causes the PY-nuclear localization signal of this protein to become buried in the fiber **(unpublished)** core while in the wild-type structure it is exposed, so the mutation may be encouraging an aggregated form which precludes relocalization to the nucleus after the formation of cytoplasmic stress granules, so subcellular mislocalization was also assigned as a mutation mechanism.

It is worth noting here that another RNA-binding protein, heterogeneous nuclear ribonucleoprotein D-like (HNRNPDL) is a functional amyloid with a cryo-EM structure of its reversible amyloid form which also has a disease-causing mutation in an aspartic acid residue in its low-complexity domain. However, the mutant forms of this protein are actually less prone to aggregate than the wild-type and cytoplasmic inclusions are absent in those with these mutations(114), so while it is an amyloid protein, its pathogenic mechanism is not likely to be amyloid formation.

Huntingtin (HTT)

Huntingtin is a protein whose function is not explicitly known, but it potentially has various roles including mediating trafficking of vesicles and organelles, regulating transcription, and acting as an antiapoptotic agent(115). Regardless, this protein is essential, as double-knockouts in mice are embryonic lethal(116), and haploinsufficient. Huntingtin is found in an aggregated state in the brains of individuals with Huntington's disease (HD)(117–119). Brain samples from HD patients display positive Congo red staining and cellulose acetate filter assays, which capture insoluble

protein aggregates, show that insoluble material from patient brains contains huntingtin protein in a conformation distinct from its soluble form(117). Also, *in vitro* studies have shown that huntingtin exon 1 recombinant protein aggregates into congophilic aggregates with ultrastructural features typical of amyloid(120–122). HD is caused by a polyglutamine expansion in exon 1 of the HTT gene which codes for huntingtin(120), and aggregation of huntingtin exon 1 recombinant protein into amyloid fibers is dependent on having a pathological number of glutamine repeats, making huntingtin a proto-amyloid. Aggregates in patients' brains are mainly composed of N-terminal fragments of huntingtin(118, 119), so *in vitro* experiments on N-terminal constructs of huntingtin have disease relevance. Polyglutamine tracts tend to be intrinsically disordered(123), but at a critical length of repeats, ~40 minimum for huntingtin, the formation of β-sheets with polar zippers becomes energetically favorable(124, 125), so only fiber stabilization was assigned as the amyloidogenic mechanism of the mutation. The evidence of amyloid formation by other proteins containing pathogenic polyglutamine expansions is not as strong as for huntingtin.

Immunoglobulin heavy chain (IGH)

The immunoglobulin heavy chain is the large subunit of an antibody, or immunoglobulin, and is linked to another heavy chain and a light chain by disulfide bonds. The heavy chain consists of a "variable" region, which are different between individual antibodies, and multiple "constant" regions, which are conserved between individual antibodies. The amyloid aggregation of this protein is associated with what is called "primary amyloidosis" or multiple myeloma-associated amyloidosis(126–129), the same disease caused by amyloid aggregation of the antibody light chain. When caused by the heavy chain, it is referred to as AH amyloidosis. This protein was first found to form amyloids in this disease through immunoblotting and amino acid sequence analysis of congophilic amyloid material extracted from a patient's spleen(126). Heavy chain amyloidosis has been reported several times since the initial report(127–129) and in all but one case(128) the

amyloid fibers were composed of a heavy chain fragment which included the variable domain. Interestingly, in the first reported case the amyloid protein was the heavy chain variable region connected directly to the third constant region, constituting a large internal deletion(126). This is the reason this protein was one of the “special cases”: the fragments forming amyloids in people all ostensibly have distinct amino acid sequences from each other and even from other antibodies within the same patient. This makes it difficult to connect the protein’s amino acid sequence to its amyloidogenicity.

Immunoglobulin light chain (IGL or IGK)

The immunoglobulin light chain is the small subunit of an antibody, or immunoglobulin, and is linked to a heavy chain by a disulfide bond. The light chain consists of a “variable” region, which are different between individual antibodies, and a “constant” region, which are conserved between individual antibodies. The amyloid aggregation of this protein is associated with what is called “primary amyloidosis” or multiple myeloma-associated amyloidosis(130–134), the same disease caused by amyloid aggregation of the antibody heavy chain. When caused by the light chain, it is referred to as AL amyloidosis. This kind of amyloidosis has long been associated with conditions like myeloma, and so a connection to immunoglobulin proteins had been hypothesized long before it was confirmed. Gamma globulin was shown to be a main component of the congophilic amyloid material in human patients as early as 1956(135). The sequence of the protein component of this amyloid material was later confirmed to be the sequence of the antibody light chain(131, 132). The amyloid component always contains the variable region of the protein which is the reason this protein was grouped into “special cases”: the fragments forming amyloids in people all ostensibly have distinct amino acid sequences from each other and even from other antibodies within the same patient. However, certain amino acid compositions have been

associated with higher incidence of amyloidosis(133, 134, 136, 137) and λ light chains form amyloids more often than κ light chains(138).

Insulin (INS)

Insulin is a hormone secreted by the beta cells of the pancreas which functions to regulate blood glucose levels by decreasing blood glucose through signaling cells to uptake blood glucose and store it(139). Insulin is an iatrogenic amyloid, so its amyloidosis is associated with drug forms of the protein, not the native protein. These drugs include porcine insulin, glargine, lispro, and others(140–143). Insulin forms subcutaneous amyloid deposits at sites of injection. This amyloidosis has been referred to as “insulin ball”, but is more commonly referred to as insulin-derived amyloidosis(141–144). There is some evidence to suggest these amyloid deposits are toxic to surrounding tissue(141) and there is at least one case of insulin amyloid deposits increasing in size even after decreasing insulin dosage and cessation of injections into existing amyloid deposits(144). Otherwise, adverse effects are mainly interference with insulin absorption leading to reduced efficacy of insulin drugs(142, 143). Insulin was first shown to be amyloidogenic for the case of porcine insulin through immunohistochemistry and amino acid sequence analysis of congophilic material from a patient’s thigh biopsy(140).

Integral membrane protein 2B (ITM2B or BRI or BRI2)

Integral membrane protein 2B is a protein whose function is not entirely clear, but there is evidence to suggest it has roles in triggering apoptosis as well as inhibition of the buildup and aggregation of amyloid-β peptide(145–148). It has a furin cleavage site near the C-terminal of the protein, and the normal protein is cleaved here during its processing(149). The C-terminal cleavage product forms amyloid deposits in familial British dementia (FBD) and familial Danish dementia (FDD), as identified by mass spectrometry analysis of isolated congophilic amyloid

material from patients and immunohistochemistry(150, 151). These two diseases are caused by two different, but related mutations. FBD is caused by a stop codon mutation which changes the normal stop codon (codon 267) to a codon for arginine, extending the protein from 267 to 277 amino acids(151). FDD is caused by a frame-shift mutation caused by a decamer duplication in the DNA sequence between codons 265 and 266 also extending the protein to 277 amino acids(150). Since only mutation results in this protein forming amyloids, it is a proto-amyloid. Though each mutation results in a different C-terminal amino acid sequence, both cause the resulting extended C-terminal cleavage product (both the same length) to become amyloidogenic. Since the cleavage product is a 34-amino acid peptide, which likely lacks significant secondary structure, and production of this peptide is enhanced when a mutation is present, the amyloidogenic mechanisms of fiber stabilization and altered processing were assigned to this protein's mutations.

Interleukin-1 receptor antagonist protein (Anakinra)

Anakinra is a recombinant protein drug which acts as an IL-1 blocker for the treatment of rheumatoid arthritis and neonatal onset multisytem inflammatory disease (NOMID). This is an iatrogenic amyloid: anakinra-associated amyloidosis is caused by subcutaneous injection of the drug. Anakinra was confirmed as the amyloidogenic agent by mass spectrometry of laser dissected congophilic material from biopsies from two patients with NOMID(152). As an iatrogenic amyloid, its aggregation is probably a result of increased local concentration of protein at the injection site.

Islet amyloid polypeptide (IAPP)

Islet amyloid polypeptide, or amylin, is a peptide hormone secreted by the beta cells of the pancreas. Its function is not entirely understood, but its main function seems to be regulation of

insulin activity(153). The amyloid aggregation of this protein is associated with type 2 diabetes. The formation of amyloid in type 2 diabetes had been noticed as early as 1900(154) and later confirmed(155), but only much later was the protein responsible for the amyloid deposits identified as a novel amyloid protein(156–158) through amino acid sequence analysis of congophilic material extracted from insulin-producing tumors and pancreas samples from patients with type 2 diabetes and also immunohistochemical analysis. The fiber structures of wild-type and mutant islet amyloid polypeptide reveal that the mutant fibers are not necessarily more stable than the wild-type(159) (although this conclusion is not universal(160)), so the amyloidogenic mutation (S20G) is hypothesized to act mainly through rearrangement of the monomer, so native structure destabilization was assigned as a mutation mechanism. Also, gene promoter mutations have been reported for IAPP which are associated with type 2 diabetes(161), so altered processing was also assigned as a mechanism.

Keratin-5 (KRT5)

Keratin-5, like other cytokeratin proteins, is a protein which forms heteropolymer intermediate filaments in epithelial tissue, and keratin-5 is found in the epidermis(162). This protein is associated with localized cutaneous amyloidosis, which is a type of amyloidosis with two types of presentations: primary (sometimes called lichen or macular amyloidosis) or secondary which is associated with skin neoplasms(163–165). Evidence of its amyloid nature mainly comes in the form of immunohistological studies of the keratin profiles in congophilic amyloid deposits of patients, the results of which always show positive staining for keratin-5, but variable staining for other keratins(163–165). It should be noted that keratin-5 normally interacts with keratin-14 and keratin-14 was detected immunohistochemically in some amyloid deposits, but was not as ubiquitous as keratin-5**.** Interestingly, mutations in keratin-5 lead to non-amyloid conditions, namely epidermolysis bullosa simplex (EBS) and Dowling-Degos disease(162, 166). However,

one mutation (V324A) was associated with a case of Weber-Cockayne type EBS presenting with cutaneous amyloidosis and the congophilic amyloid deposits stained positive for an anti-keratin antibody which reacts with keratin-1, -5, -10, and -14(167). There is not enough biochemical data on the amyloid nature of this protein in the wild-type or mutant state to determine an amyloidogenic mechanism for the V324A mutation.

Keratin-8 (KRT8)

Keratin-8, like other cytokeratin proteins, is a protein which forms heteropolymer intermediate filaments in hepatocytes(168). This protein is thought to form amyloids in the form of aggregates called Mallory-Denk bodies in the liver of patients with alcoholic steatohepatitis(169). Also, amyloidogenic mutations in keratin-8 are associated with cryptogenic liver disease(170). Keratin-8 was identified as an amyloid protein through a computational screen intending to identify mutations which modify a segment capable of reversible aggregation into a segment prone to irreversible aggregation(169). The amyloid nature of keratin-8 was confirmed through ThT fluorescence, x-ray fiber diffraction, and electron microscopy-based ultrastructural analysis of the head domain of wild-type and mutant keratin-8 as well as peptide crystal structures of wild-type and mutant segments of the protein. The head domain, where the amyloidogenic mutations are, is intrinsically disordered and the mutations significantly increase amyloidogenicity, based on the kinetics observed in ThT assays, and crystal structures of the mutants reveal stronger side chain interactions and more stable secondary structure. For these reasons, the amyloidogenic mechanism of fiber stabilization was the only one assigned to mutations in keratin-8.

Lactadherin (MFGE8)

Lactadherin is a glycoprotein which is secreted into milk and binds to milk-fat-globule membranes. It has a variety of functions, many related to immune response, such as playing a role in

phagocytosis, and other cellular functions like cell adhesion(171). A fragment of this protein spanning residues 245-294, called medin, is the main constituent of aortic medial amyloid. This was elucidated through amino acid sequence analysis of congophilic material from patient aortic tissue and immunohistochemistry(172). A synthetic octapeptide consisting of part of the medin sequence was also shown to form amyloid fibers *in vitro*(172). No familial mutations have been associated with amyloidosis lactadherin, and, in fact, aortic medial amyloid is found in the vast majority of individuals over 60 years old(172, 173). The health impact of these amyloid deposits is not fully understood.

Lactotransferrin (LTF)

Lactotransferrin, also called lactoferrin, is a glycoprotein found in secretory fluids, such as milk and saliva, and also in granulocytes. It has various immune functions, mainly as an antimicrobial agent, particularly by binding to free iron which is required for bacterial growth(174). Lactoferrin has been found in amyloid deposits in familial subepithelial corneal amyloidosis, also called gelatinous drop-like corneal dystrophy, (an hereditary corneal dystrophy similar to lattice corneal dystrophy which is caused by amyloidosis of gelsolin)(175), secondary corneal amyloidosis associated with trichiasis(176, 177), along with localized amyloidosis in various other organs (pancreas, bronchus, seminal vesicle)(178–180). It was first proposed to be an amyloid protein when congophilic material from localized amyloidosis of the seminal vesicle was positively immunostained with antibodies against lactotransferrin, and demonstration through electron microscopy that the amyloid fibrils themselves were being decorated by the antibodies(180). Soon after, it was shown to be present in amyloid deposits in gelatinous drop-like corneal dystrophy through Edman degradation amino acid sequence analysis of proteins extracted from congophilic corneal tissue and immunohistochemistry(175). The result for gelatinous drop-like corneal dystrophy was somewhat doubted, however, since mutations in proteins besides lactoferrin cause

this hereditary disease(176, 181), and the protein identified in that study was ostensibly wild-type. Although, mutations in one protein causing a disease in which a different protein forms amyloids is not uncommon, so this skepticism may be unwarranted. Variant lactoferrin (E579D) was found in amyloid deposits of patients with trichiasis-associated secondary corneal amyloidosis(176, 177), and wild-type lactotransferrin can only form amyloids *in vitro* under conditions which are far from physiological(176, 182), but the genetics of the patients with other forms of localized lactoferrin amyloidosis are unknown(175, 178–180). Although, the variant found in individuals with secondary corneal amyloidosis associated with trichiasis is also present in healthy individuals, although at lower frequencies(177), making this variant allele a polymorphism rather than a mutation. For this reason, this protein was classified as a sporadic amyloid. It should be noted, however, that all individuals studied with secondary corneal amyloidosis associated with trichiasis harbored this polymorphism. And although a polymorphism is not considered a mutation, an amyloidogenic mechanism for this variant allele has been proposed: the variant residue is hypothesized to disrupt a stabilizing hydrogen bond interaction and increase flexibility enough to expose a hydrophobic patch(177).

Leukocyte cell-derived chemotaxin-2 (LECT2)

Leukocyte cell-derived chemotaxin-2 is a protein with a wide variety of functions including chemotaxis, liver regeneration, immune modulation, bone growth, neuronal development, glucose metabolism, and more(183). This protein forms amyloid deposits in cases of renal amyloidosis and hepatic amyloidosis(184–187). It was first identified as an amyloid protein through Edman degradation sequence analysis of congophilic material from kidney tissue and immunohistochemistry(184). All patients with amyloidosis of this protein who have been genetically sequenced are homozygous, or very rarely heterozygous, for the same polymorphism coding for a valine at position 58 (40 in the mature protein) rather than an isoleucine(184–190).

Since this is a polymorphism and not a mutation, and thus the variant residue is seen in healthy individuals in the population, this protein was classified as a sporadic amyloid. However, it should be noted that the valine residue may destabilize the native structure relative to an isoleucine residue(187, 191), but the polymorphic valine residue was not resolved in a recombinant protein structure of the fibril core(192).

Lysozyme (LYZ)

Lysozyme is a bacteriolytic enzyme found in mucosal secretions(193). It forms amyloids in hereditary non-neuropathic systemic amyloidosis Ostertag type, now known as hereditary lysozyme amyloidosis(194). Lysozyme was first found to be the amyloid component in this disease through amino acid sequence analysis of protein extracted from congophilic amyloid deposits from a patient's kidney, and also immunohistochemistry(194). Lysozyme is only found to form amyloid deposits if it harbors one of the documented dominant hereditary mutations, and, in fact, wild-type lysozyme is not detectable in the amyloid deposits(195), making lysozyme a proto-amyloid. Although, at least one study has shown that in unphysiologically low pH conditions (even for lysosomes) wild-type lysozyme is destabilized and is able to form amyloid fibers(196). The initial report of lysozyme being the amyloid protein hypothesized that the mutation was destabilizing the native structure of the protein and later molecular dynamics simulations lead to the same conclusion for many of the known lysozyme mutations(197), so native structure destabilization was assigned as the mutation mechanism for this protein.

Major prion protein (PRNP)

Major prion protein glycosylphosphatidylinositol anchored membrane protein whose total suite of functions is not entirely clear, although it is a highly conserved protein in mammals and is known to have functions in cell signaling, neuritogenesis, neuronal homeostasis, and others(198). This

protein forms amyloids in humans in Creutzfeldt-Jakob disease (CJD) (familial, sporadic, and iatrogenic), fatal familial insomnia (FFI) (and sporadic fatal insomnia), Gertsmann–Sträussler–Scheinker (GSS), and Kuru(199, 200). The prion protein was first proposed to be the infectious agent of the animal disease scrapie by Prusiner in 1982(201) and was later hypothesized to also be the component of the amyloid deposits in this disease evidenced by Congo red staining and electron microscopy-based ultrastructural characterization of purified scrapie prion protein(202) as well as immunoelectron microscopy and immunohistochemistry on scrapie-infected brains(203). Prion protein was first shown to exist as an amyloid in humans through immunostaining of congophilic plaques in human brains with CJD and GSS(204). There are over 60 known pathogenic mutations in major prion protein(200, 205). Wild-type prion protein is able to undergo a transition from α-helical secondary structure to β-sheet secondary structure and form very stable fibers. Some mutant fibers have even been shown to be less stable than wild-type fibers(206). The amyloidogenic mechanism of mutations in this protein seem to mainly be destabilization of the native fold through disruption of important intramolecular interactions like salt bridges(199, 206, 207) (and truncating mutations certainly disrupt the native structure of the protein) or altered processing through destabilization of the native structure making the protein vulnerable to aberrant proteolytic processing, leading to production of amyloidogenic fragments of the protein(208–211).

Melanocyte protein PMEL (PMEL)

Melanocyte protein PMEL is a protein which, after extensive post-translational modification, forms functional amyloid fibers inside melanosomes. These amyloid fibers form a structural foundation for the organelle to store melanin pigments(212). Mutations in this protein in humans cause pigmentary dispersion syndrome (PDS), characterized by shedding of pigmented material from the iris, which can lead to pigmentary glaucoma (PG), which can lead to blindness(213). Many of

the PMEL variants linked to this disease have been shown to lead to formation of abnormal fibers, rather than abolishing fiber formation altogether, ostensibly forming a pathogenic amyloid rather than a functional amyloid(213, 214). This abnormal fiber formation is seen by electron microscopy of pseudomelanosomes formed in HeLa cells expressing PMEL variants. Western blot analysis of the lysate of these HeLa cells reveals defects in proteolytic processing and post-translational modification of the variant forms of the protein(213), so altered processing was assigned as the amyloidogenic mutation mechanism.

Microtubule-associated protein tau (MAPT)

Microtubule-associated protein tau is a neuronal protein which binds to and stabilizes microtubules(215), but it may have other biological roles as well such as RNA-binding(216). Tau protein is found in amyloid deposits in over 20 human diseases, collectively called tauopathies(217), and mutant forms of tau cause diseases with a wide variety of presentations collectively referred to by the umbrella term “frontotemporal dementia and parkinsonism linked to chromosome 17” (FTDP-17)(215, 218, 219), but at least one has been specifically identified as Pick’s disease (PiD) mutations(220) and some polymorphisms are risk factors for other tauopathies. Microtubule-associated protein tau was shown to be an amyloid protein when it was identified as the constituent protein of Alzheimer’s disease paired helical filaments (PHFs) through immunoblotting of tau with anti-microtubule antibodies cross-reactive for PHFs, and also immunostaining of Alzheimer’s tangles and plaque neurites with affinity-purified tau antibodies(221). Mutations in Microtubule-associated protein tau can have different amyloidogenic mechanisms from each other. Some mutations operate by altered processing through affecting the alternative splicing of MAPT, specifically exon 10 which contains the fourth tandem repeat of the four microtubule binding domain imperfect repeats(215, 219, 222–226). Mutations can either increase or decrease the inclusion of this exon in transcripts, but the ratio of

tau protein with four repeats to tau protein with three repeats seems to be tightly regulated and disruption of this ratio leads to amyloid aggregation. The mechanism is potentially related to limited binding availability of microtubules to certain isoforms of tau protein(215, 223). Other mutations disrupt binding to microtubules directly, releasing free tau protein to aggregate(215, 217–219, 227). However, there are examples of mutations which actually increase binding to microtubules, but this may encourage pathological hyperphosphorylation leading to aggregation(228). In either case, dysregulation of binding to the native binding partner leads to amyloid aggregation. There are also many mutations which have been shown to accelerate aggregation *in vitro* where the protein is ostensibly disordered(229), meaning these mutations must stabilize the fiber form in some way(215, 228, 230–234). Since the protein is intrinsically disordered when not bound to microtubules(229), native structure destabilization could not be assigned as a mutation mechanism. It should be noted that post-translational modifications like phosphorylation and acetylation seem to be important for the amyloid aggregation of tau protein(219, 235–238), however the relationship between mutations and these features is not clear.

Odontogenic ameloblast-associated protein (ODAM)

Odontogenic ameloblast-associated protein is a protein secreted by ameloblasts which plays a role in odontogenesis and is incorporated into the enamel matrix of mature enamel layers(239). This protein is found in the amyloid deposits associated with Calcifying epithelial odontogenic tumors (CEOTs), also known as Pindborg tumors(240, 241). The protein was first identified to be the constituent of the amyloid deposits by Edman degradation amino acid sequence analysis of amyloid material extracted from congophilic tumors, reverse transcription-PCR analysis of mRNA from tumor samples, and immunohistochemistry(240, 241). This protein does not have any associated amyloidogenic mutations, making it a sporadic amyloid.

Parathyroid hormone (PTH)

Parathyroid hormone is a hormone secreted by the parathyroid glands which regulates blood calcium levels(242). This protein makes up the amyloid deposits associated with parathyroid adenoma and parathyroid hyperplasia(243). Parathyroid hormone had been shown to be able to form amyloid fibers *in vitro*(244, 245) before its identification as the component of parathyroid tumors, but it was confirmed *in vivo* through mass spectrometry proteomic analysis of microdissected congophilic parathyroid adenoma samples and immunohistochemistry(243). Parathyroid hormone does not have any associated amyloidogenic mutations, making it a sporadic amyloid, but it should be noted that parathyroid adenomas are associated with elevated parathyroid hormone levels, and this increase in production of the amyloidogenic protein may be necessary for amyloidogenesis.

Polyadenylate-binding protein 2 (PABPN1)

Polyadenylate-binding protein 2 is a nuclear protein which stimulates the addition of poly(A) tails on mRNA(246). This protein forms fibrous nuclear aggregates in oculopharyngeal muscular dystrophy, confirmed by immunofluorescent labeling and immunoelectron microscopy(247), although these aggregates were not characterized as amyloids. This protein was later shown to be able to form fibers with amyloid characteristics *in vitro*, as evidenced by affinity to ThT and ultrastructural characterization by electron microscopy(248). This protein causes disease due to a trinucleotide expansion which extends a polyalanine region near the N-terminal of the protein(247, 248). It has been shown that the wild-type protein can form ThT-positive aggregates, but the polyalanine expansion greatly accelerates their formation(248), so this protein is classified as an ambimorph amyloid. Also, this protein can cause disease through a mutation which mimics the polyalanine expansion by substituting a glycine, which interrupts the polyalanine region of the

protein, for an alanine(249), and this presumably induces the disease through the same mechanism as the polyalanine expansion, which is formation of fibrillar nuclear aggregates. In N-terminal fragments, the polyalanine expansion seems to induce α-helical secondary structure in an otherwise unstructured region of the protein(248), but how this influences fiber formation is unclear, and since this region in the native protein is unstructured we did not consider this to fall under the mechanism of native structure destabilization. However, since this mutation accelerates the *in vitro* fiber formation of an otherwise unstructured protein, the mutation must be stabilizing the fiber form in some way, possibly through a capacity to transition from the induced α-helical secondary structure to β-sheet secondary structure, so fiber stabilization was assigned as the mutation mechanism.

Prolactin (PRL)

Prolactin is a hormone secreted by the pituitary gland with various physiological functions, but mainly promotion of milk production and development of mammary glands in breast tissue(250). Prolactin forms amyloid deposits in prolactin-producing pituitary adenomas and also tumor-free pituitary glands of individuals of advanced age(251, 252). Prolactin was identified to be the component of the amyloid fibers in both cases through amino acid sequence analysis of amyloid material extracted from congophilic deposits in pituitary gland samples. Interestingly, in both studies, commercial anti-prolactin antibodies were not reactive with the amyloid material. This is either due to aberrant proteolytic cleavage of prolactin (so the amyloid is composed of a fragment of the protein lacking the epitope recognized by the antibody) or the conformational change accompanying amyloid formation buries or alters the epitope recognized by the antibody. Prolactin amyloidogenesis is not associated with any mutations, so it is classified as a sporadic amyloid.

Protein TFG (TFG)

Protein TFG functions in the trafficking of secretory vesicles between the endoplasmic reticulum (ER) and ER-Golgi intermediate compartments(253). Mutations in protein TFG are associated with Charcot-Marie-Tooth disease type 2 and hereditary motor and sensory neuropathy with proximal dominant involvement(254, 255). The mutations associated with both these diseases have been shown to induce amyloid aggregation of the protein *in vitro* as evidenced by ThT fluorescence, x-ray fiber diffraction, and electron microscopy-based ultrastructural analysis(256). Wild-type recombinant protein was also able to form amyloid fibers *in vitro*, although at a slower rate, so this protein is classified as an ambimorph amyloid. The mutations occur in the disordered low-complexity domain of the protein and cryo-EM structures reveal that the mutant residues form key stabilizing interactions in the fiber core **(unpublished)**, so fiber stabilization is the only mutation mechanism assigned.

Pulmonary surfactant-associated protein C (SFTPC)

Pulmonary surfactant-associated protein C, also called lung surfactant protein C, is a transmembrane lipopeptide which functions to lower alveolar surface tension at the air-liquid interface(257). This protein forms amyloids in pulmonary alveolar proteinosis (PAP), confirmed by Edman degradation amino acid sequence analysis of congophilic amyloid material extracted from bronchoalveolar lavage (BAL) fluid from a PAP patient(258). No mutations are associated with the amyloid formation of this protein, making it a sporadic amyloid. Interestingly, this protein, despite being a 35-residue peptide, exists as a stable α-helix in lipid membranes, but transitions to β-sheet aggregates in solution(258, 259). This transition is dependent on removal of palmitoyl groups from the protein's cysteine residues, and this modification along with increased levels of the protein seem to strongly contribute to its amyloid conversion, although the cause of the protein's depalmitoylation is unknown.

RNA-binding protein FUS (FUS)

RNA-binding protein FUS is a nuclear protein involved in transcription and DNA repair, but also forms cytosolic stress granules through liquid-liquid phase separation(260, 261). In stress granules, this protein forms reversible aggregates consisting of fibers with amyloid qualities, demonstrated *in vitro* through electron microscopy-based ultrastructural analysis and x-ray fiber diffraction(262, 263). This protein is also found in cytoplasmic inclusions in diseases including frontotemporal lobar degeneration (FTLD-FUS) and amyotrophic lateral sclerosis (ALS)(260, 264–266). The low-complexity domain of RNA-binding protein FUS has been shown to form reversible, liquid-like aggregates *in vitro* which transition to solid, irreversible, cytotoxic amyloid fibers over time(267), and disease mutations have been shown to accelerate this transition(268). Since these mutations promote aggregation of an otherwise disordered region of the protein, fiber stabilization was assigned as the mechanism. Other mutations seem to not directly increase aggregation propensity(265), but rather contribute to cytoplasmic mislocalization of the protein, which contributes to its pathological aggregation, so subcellular mislocalization was also assigned as a mechanism.

S100-A8/A9 (S100A8/A9)

S100-A8 and S100-A9, also known as calgranulin-A and calgranulin-B, respectively, are calcium and zinc binding proteins which have various biological roles including pro-inflammatory roles and acting as an antifungal agent(269). These proteins can form homodimers and heterodimers, but also a heterotetrameric form called calprotectin. These proteins were found in congophilic corpora amylacea, a type of extracellular inclusion found in various tissues, from prostate tissue extracted from patients with prostate cancer(270), although corpora amylacea can exist in noncancerous aged prostate as well. S100-A8 and S100-A9 were identified as the amyloid component by mass

spectrometry analysis and immunostaining, along with atomic force microscopy analysis of the amyloid material. We will note, however, that the *ex vivo* fibers and those generated *in vitro* in this work do not, in our view, necessarily look like typical amyloid fibers which are explicitly unbranched. The histology and mass spectrometry analysis, however, provide evidence consistent with other established amyloid proteins. It is also unclear which form the proteins take in these aggregates: homopolymers, polymers of the heterodimers, or polymers of calprotectin (a heterotetramer). Both proteins are evidently present in the amyloid aggregates, but the segments of the proteins which are predicted to be most aggregation-prone also seem to be involved in their native oligomerization, so it may be the case that disruption of the oligomeric states of these proteins leads to their amyloid aggregation. If this is the case, it is unlikely that the amyloid fibers are heteropolymers, but whether or not this is the case is not clear from the evidence. These proteins have no mutations associated with their amyloid aggregation, making them a sporadic amyloid, but their aggregation may be linked to increased local concentration due to chronic inflammation.

Semenogelin 1 (SEMG1)

Semenogelin 1 is the main protein component of human semen and promotes sperm survival, motility, and fertility(271). This protein forms amyloid deposits in senile seminal vesicle amyloid, a localized amyloidosis associated with male aging. Semenogelin 1 was confirmed as the amyloidogenic protein through mass spectrometry analysis of congophilic material from seminal vesicle samples with amyloid and immunohistochemistry(272). No mutations are associated with the amyloid formation of this protein, so it is a sporadic amyloid. Of note, the amyloid component of this protein seems to be an N-terminal fragment of semenogelin 1.

Serum amyloid A (SAA1)

Serum amyloid A is an acute-phase response protein which is secreted by the liver into the blood in response to inflammatory conditions(273). This protein is found in amyloid deposits in individuals with amyloid A amyloidosis, a systemic secondary amyloidosis resulting from chronic inflammation(274). This disease can also manifest as a primary amyloidosis due to a mutation in the SAA1 promoter region inducing overexpression(275), making this protein an ambimorph amyloid. This protein was first identified as a unique amyloid protein through amino acid sequence analysis of congophilic amyloid material from livers and spleens of patients with secondary amyloidosis associated with familial Mediterranean fever, tuberculosis, Hodgkin's lymphoma, and bronchiectasis(276), and this result was corroborated by other groups later(277, 278). Since the mutation associated with primary amyloid A amyloidosis is a promoter mutation which causes overexpression, altered processing was assigned as the mutation mechanism.

Somatostatin (SST)

Somatostatin is a pancreatic prohormone which is cleaved into two small peptide hormones, somatostatin-14 and somatostatin-28, which regulate the production of pituitary hormones(279). One of these peptide hormones, somatostatin-14, was shown to form amyloid fibers *in vitro* as evidenced by Congo red staining, electron microscopy-based ultrastructural analysis, and x-ray fiber diffraction(280). Somatostatin was found in amyloid deposits in somatostatin-producing neuroendocrine tumors (somatostatinomas)(281, 282). Somatostatin was confirmed as the amyloid protein through mass spectrometry analysis of microdissected congophilic tissue and immunohistochemistry. Interestingly, somatostatin-28 was the major species present in the *in vivo* amyloid deposits, not somatostatin-14, based on immunostaining results. Amyloidosis of somatostatin is not associated with any mutations, making it a sporadic amyloid. However, it should be noted that the amyloid formation in the case of neuroendocrine tumors may be reliant on the increased production of the protein by the tumor.

Superoxide dismutase (SOD1)

Superoxide dismutase is a metalloenzyme which catalyzes a dismutation reaction of superoxide radicals into O_2 and H_2O_2(283). Superoxide dismutase is found in pathological inclusions in both familial and sporadic ALS(284, 285). Mutant superoxide dismutase has been shown to form fibers in transgenic mice with amyloid characteristics as evidenced by immunoelectron microscopy and thioflavin S staining(286, 287). Wild type and mutant recombinant protein can be induced to form thioflavin T-positive, fibrillar aggregates (amyloids) under reducing conditions *in vitro*, mimicking the reducing environment of the cell(288, 289). There are over 200 documented mutations in superoxide dismutase which are linked to familial ALS (https://alsod.ac.uk/output/gene.php/SOD1), although not all of them are necessarily amyloidogenic. These mutations likely induce aggregation by interrupting its ability to bind metal ions, since the mature, metal-bound protein is very resistant to aggregation and inclusions in transgenic mice and cell lines expressing mutant superoxide dismutase contain metal-deficient, disulfide reduced protein(284, 287, 289, 290). Since mutations in this protein likely disrupt an important binding site, the mutation mechanisms assigned are native structure destabilization and decreased binding to native partners.

TAR DNA-binding protein 43 (TARDBP)

TAR DNA-binding protein 43 is a nuclear DNA- and RNA-binding protein with roles in regulating transcription of RNA(260). It can also be found in the cytoplasm as a constituent of pathological inclusions in FTLD-TDP and ALS(260, 291–295). Inclusions in the brains of individuals with FTLD-TDP have been shown to bind thioflavin S and also be immunoreactive to antibodies against TAR DNA-binding protein 43(296). Immunoelectron microscopy studies also reveal this protein exists in the form of filamentous inclusions in tissue samples from a variety of neurodegenerative

diseases(297). It has also been shown that fibrillar aggregates of TAR DNA-binding protein 43 from patient brains can act as seeds which induce aggregation in cultured cell lines(298). The structure of the amyloid fiber has also been solved from material extracted from the brain of a patient with ALS with FTLD(299). Disease mutations in this protein concentrate in the low-complexity C-terminal region, which is disordered(260), and many of these mutations have been shown to accelerate aggregation(291, 300), so the mutation mechanism of fiber stabilization was assigned. Since aggregates are mislocalized to the cytoplasm and this mislocalization can be enhanced by mutations(301), subcellular mislocalization was also assigned as a mechanism. There are also noncoding variants for this protein associated with ALS(302), but many of them are polymorphisms, not mutations, and their effects on protein production are not entirely clear, so altered processing was not included as a mutation mechanism for this protein.

Transcription elongation regulator 1 (TCERG1 or CA150)

Transcription elongation regulator 1, or CA150, is a transcription factor which codeposits with huntingtin aggregates in Huntington's disease and seems to be a modifier of the age of onset(303). CA150 rapidly forms amyloid fibers *in vitro* as evidenced by thioflavin T fluorescence, light scattering, electron microscopy-based ultrastructural analysis, optical diffraction of fibers, and solid state NMR structure determination(304, 305). There are no pathogenic mutations associated with CA150, making it a sporadic amyloid.

Transforming growth factor-beta-induced protein ig-h3 (TGFBI)

Transforming growth factor-beta-induced protein ig-h3, also called kerato-epithelin, is an extracellular matrix protein which is abundant in the corneal stroma(306). It is found in amyloid deposits in lattice corneal dystrophy, but also non-amyloid aggregates in other corneal dystrophies(307–311). The contribution of kerato-epithelin to amyloid deposits was confirmed by

immunohistochemical staining of congophilic deposits in patient corneal tissue(311). Peptide fragments of the protein were also shown to be able to form amyloid fibers *in vitro* evidenced by circular dichroism spectra, thioflavin T fluorescence, and electron microscopy-based ultrastructural analysis(308, 312). Mutations in this protein which result in amyloid deposition seem to lead to deposition of unique proteolytic fragments of the protein not found in wild-type corneas, specifically from the fourth fasciclin-1 domain (FAS1-4)(313–315), and for this reason the mutation mechanism assigned to this protein is altered processing. The wild-type form of this protein does not seem to develop into amyloid fibers, so kerato-epithelin is classified as a proto-amyloid.

Transmembrane protein 106B (TMEM106B)

Transmembrane protein 106B is a transmembrane glycoprotein which localizes to the membranes of lysosomes and interacts with progranulin(316). This protein is found as an amyloid fiber in a wide variety of neurodegenerative diseases including FTLD-TDP, progressive supranuclear palsy (PSP), dementia with lewy bodies (DLB), Alzheimer's disease, corticobasal degeneration, FTDP-17, argyrophilic grain disease, Parkinson's disease, limbic-predominant neuronal inclusion body four-repeat tauopathy, aging-related tau astrogliopathy, MSA, and ALS(317, 318). Although, the connection of the protein's amyloidogenesis to disease is unclear because it can also be found in fibrillar form in healthy, aged brains(319). It was independently shown to be an amyloid protein by three separate groups at the same time by the same method: cryo-EM structure determination of brain-extracted amyloid fibers(317, 318, 320).There are no familial mutations associated with transmembrane protein 106B, making it a sporadic amyloid, but there is a polymorphism at residue 185 (threonine or serine)(321). This polymorphism may influence expression of the protein(316) and having a serine at this position is hypothesized to be

protective against disease due to more rapid degradation of the protein with the serine polymorph(316, 322).

Transthyretin (TTR)

Transthyretin, also called prealbumin, is a thyroid hormone distributor protein which is secreted into the blood by the liver and into the cerebro-spinal fluid (CSF) by the epithelial cells of the choroid plexus(323). Transthyretin functions as a tetramer and binds to thyroid hormones and retinol-binding protein, as well as certain drugs and pollutants(323). This protein forms amyloid deposits in systemic transthyretin amyloidosis, which is characterized by amyloid deposition in multiple organs and commonly manifests as cardiomyopathy and/or polyneuropathy(324, 325). Transthyretin was first identified as an amyloid protein by matching the immunoreactivity of antisera raised against prealbumin to an antisera raised against amyloid fiber protein extracted from the kidneys of patients with familial amyloidotic polyneuropathy(326). This result was later repeated in other cases of familial amyloidosis as well as sporadic amyloidosis(327). There are over 120 amyloidogenic mutations in transthyretin, but the most common one is V50M (V30M with the numbering of the mature protein)(324, 328). Comparison of wild-type and mutant structures of transthyretin amyloid fibers(329, 330) reveals that they have nearly identical structures, meaning the mutation mechanism is solely disruption of the native tetramer, so native structure destabilization is the only assigned mutation mechanism.

Ubiquilin-2 (UBQLN2)

Ubiquilin-2 is a protein which interacts with ubiquitinated proteins and delivers them to the proteasome for degradation(331). Intracellular aggregates of this protein are found in various neurodegenerative diseases including ALS, synucleinopathies, and polyglutamine diseases(332–336), and mutations in its proline-rich repeat (PXX) domain cause dominant X-linked ALS and

FTD. Ubiquilin-2 was shown to form amyloids in vitro through ThT fluorescence and electron microscopy-based ultrastructural analysis(332). In regard to the full-length protein, only the construct with the P506T mutation was able to form amyloid fibers and the wild-type could not. For this reason, we have classified this protein as a proto-amyloid. However, it has not escaped our attention that the wild-type protein with its N-terminal ubiquitin-like (UBL) domain (residues 1-106) deleted was able to form amyloid fibers. Also, the protein's C-terminal ubiquitin-associated (UBA) domain alone (residues 575-624) was able to form amyloid fibers and the P506T ubiquilin-2 had reduced fiber formation with this region deleted. The UBA domain is responsible for binding to ubiquitinated substrates, and disruption of this binding capability has been shown to lead to aggregation in a cell model(332). The P506T mutation has also been shown to increase cellular aggregation, so it likely disrupts the UBA domain's binding activity, so the mutation mechanism of decreased binding to native partner was assigned. However, differences between the aggregation behavior of wild-type and mutant protein in vitro (in the absence of binding partners) cannot be explained by this mechanism. The PXX domain is intrinsically disordered(337), and proline residues discourage β-strand formation because of the geometry of their peptide bonds and discourage amyloid fiber formation by reducing the capacity for interstrand backbone hydrogen bonding. Point mutations away from proline may be sufficient to permit the PXX domain to be incorporated into the core of an amyloid fiber, although this is only speculation. Nevertheless, for this reason we also assigned fiber stabilization as a mutation mechanism.

α-synuclein (SNCA)

α-synuclein is a protein whose function is not entirely clear, but localizes to presynaptic terminals and interacts with lipid membranes, i.e. vesicles, and is able to adopt an α-helical secondary structure when associated with membranes despite being disordered in solution(338). This protein is the main component of Lewy bodies and Lewy neurites which are the hallmarks of

Parkinson's disease (PD) and dementia with Lewy bodies (DLB)(339) and also aggregates in multiple system atrophy (MSA) as well as other "synucleinopathies". Both wild-type and mutant α-synuclein has been observed to form amyloid fibers *in vitro*, making it an ambimorph amyloid, as evidenced by circular dichroism spectrometry, thioflavin T fluorescence, electron microscopy and atomic force microscopy ultrastructural analysis, immunoelectron microscopy, x-ray and electron fiber diffraction, and cryo-EM structure determinations(340–349). α-synuclein was confirmed to exist as an amyloid fiber *in vivo* in Lewy bodies through microbeam X-ray diffraction of thin sections of Parkinson's disease brain samples(350). Many mutations in this protein have been shown to accelerate fiber formation *in vitro* and form more stable fibers than the wild-type protein(342, 343, 351), and since the protein is disordered in solution(352, 353), only the mutation mechanism of fiber stabilization could be assigned for these. α-synuclein has also been shown to have increased seeding capacity when it has certain mutations, so increased seeding was also assigned(346, 347, 354). There are also duplications and triplications of the SNCA gene which increases expression leading to disease(355, 356), so altered processing was also assigned as a mechanism. Lastly, some mutations have been shown to reduce binding to lipid membranes(340, 357) so reduced binding to native partners was also assigned as a mechanism.

β2-microglobulin (B2M)

β2-microglobulin is a component of the class 1 major histocompatibility complex and, when it dissociates from the complex, is cleared from the body by the kidneys(358). Patients with renal failure requiring dialysis can build up β2-microglobulin in their blood because dialysis machines are not able to clear it efficiently(359, 360), although this problem has been mitigated significantly (but not entirely) by modern machinery(360, 361). This can lead to the formation of amyloid fibers by β2-microglobulin in a condition known as dialysis-related amyloidosis which manifests as

painful bone and joint-related ailments like carpal tunnel syndrome and arthritis(359, 362). β2-microglobulin was first identified as an amyloidogenic protein through amino acid sequence analysis of congophilic amyloid material extracted during a carpal tunnel release operation on a patient who had been on dialysis for 13 years(362). There is also a documented mutation in this protein associated with an hereditary systemic amyloidosis: D96N (D76N with the numbering of the mature protein)(363). This mutation alters the surface charge landscape of the protein, as revealed by the crystal structure of the mutant protein, and reduces its denaturation resistance in guanidine hydrochloride(363), so native structure destabilization was assigned as a mechanism. The mutation also greatly increased the aggregation propensity of the protein compared to the wild-type under physiological conditions. The destabilization of the monomer by the mutation does not seem drastic enough to account for this result, so fiber stabilization was also assigned as a mechanism.

References

1. Brothers, H. M., Gosztyla, M. L., and Robinson, S. R. (2018) The Physiological Roles of Amyloid-β Peptide Hint at New Ways to Treat Alzheimer's Disease. *Front Aging Neurosci*. **10**, 118
2. Nilsberth, C., Westlind-Danielsson, A., Eckman, C. B., Condron, M. M., Axelman, K., Forsell, C., Stenh, C., Luthman, J., Teplow, D. B., Younkin, S. G., Näslund, J., and Lannfelt, L. (2001) The "Arctic" APP mutation (E693G) causes Alzheimer's disease by enhanced Abeta protofibril formation. *Nat Neurosci*. **4**, 887–893
3. Masters, C. L., Simms, G., Weinman, N. A., Multhaup, G., McDonald, B. L., and Beyreuther, K. (1985) Amyloid plaque core protein in Alzheimer disease and Down syndrome. *Proc Natl Acad Sci U S A*. **82**, 4245–4249

4. Melchor, J. P., McVoy, L., and Van Nostrand, W. E. (2000) Charge alterations of E22 enhance the pathogenic properties of the amyloid beta-protein. *J Neurochem*. **74**, 2209–2212

5. Van Nostrand, W. E., Melchor, J. P., Cho, H. S., Greenberg, S. M., and Rebeck, G. W. (2001) Pathogenic effects of D23N Iowa mutant amyloid beta -protein. *J Biol Chem*. **276**, 32860–32866

6. Glenner, G. G., and Wong, C. W. (1984) Alzheimer's disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein. *Biochem Biophys Res Commun*. **120**, 885–890

7. Kang, J., Lemaire, H. G., Unterbeck, A., Salbaum, J. M., Masters, C. L., Grzeschik, K. H., Multhaup, G., Beyreuther, K., and Müller-Hill, B. (1987) The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface receptor. *Nature*. **325**, 733–736

8. Selkoe, D. J. (1999) Translating cell biology into therapeutic advances in Alzheimer's disease. *Nature*. **399**, A23-31

9. Haass, C., Hung, A. Y., Selkoe, D. J., and Teplow, D. B. (1994) Mutations associated with a locus for familial Alzheimer's disease result in alternative processing of amyloid beta-protein precursor. *J Biol Chem*. **269**, 17741–17748

10. Watson, D. J., Selkoe, D. J., and Teplow, D. B. (1999) Effects of the amyloid precursor protein Glu693-->Gln "Dutch" mutation on the production and stability of amyloid beta-protein. *Biochem J*. **340**, 703–709

11. Mullan, M., Crawford, F., Axelman, K., Houlden, H., Lilius, L., Winblad, B., and Lannfelt, L. (1992) A pathogenic mutation for probable Alzheimer's disease in the APP gene at the N-terminus of beta-amyloid. *Nat Genet*. **1**, 345–347

12. Citron, M., Vigo-Pelfrey, C., Teplow, D. B., Miller, C., Schenk, D., Johnston, J., Winblad, B., Venizelos, N., Lannfelt, L., and Selkoe, D. J. (1994) Excessive production of amyloid

beta-protein by peripheral cells of symptomatic and presymptomatic patients carrying the Swedish familial Alzheimer disease mutation. *Proc Natl Acad Sci U S A*. **91**, 11993–11997

13. Hardy, J. (1997) Amyloid, the presenilins and Alzheimer's disease. *Trends Neurosci*. **20**, 154–159
14. Schütz, A. K., Vagt, T., Huber, M., Ovchinnikova, O. Y., Cadalbert, R., Wall, J., Güntert, P., Böckmann, A., Glockshuber, R., and Meier, B. H. (2015) Atomic-resolution three-dimensional structure of amyloid β fibrils bearing the Osaka mutation. *Angew Chem Int Ed Engl*. **54**, 331–335
15. Yang, Y., Zhang, W., Murzin, A. G., Schweighauser, M., Huang, M., Lövestam, S., Peak-Chew, S. Y., Saito, T., Saido, T. C., Macdonald, J., Lavenir, I., Ghetti, B., Graff, C., Kumar, A., Nordberg, A., Goedert, M., and Scheres, S. H. W. (2023) Cryo-EM structures of amyloid-β filaments with the Arctic mutation (E22G) from human and mouse brains. *Acta Neuropathol*. **145**, 325–333
16. Arciello, A., Piccoli, R., and Monti, D. M. (2016) Apolipoprotein A-I: the dual face of a protein. *FEBS Lett*. **590**, 4171–4179
17. Nichols, W. C., Dwulet, F. E., Liepnieks, J., and Benson, M. D. (1988) Variant apolipoprotein AI as a major constituent of a human hereditary amyloid. *Biochem Biophys Res Commun*. **156**, 762–768
18. Westermark, P., Mucchiano, G., Marthin, T., Johnson, K. H., and Sletten, K. (1995) Apolipoprotein A1-derived amyloid in human aortic atherosclerotic plaques. *Am J Pathol*. **147**, 1186–1192
19. 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

20. Lagerstedt, J. O., Cavigiolio, G., Roberts, L. M., Hong, H.-S., Jin, L.-W., Fitzgerald, P. G., Oda, M. N., and Voss, J. C. (2007) Mapping the structural transition in an amyloidogenic apolipoprotein A-I. *Biochemistry*. **46**, 9693–9699

21. Raimondi, S., Guglielmi, F., Giorgetti, S., Di Gaetano, S., Arciello, A., Monti, D. M., Relini, A., Nichino, D., Doglia, S. M., Natalello, A., Pucci, P., Mangione, P., Obici, L., Merlini, G., Stoppini, M., Robustelli, P., Tartaglia, G. G., Vendruscolo, M., Dobson, C. M., Piccoli, R., and Bellotti, V. (2011) Effects of the known pathogenic mutations on the aggregation pathway of the amyloidogenic peptide of apolipoprotein A-I. *J Mol Biol*. **407**, 465–476

22. Brewer, H. B., Lux, S. E., Ronan, R., and John, K. M. (1972) Amino Acid Sequence of Human apoLp-Gln-II (apoA-II), an Apolipoprotein Isolated from the High-Density Lipoprotein Complex. *Proc Natl Acad Sci U S A*. **69**, 1304–1308

23. Benson, M. D., Liepnieks, J. J., Yazaki, M., Yamashita, T., Hamidi Asl, K., Guenther, B., and Kluve-Beckerman, B. (2001) A new human hereditary amyloidosis: the result of a stop-codon mutation in the apolipoprotein AII gene. *Genomics*. **72**, 272–277

24. Prokaeva, T., Akar, H., Spencer, B., Havasi, A., Cui, H., O'Hara, C. J., Gursky, O., Leszyk, J., Steffen, M., Browning, S., Rosenberg, A., and Connors, L. H. (2017) Hereditary Renal Amyloidosis Associated With a Novel Apolipoprotein A-II Variant. *Kidney Int Rep*. **2**, 1223–1232

25. Gursky, O. (2014) Hot spots in apolipoprotein A-II misfolding and amyloidosis in mice and men. *FEBS Lett*. **588**, 845–850

26. De Gracia, R., Fernández, E. J., Riñón, C., Selgas, R., and Garcia-Bustos, J. (2006) Hereditary renal amyloidosis associated with a novel mutation in the apolipoprotein AII gene. *QJM*. **99**, 274

27. Yazaki, M., Liepnieks, J. J., Yamashita, T., Guenther, B., Skinner, M., and Benson, M. D. (2001) Renal amyloidosis caused by a novel stop-codon mutation in the apolipoprotein A-

II gene. *Kidney Int*. **60**, 1658–1665

28. Yazaki, M., Liepnieks, J. J., Barats, M. S., Cohen, A. H., and Benson, M. D. (2003) Hereditary systemic amyloidosis associated with a new apolipoprotein AII stop codon mutation Stop78Arg. *Kidney Int*. **64**, 11–16
29. Qu, J., Ko, C.-W., Tso, P., and Bhargava, A. (2019) Apolipoprotein A-IV: A Multifunctional Protein Involved in Protection against Atherosclerosis and Diabetes. *Cells*. **8**, 319
30. Bergström, J., Murphy, C., Eulitz, M., Weiss, D. T., Westermark, G. T., Solomon, A., and Westermark, P. (2001) Codeposition of apolipoprotein A-IV and transthyretin in senile systemic (ATTR) amyloidosis. *Biochem Biophys Res Commun*. **285**, 903–908
31. Canetti, D., Nocerino, P., Rendell, N. B., Botcher, N., Gilbertson, J. A., Blanco, A., Rowczenio, D., Morelli, A., Mangione, P. P., Corazza, A., Verona, G., Giorgetti, S., Marchese, L., Westermark, P., Hawkins, P. N., Gillmore, J. D., Bellotti, V., and Taylor, G. W. (2021) Clinical ApoA-IV amyloid is associated with fibrillogenic signal sequence. *J Pathol*. **255**, 311–318
32. Wolska, A., Dunbar, R. L., Freeman, L. A., Ueda, M., Amar, M. J., Sviridov, D. O., and Remaley, A. T. (2017) Apolipoprotein C-II: New findings related to genetics, biochemistry, and role in triglyceride metabolism. *Atherosclerosis*. **267**, 49–60
33. Nasr, S. H., Dasari, S., Hasadsri, L., Theis, J. D., Vrana, J. A., Gertz, M. A., Muppa, P., Zimmermann, M. T., Grogg, K. L., Dispenzieri, A., Sethi, S., Highsmith, W. E., Merlini, G., Leung, N., and Kurtin, P. J. (2017) Novel Type of Renal Amyloidosis Derived from Apolipoprotein-CII. *J Am Soc Nephrol*. **28**, 439–445
34. Hatters, D. M., MacPhee, C. E., Lawrence, L. J., Sawyer, W. H., and Howlett, G. J. (2000) Human apolipoprotein C-II forms twisted amyloid ribbons and closed loops. *Biochemistry*. **39**, 8276–8283
35. Sethi, S., Dasari, S., Plaisier, E., Ronco, P., Nasr, S. H., Brocheriou, I., Theis, J. D.,

Vrana, J. A., Zimmermann, M. T., Quint, P. S., McPhail, E. D., and Kurtin, P. J. (2018) Apolipoprotein CII Amyloidosis Associated With p.Lys41Thr Mutation. *Kidney Int Rep*. **3**, 1193–1201

36. Kohan, A. B. (2015) Apolipoprotein C-III: a potent modulator of hypertriglyceridemia and cardiovascular disease. *Curr Opin Endocrinol Diabetes Obes*. **22**, 119–125
37. Valleix, S., Verona, G., Jourde-Chiche, N., Nédelec, B., Mangione, P. P., Bridoux, F., Mangé, A., Dogan, A., Goujon, J.-M., Lhomme, M., Dauteuille, C., Chabert, M., Porcari, R., Waudby, C. A., Relini, A., Talmud, P. J., Kovrov, O., Olivecrona, G., Stoppini, M., Christodoulou, J., Hawkins, P. N., Grateau, G., Delpech, M., Kontush, A., Gillmore, J. D., Kalopissis, A. D., and Bellotti, V. (2016) D25V apolipoprotein C-III variant causes dominant hereditary systemic amyloidosis and confers cardiovascular protective lipoprotein profile. *Nat Commun*. **7**, 10353
38. de Messieres, M., Huang, R. K., He, Y., and Lee, J. C. (2014) Amyloid Triangles, Squares, and Loops of Apolipoprotein C-III. *Biochemistry*. **53**, 3261–3263
39. Maack, T. (1996) Role of atrial natriuretic factor in volume control. *Kidney Int*. **49**, 1732–1737
40. Song, W., Wang, H., and Wu, Q. (2015) Atrial Natriuretic Peptide in Cardiovascular Biology and Disease (NPPA). *Gene*. **569**, 1–6
41. Kaye, G. C., Butler, M. G., D'Ardenne, A. J., Edmondson, S. J., Camm, A. J., and Slavin, G. (1986) Identification of immunoreactive atrial natriuretic peptide in atrial amyloid. *J Clin Pathol*. **39**, 581–582
42. Podduturi, V., Armstrong, D. R., Hitchcock, M. A., Roberts, W. C., and Guileyardo, J. M. (2013) Isolated atrial amyloidosis and the importance of molecular classification. *Proc (Bayl Univ Med Cent)*. **26**, 387–389
43. Pucci, A., Wharton, J., Arbustini, E., Grasso, M., Diegoli, M., Needleman, P., Viganò, M.,

and Polak, J. M. (1991) Atrial amyloid deposits in the failing human heart display both atrial and brain natriuretic peptide-like immunoreactivity. *J Pathol*. **165**, 235–241

44. DeJesus-Hernandez, M., Mackenzie, I. R., Boeve, B. F., Boxer, A. L., Baker, M., Rutherford, N. J., Nicholson, A. M., Finch, N. A., Flynn, H., Adamson, J., Kouri, N., Wojtas, A., Sengdy, P., Hsiung, G.-Y. R., Karydas, A., Seeley, W. W., Josephs, K. A., Coppola, G., Geschwind, D. H., Wszolek, Z. K., Feldman, H., Knopman, D. S., Petersen, R. C., Miller, B. L., Dickson, D. W., Boylan, K. B., Graff-Radford, N. R., and Rademakers, R. (2011) Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron*. **72**, 245–256

45. Renton, A. E., Majounie, E., Waite, A., Simón-Sánchez, J., Rollinson, S., Gibbs, J. R., Schymick, J. C., Laaksovirta, H., van Swieten, J. C., Myllykangas, L., Kalimo, H., Paetau, A., Abramzon, Y., Remes, A. M., Kaganovich, A., Scholz, S. W., Duckworth, J., Ding, J., Harmer, D. W., Hernandez, D. G., Johnson, J. O., Mok, K., Ryten, M., Trabzuni, D., Guerreiro, R. J., Orrell, R. W., Neal, J., Murray, A., Pearson, J., Jansen, I. E., Sondervan, D., Seelaar, H., Blake, D., Young, K., Halliwell, N., Callister, J. B., Toulson, G., Richardson, A., Gerhard, A., Snowden, J., Mann, D., Neary, D., Nalls, M. A., Peuralinna, T., Jansson, L., Isoviita, V.-M., Kaivorinne, A.-L., Hölttä-Vuori, M., Ikonen, E., Sulkava, R., Benatar, M., Wuu, J., Chiò, A., Restagno, G., Borghero, G., Sabatelli, M., ITALSGEN Consortium, Heckerman, D., Rogaeva, E., Zinman, L., Rothstein, J. D., Sendtner, M., Drepper, C., Eichler, E. E., Alkan, C., Abdullaev, Z., Pack, S. D., Dutra, A., Pak, E., Hardy, J., Singleton, A., Williams, N. M., Heutink, P., Pickering-Brown, S., Morris, H. R., Tienari, P. J., and Traynor, B. J. (2011) A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron*. **72**, 257–268

46. Zu, T., Gibbens, B., Doty, N. S., Gomes-Pereira, M., Huguet, A., Stone, M. D., Margolis, J., Peterson, M., Markowski, T. W., Ingram, M. A. C., Nan, Z., Forster, C., Low, W. C.,

Schoser, B., Somia, N. V., Clark, H. B., Schmechel, S., Bitterman, P. B., Gourdon, G., Swanson, M. S., Moseley, M., and Ranum, L. P. W. (2011) Non-ATG-initiated translation directed by microsatellite expansions. *Proc Natl Acad Sci U S A*. **108**, 260–265

47. Mori, K., Arzberger, T., Grässer, F. A., Gijselinck, I., May, S., Rentzsch, K., Weng, S.-M., Schludi, M. H., van der Zee, J., Cruts, M., Van Broeckhoven, C., Kremmer, E., Kretzschmar, H. A., Haass, C., and Edbauer, D. (2013) Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathol*. **126**, 881–893
48. Flores, B. N., Dulchavsky, M. E., Krans, A., Sawaya, M. R., Paulson, H. L., Todd, P. K., Barmada, S. J., and Ivanova, M. I. (2016) Distinct C9orf72-Associated Dipeptide Repeat Structures Correlate with Neuronal Toxicity. *PLoS One*. **11**, e0165084
49. Chang, Y.-J., Jeng, U.-S., Chiang, Y.-L., Hwang, I.-S., and Chen, Y.-R. (2016) The Glycine-Alanine Dipeptide Repeat from C9orf72 Hexanucleotide Expansions Forms Toxic Amyloids Possessing Cell-to-Cell Transmission Properties. *J Biol Chem*. **291**, 4903–4911
50. Brasseur, L., Coens, A., Waeytens, J., Melki, R., and Bousset, L. (2020) Dipeptide repeat derived from C9orf72 hexanucleotide expansions forms amyloids or natively unfolded structures in vitro. *Biochem Biophys Res Commun*. **526**, 410–416
51. Felsenfeld, A. J., and Levine, B. S. (2015) Calcitonin, the forgotten hormone: does it deserve to be forgotten? *Clin Kidney J*. **8**, 180–187
52. Khurana, R., Agarwal, A., Bajpai, V. K., Verma, N., Sharma, A. K., Gupta, R. P., and Madhusudan, K. P. (2004) Unraveling the amyloid associated with human medullary thyroid carcinoma. *Endocrinology*. **145**, 5465–5470
53. Koopman, T., Niedlich-den Herder, C., Stegeman, C. A., Links, T. P., Bijzet, J., Hazenberg, B. P. C., and Diepstra, A. (2017) Kidney Involvement in Systemic Calcitonin Amyloidosis Associated With Medullary Thyroid Carcinoma. *Am J Kidney Dis*. **69**, 546–

549

54. Tan, Y., Li, D.-Y., Ma, T.-T., Xu, R., Zhou, F., Wang, S.-X., and Zhao, M.-H. (2020) Renal calcitonin amyloidosis in a patient with disseminated medullary thyroid carcinoma. *Amyloid*. **27**, 213–214

55. Sletten, K., Westermark, P., and Natvig, J. B. (1976) Characterization of amyloid fibril proteins from medullary carcinoma of the thyroid. *J Exp Med*. **143**, 993–998

56. Butler, M., and Khan, S. (1986) Immunoreactive calcitonin in amyloid fibrils of medullary carcinoma of the thyroid gland. An immunogold staining technique. *Arch Pathol Lab Med*. **110**, 647–649

57. Dai, R., Wu, Z., Chu, H. Y., Lu, J., Lyu, A., Liu, J., and Zhang, G. (2020) Cathepsin K: The Action in and Beyond Bone. *Front Cell Dev Biol*. **8**, 433

58. Röcken, C., Stix, B., Brömme, D., Ansorge, S., Roessner, A., and Bühling, F. (2001) A putative role for cathepsin K in degradation of AA and AL amyloidosis. *Am J Pathol*. **158**, 1029–1038

59. Linke, R. P., Serpell, L. C., Lottspeich, F., and Toyoda, M. (2017) Cathepsin K as a novel amyloid fibril protein in humans. *Amyloid*. **24**, 68–69

60. Muller, P. A. J., and Vousden, K. H. (2013) p53 mutations in cancer. *Nat Cell Biol*. **15**, 2–8

61. Navalkar, A., Pandey, S., Singh, N., Patel, K., Datta, D., Mohanty, B., Jadhav, S., Chaudhari, P., and Maji, S. K. (2021) Direct evidence of cellular transformation by prion-like p53 amyloid infection. *J Cell Sci*. **134**, jcs258316

62. Ghosh, S., Salot, S., Sengupta, S., Navalkar, A., Ghosh, D., Jacob, R., Das, S., Kumar, R., Jha, N. N., Sahay, S., Mehra, S., Mohite, G. M., Ghosh, S. K., Kombrabail, M., Krishnamoorthy, G., Chaudhari, P., and Maji, S. K. (2017) p53 amyloid formation leading to its loss of function: implications in cancer pathogenesis. *Cell Death Differ*. **24**, 1784–

1798

63. Ano Bom, A. P. D., Rangel, L. P., Costa, D. C. F., de Oliveira, G. A. P., Sanches, D., Braga, C. A., Gava, L. M., Ramos, C. H. I., Cepeda, A. O. T., Stumbo, A. C., De Moura Gallo, C. V., Cordeiro, Y., and Silva, J. L. (2012) Mutant p53 aggregates into prion-like amyloid oligomers and fibrils: implications for cancer. *J Biol Chem*. **287**, 28152–28162
64. Navalkar, A., Ghosh, S., Pandey, S., Paul, A., Datta, D., and Maji, S. K. (2020) Prion-like p53 Amyloids in Cancer. *Biochemistry*. **59**, 146–155
65. Bullock, A. N., Henckel, J., DeDecker, B. S., Johnson, C. M., Nikolova, P. V., Proctor, M. R., Lane, D. P., and Fersht, A. R. (1997) Thermodynamic stability of wild-type and mutant p53 core domain. *Proc Natl Acad Sci U S A*. **94**, 14338–14342
66. Lee, A. S., Galea, C., DiGiammarino, E. L., Jun, B., Murti, G., Ribeiro, R. C., Zambetti, G., Schultz, C. P., and Kriwacki, R. W. (2003) Reversible amyloid formation by the p53 tetramerization domain and a cancer-associated mutant. *J Mol Biol*. **327**, 699–709
67. Simon, M., Montézin, M., Guerrin, M., Durieux, J. J., and Serre, G. (1997) Characterization and purification of human corneodesmosin, an epidermal basic glycoprotein associated with corneocyte-specific modified desmosomes. *J Biol Chem*. **272**, 31770–31776
68. Caubet, C., Bousset, L., Clemmensen, O., Sourigues, Y., Bygum, A., Chavanas, S., Coudane, F., Hsu, C.-Y., Betz, R. C., Melki, R., Simon, M., and Serre, G. (2010) A new amyloidosis caused by fibrillar aggregates of mutated corneodesmosin. *FASEB J*. **24**, 3416–3426
69. Dávalos, N. O., García-Vargas, A., Pforr, J., Dávalos, I. P., Picos-Cárdenas, V. J., García-Cruz, D., Kruse, R., Figuera, L. E., Nöthen, M. M., and Betz, R. C. (2005) A non-sense mutation in the corneodesmosin gene in a Mexican family with hypotrichosis simplex of the scalp. *Br J Dermatol*. **153**, 1216–1219

70. Levy-Nissenbaum, E., Betz, R. C., Frydman, M., Simon, M., Lahat, H., Bakhan, T., Goldman, B., Bygum, A., Pierick, M., Hillmer, A. M., Jonca, N., Toribio, J., Kruse, R., Dewald, G., Cichon, S., Kubisch, C., Guerrin, M., Serre, G., Nöthen, M. M., and Pras, E. (2003) Hypotrichosis simplex of the scalp is associated with nonsense mutations in CDSN encoding corneodesmosin. *Nat Genet*. **34**, 151–153

71. Levy, E., Jaskolski, M., and Grubb, A. (2006) The Role of Cystatin C in Cerebral Amyloid Angiopathy and Stroke: Cell Biology and Animal Models. *Brain Pathology*. **16**, 60

72. Whelly, S., Muthusubramanian, A., Powell, J., Johnson, S., Hastert, M. C., and Cornwall, G. A. (2016) Cystatin-related epididymal spermatogenic subgroup members are part of an amyloid matrix and associated with extracellular vesicles in the mouse epididymal lumen. *Mol Hum Reprod*. **22**, 729–744

73. Wahlbom, M., Wang, X., Lindström, V., Carlemalm, E., Jaskolski, M., and Grubb, A. (2007) Fibrillogenic oligomers of human cystatin C are formed by propagated domain swapping. *J Biol Chem*. **282**, 18318–18326

74. Emilsson, V., Thorsteinsson, L., Jensson, O., and Gudnnundsson, G. (1996) Human cystatin C expression and regulation by TGF-β1: Implications for the pathogenesis of hereditary cystatin C amyloid angiopathy causing brain hemorrhage. *Amyloid*. **3**, 110–118

75. Ghiso, J., Jensson, O., and Frangione, B. (1986) Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Icelandic type is a variant of gamma-trace basic protein (cystatin C). *Proc Natl Acad Sci U S A*. **83**, 2974–2978

76. March, M. E., Gutierrez-Uzquiza, A., Snorradottir, A. O., Matsuoka, L. S., Balvis, N. F., Gestsson, T., Nguyen, K., Sleiman, P. M. A., Kao, C., Isaksson, H. J., Bragason, B. T., Olafsson, E., Palsdottir, A., and Hakonarson, H. (2021) NAC blocks Cystatin C amyloid complex aggregation in a cell system and in skin of HCCAA patients. *Nat Commun*. **12**, 1827

77. Palsdottir, A., Snorradottir, A. O., and Thorsteinsson, L. (2006) Hereditary cystatin C amyloid angiopathy: genetic, clinical, and pathological aspects. *Brain Pathol*. **16**, 55–59

78. Clemen, C. S., Herrmann, H., Strelkov, S. V., and Schröder, R. (2013) Desminopathies: pathology and mechanisms. *Acta Neuropathol*. **125**, 47–75

79. De Bleecker, J. L., Engel, A. G., and Ertl, B. B. (1996) Myofibrillar myopathy with abnormal foci of desmin positivity. II. Immunocytochemical analysis reveals accumulation of multiple other proteins. *J Neuropathol Exp Neurol*. **55**, 563–577

80. Selcen, D., Ohno, K., and Engel, A. G. (2004) Myofibrillar myopathy: clinical, morphological and genetic studies in 63 patients. *Brain*. **127**, 439–451

81. Kedia, N., Arhzaouy, K., Pittman, S. K., Sun, Y., Batchelor, M., Weihl, C. C., and Bieschke, J. (2019) Desmin forms toxic, seeding-competent amyloid aggregates that persist in muscle fibers. *Proc Natl Acad Sci U S A*. **116**, 16835–16840

82. Brodehl, A., Dieding, M., Klauke, B., Dec, E., Madaan, S., Huang, T., Gargus, J., Fatima, A., Saric, T., Cakar, H., Walhorn, V., Tönsing, K., Skrzipczyk, T., Cebulla, R., Gerdes, D., Schulz, U., Gummert, J., Svendsen, J. H., Olesen, M. S., Anselmetti, D., Christensen, A. H., Kimonis, V., and Milting, H. (2013) The novel desmin mutant p.A120D impairs filament formation, prevents intercalated disk localization, and causes sudden cardiac death. *Circ Cardiovasc Genet*. **6**, 615–623

83. Schröder, R., Goudeau, B., Simon, M. C., Fischer, D., Eggermann, T., Clemen, C. S., Li, Z., Reimann, J., Xue, Z., Rudnik-Schöneborn, S., Zerres, K., van der Ven, P. F. M., Fürst, D. O., Kunz, W. S., and Vicart, P. (2003) On noxious desmin: functional effects of a novel heterozygous desmin insertion mutation on the extrasarcomeric desmin cytoskeleton and mitochondria. *Hum Mol Genet*. **12**, 657–669

84. Camaj, P., Seeliger, H., Ischenko, I., Krebs, S., Blum, H., De Toni, E. N., Faktorova, D., Jauch, K.-W., and Bruns, C. J. (2009) EFEMP1 binds the EGF receptor and activates

MAPK and Akt pathways in pancreatic carcinoma cells. *Biol Chem*. **390**, 1293–1302

85. Dao, L. N., Kurtin, P. J., Smyrk, T. C., Theis, J. D., Dasari, S., Vrana, J. A., Dispenzieri, A., Nasr, S. H., and McPhail, E. D. (2021) The novel form of amyloidosis derived from EGF-containing fibulin-like extracellular matrix protein 1 (EFEMP1) preferentially affects the lower gastrointestinal tract of elderly femalesa. *Histopathology*. **78**, 459–463
86. Tasaki, M., Ueda, M., Hoshii, Y., Mizukami, M., Matsumoto, S., Nakamura, M., Yamashita, T., Ueda, A., Misumi, Y., Masuda, T., Inoue, Y., Torikai, T., Nomura, T., Tsuda, Y., Kanenawa, K., Isoguchi, A., Okada, M., Matsui, H., Obayashi, K., and Ando, Y. (2019) A novel age-related venous amyloidosis derived from EGF-containing fibulin-like extracellular matrix protein 1. *J Pathol*. **247**, 444–455
87. Chapman, J. R., and Dogan, A. (2019) Fibrinogen alpha amyloidosis: insights from proteomics. *Expert Rev Proteomics*. **16**, 783–793
88. Benson, M. D., Liepnieks, J., Uemichi, T., Wheeler, G., and Correa, R. (1993) Hereditary renal amyloidosis associated with a mutant fibrinogen alpha-chain. *Nat Genet*. **3**, 252–255
89. Uemichi, T., Liepnieks, J. J., Yamada, T., Gertz, M. A., Bang, N., and Benson, M. D. (1996) A frame shift mutation in the fibrinogen A alpha chain gene in a kindred with renal amyloidosis. *Blood*. **87**, 4197–4203
90. Serpell, L. C., Benson, M., Liepnieks, J. J., and Fraser, P. E. (2007) Structural analyses of fibrinogen amyloid fibrils. *Amyloid*. **14**, 199–203
91. Sewgobind, N. V., Albers, S., and Pieters, R. J. (2021) Functions and Inhibition of Galectin-7, an Emerging Target in Cellular Pathophysiology. *Biomolecules*. **11**, 1720
92. Miura, Y., Harumiya, S., Ono, K., Fujimoto, E., Akiyama, M., Fujii, N., Kawano, H., Wachi, H., and Tajima, S. (2013) Galectin-7 and actin are components of amyloid deposit of localized cutaneous amyloidosis. *Exp Dermatol*. **22**, 36–40

93. Chapman, J. R., Liu, A., Yi, S. S., Hernandez, E., Ritorto, M. S., Jungbluth, A. A., Pulitzer, M., and Dogan, A. (2021) Proteomic analysis shows that the main constituent of subepidermal localised cutaneous amyloidosis is not galectin-7. *Amyloid*. **28**, 35–41

94. Ono, K., Fujimoto, E., Fujimoto, N., Akiyama, M., Satoh, T., Maeda, H., Fujii, N., and Tajima, S. (2014) In vitro amyloidogenic peptides of galectin-7: possible mechanism of amyloidogenesis of primary localized cutaneous amyloidosis. *J Biol Chem*. **289**, 29195–29207

95. Solomon, J. P., Page, L. J., Balch, W. E., and Kelly, J. W. (2012) Gelsolin amyloidosis: genetics, biochemistry, pathology and possible strategies for therapeutic intervention. *Crit Rev Biochem Mol Biol*. **47**, 282–296

96. de la Chapelle, A., Tolvanen, R., Boysen, G., Santavy, J., Bleeker-Wagemakers, L., Maury, C. P., and Kere, J. (1992) Gelsolin-derived familial amyloidosis caused by asparagine or tyrosine substitution for aspartic acid at residue 187. *Nat Genet*. **2**, 157–160

97. Schmidt, E.-K., Mustonen, T., Kiuru-Enari, S., Kivelä, T. T., and Atula, S. (2020) Finnish gelsolin amyloidosis causes significant disease burden but does not affect survival: FIN-GAR phase II study. *Orphanet J Rare Dis*. **15**, 19

98. Kiuru-Enari, S., and Haltia, M. (2013) Hereditary gelsolin amyloidosis. *Handb Clin Neurol*. **115**, 659–681

99. Maury, C. P., Alli, K., and Baumann, M. (1990) Finnish hereditary amyloidosis. Amino acid sequence homology between the amyloid fibril protein and human plasma gelsoline. *FEBS Lett*. **260**, 85–87

100. Haltia, M., Prelli, F., Ghiso, J., Kiuru, S., Somer, H., Palo, J., and Frangione, B. (1990) Amyloid protein in familial amyloidosis (Finnish type) is homologous to gelsolin, an actin-binding protein. *Biochem Biophys Res Commun*. **167**, 927–932

101. Maury, C. P. (1991) Gelsolin-related amyloidosis. Identification of the amyloid protein in Finnish hereditary amyloidosis as a fragment of variant gelsolin. *J Clin Invest*. **87**, 1195–1199

102. Maury, C. P., and Nurmiaho-Lassila, E. L. (1992) Creation of amyloid fibrils from mutant Asn187 gelsolin peptides. *Biochem Biophys Res Commun*. **183**, 227–231

103. Maury, C. P., Nurmiaho-Lassila, E. L., and Rossi, H. (1994) Amyloid fibril formation in gelsolin-derived amyloidosis. Definition of the amyloidogenic region and evidence of accelerated amyloid formation of mutant Asn-187 and Tyr-187 gelsolin peptides. *Lab Invest*. **70**, 558–564

104. G, J., and Bb, Z. (2003) Glucagon and regulation of glucose metabolism. *American journal of physiology. Endocrinology and metabolism*. 10.1152/ajpendo.00492.2002

105. Ichimata, S., Katoh, N., Abe, R., Yoshinaga, T., Kametani, F., Yazaki, M., Uehara, T., and Sekijima, Y. (2021) A case of novel amyloidosis: glucagon-derived amyloid deposition associated with pancreatic neuroendocrine tumour. *Amyloid*. **28**, 72–73

106. Gelenter, M. D., Smith, K. J., Liao, S.-Y., Mandala, V. S., Dregni, A. J., Lamm, M. S., Tian, Y., Xu, W., Pochan, D. J., Tucker, T. J., Su, Y., and Hong, M. (2019) The peptide hormone glucagon forms amyloid fibrils with two coexisting β-strand conformations. *Nat Struct Mol Biol*. **26**, 592–598

107. Martins, C. O., Lezcano, C., Yi, S. S., Landau, H. J., Chapman, J. R., and Dogan, A. (2018) Novel iatrogenic amyloidosis caused by peptide drug liraglutide: a clinical mimic of AL amyloidosis. *Haematologica*. **103**, e610–e612

108. Gui, X., Luo, F., Li, Y., Zhou, H., Qin, Z., Liu, Z., Gu, J., Xie, M., Zhao, K., Dai, B., Shin, W. S., He, J., He, L., Jiang, L., Zhao, M., Sun, B., Li, X., Liu, C., and Li, D. (2019) Structural basis for reversible amyloids of hnRNPA1 elucidates their role in stress granule assembly. *Nat Commun*. **10**, 2006

109. Clarke, J. P., Thibault, P. A., Salapa, H. E., and Levin, M. C. (2021) A Comprehensive Analysis of the Role of hnRNP A1 Function and Dysfunction in the Pathogenesis of Neurodegenerative Disease. *Front Mol Biosci*. **8**, 659610

110. Kim, H. J., Kim, N. C., Wang, Y.-D., Scarborough, E. A., Moore, J., Diaz, Z., MacLea, K. S., Freibaum, B., Li, S., Molliex, A., Kanagaraj, A. P., Carter, R., Boylan, K. B., Wojtas, A. M., Rademakers, R., Pinkus, J. L., Greenberg, S. A., Trojanowski, J. Q., Traynor, B. J., Smith, B. N., Topp, S., Gkazi, A.-S., Miller, J., Shaw, C. E., Kottlors, M., Kirschner, J., Pestronk, A., Li, Y. R., Ford, A. F., Gitler, A. D., Benatar, M., King, O. D., Kimonis, V. E., Ross, E. D., Weihl, C. C., Shorter, J., and Taylor, J. P. (2013) Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature*. **495**, 467–473

111. Sun, Y., Zhao, K., Xia, W., Feng, G., Gu, J., Ma, Y., Gui, X., Zhang, X., Fang, Y., Sun, B., Wang, R., Liu, C., and Li, D. (2020) The nuclear localization sequence mediates hnRNPA1 amyloid fibril formation revealed by cryoEM structure. *Nat Commun*. **11**, 6349

112. Kapeli, K., Martinez, F. J., and Yeo, G. W. (2017) Genetic mutations in RNA-binding proteins and their roles in ALS. *Hum Genet*. **136**, 1193–1214

113. Lu, J., Cao, Q., Hughes, M. P., Sawaya, M. R., Boyer, D. R., Cascio, D., and Eisenberg, D. S. (2020) CryoEM structure of the low-complexity domain of hnRNPA2 and its conversion to pathogenic amyloid. *Nat Commun*. **11**, 4090

114. Garcia-Pardo, J., Bartolomé-Nafría, A., Chaves-Sanjuan, A., Gil-Garcia, M., Visentin, C., Bolognesi, M., Ricagno, S., and Ventura, S. (2023) Cryo-EM structure of hnRNPDL-2 fibrils, a functional amyloid associated with limb-girdle muscular dystrophy D3. *Nat Commun*. **14**, 239

115. Schulte, J., and Littleton, J. T. (2011) The biological function of the Huntingtin protein and its relevance to Huntington's Disease pathology. *Curr Trends Neurol*. **5**, 65–78

116. Nasir, J., Floresco, S. B., O'Kusky, J. R., Diewert, V. M., Richman, J. M., Zeisler, J., Borowski, A., Marth, J. D., Phillips, A. G., and Hayden, M. R. (1995) Targeted disruption of the Huntington's disease gene results in embryonic lethality and behavioral and morphological changes in heterozygotes. *Cell*. **81**, 811–823

117. Huang, C. C., Faber, P. W., Persichetti, F., Mittal, V., Vonsattel, J. P., MacDonald, M. E., and Gusella, J. F. (1998) Amyloid formation by mutant huntingtin: threshold, progressivity and recruitment of normal polyglutamine proteins. *Somat Cell Mol Genet*. **24**, 217–233

118. DiFiglia, M., Sapp, E., Chase, K. O., Davies, S. W., Bates, G. P., Vonsattel, J. P., and Aronin, N. (1997) Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. *Science*. **277**, 1990–1993

119. Zhou, H., Cao, F., Wang, Z., Yu, Z.-X., Nguyen, H.-P., Evans, J., Li, S.-H., and Li, X.-J. (2003) Huntingtin forms toxic NH2-terminal fragment complexes that are promoted by the age-dependent decrease in proteasome activity. *J Cell Biol*. **163**, 109–118

120. Scherzinger, E., Lurz, R., Turmaine, M., Mangiarini, L., Hollenbach, B., Hasenbank, R., Bates, G. P., Davies, S. W., Lehrach, H., and Wanker, E. E. (1997) Huntingtin-encoded polyglutamine expansions form amyloid-like protein aggregates in vitro and in vivo. *Cell*. **90**, 549–558

121. Hoop, C. L., Lin, H.-K., Kar, K., Magyarfalvi, G., Lamley, J. M., Boatz, J. C., Mandal, A., Lewandowski, J. R., Wetzel, R., and van der Wel, P. C. A. (2016) Huntingtin exon 1 fibrils feature an interdigitated β-hairpin-based polyglutamine core. *Proc Natl Acad Sci U S A*. **113**, 1546–1551

122. Scherzinger, E., Sittler, A., Schweiger, K., Heiser, V., Lurz, R., Hasenbank, R., Bates, G. P., Lehrach, H., and Wanker, E. E. (1999) Self-assembly of polyglutamine-containing huntingtin fragments into amyloid-like fibrils: implications for Huntington's disease pathology. *Proc Natl Acad Sci U S A*. **96**, 4604–4609

123. Wear, M. P., Kryndushkin, D., O'Meally, R., Sonnenberg, J. L., Cole, R. N., and Shewmaker, F. P. (2015) Proteins with Intrinsically Disordered Domains Are Preferentially Recruited to Polyglutamine Aggregates. *PLoS One*. **10**, e0136362

124. Perutz, M. F. (1995) Glutamine repeats as polar zippers: their role in inherited neurodegenerative disease. *Mol Med*. **1**, 718–721

125. Perutz, M. F. (1996) Glutamine repeats and inherited neurodegenerative diseases: molecular aspects. *Curr Opin Struct Biol*. **6**, 848–858

126. Eulitz, M., Weiss, D. T., and Solomon, A. (1990) Immunoglobulin heavy-chain-associated amyloidosis. *Proc Natl Acad Sci U S A*. **87**, 6542–6546

127. Solomon, A., Weiss, D. T., and Murphy, C. (1994) Primary amyloidosis associated with a novel heavy-chain fragment (AH amyloidosis). *Am J Hematol*. **45**, 171–176

128. Tan, S. Y., Murdoch, I. E., Sullivan, T. J., Wright, J. E., Truong, O., Hsuan, J. J., Hawkins, P. N., and Pepys, M. B. (1994) Primary localized orbital amyloidosis composed of the immunoglobulin gamma heavy chain CH3 domain. *Clin Sci (Lond)*. **87**, 487–491

129. Mai, H. L., Sheikh-Hamad, D., Herrera, G. A., Gu, X., and Truong, L. D. (2003) Immunoglobulin heavy chain can be amyloidogenic: morphologic characterization including immunoelectron microscopy. *Am J Surg Pathol*. **27**, 541–545

130. Perfetti, V., Vignarelli, M. C., Casarini, S., Ascari, E., and Merlini, G. (2001) Biological features of the clone involved in primary amyloidosis (AL). *Leukemia*. **15**, 195–202

131. Glenner, G. G., Terry, W., Harada, M., Isersky, C., and Page, D. (1971) Amyloid fibril proteins: proof of homology with immunoglobulin light chains by sequence analyses. *Science*. **172**, 1150–1151

132. Glenner, G. G., Harbaugh, J., Ohma, J. I., Harada, M., and Cuatrecasas, P. (1970) An amyloid protein: the amino-terminal variable fragment of an immunoglobulin light chain. *Biochem Biophys Res Commun*. **41**, 1287–1289

133. Perfetti, V., Palladini, G., Casarini, S., Navazza, V., Rognoni, P., Obici, L., Invernizzi, R., Perlini, S., Klersy, C., and Merlini, G. (2012) The repertoire of λ light chains causing predominant amyloid heart involvement and identification of a preferentially involved germline gene, IGLV1-44. *Blood*. **119**, 144–150

134. Kourelis, T. V., Dasari, S., Theis, J. D., Ramirez-Alvarado, M., Kurtin, P. J., Gertz, M. A., Zeldenrust, S. R., Zenka, R. M., Dogan, A., and Dispenzieri, A. (2017) Clarifying immunoglobulin gene usage in systemic and localized immunoglobulin light-chain amyloidosis by mass spectrometry. *Blood*. **129**, 299–306

135. Dixon, F. J., and Vazquez, J. J. (1956) Immunohistochemical analysis of amyloid by the fluorescence technique. *J Exp Med*. **104**, 727–736

136. Brumshtein, B., Esswein, S. R., Sawaya, M. R., Rosenberg, G., Ly, A. T., Landau, M., and Eisenberg, D. S. (2018) Identification of two principal amyloid-driving segments in variable domains of Ig light chains in systemic light-chain amyloidosis. *J Biol Chem*. **293**, 19659–19671

137. Radamaker, L., Lin, Y.-H., Annamalai, K., Huhn, S., Hegenbart, U., Schönland, S. O., Fritz, G., Schmidt, M., and Fändrich, M. (2019) Cryo-EM structure of a light chain-derived amyloid fibril from a patient with systemic AL amyloidosis. *Nat Commun*. **10**, 1103

138. Abraham, R. S., Geyer, S. M., Price-Troska, T. L., Allmer, C., Kyle, R. A., Gertz, M. A., and Fonseca, R. (2003) Immunoglobulin light chain variable (V) region genes influence clinical presentation and outcome in light chain-associated amyloidosis (AL). *Blood*. **101**, 3801–3808

139. Rahman, M. S., Hossain, K. S., Das, S., Kundu, S., Adegoke, E. O., Rahman, M. A., Hannan, M. A., Uddin, M. J., and Pang, M.-G. (2021) Role of Insulin in Health and Disease: An Update. *Int J Mol Sci*. **22**, 6403

140. Dische, F. E., Wernstedt, C., Westermark, G. T., Westermark, P., Pepys, M. B., Rennie,

J. A., Gilbey, S. G., and Watkins, P. J. (1988) Insulin as an amyloid-fibril protein at sites of repeated insulin injections in a diabetic patient. *Diabetologia*. **31**, 158–161

141. Iwaya, K., Zako, T., Fukunaga, J., Sörgjerd, K. M., Ogata, K., Kogure, K., Kosano, H., Noritake, M., Maeda, M., Ando, Y., Katsura, Y., and Nagase, T. (2019) Toxicity of insulin-derived amyloidosis: a case report. *BMC Endocr Disord*. **19**, 61
142. Nagase, T., Iwaya, K., Iwaki, Y., Kotake, F., Uchida, R., Oh-i, T., Sekine, H., Miwa, K., Murakami, S., Odaka, T., Kure, M., Nemoto, Y., Noritake, M., and Katsura, Y. (2014) Insulin-derived amyloidosis and poor glycemic control: a case series. *Am J Med*. **127**, 450–454
143. Nagase, T., Katsura, Y., Iwaki, Y., Nemoto, K., Sekine, H., Miwa, K., Oh-I, T., Kou, K., Iwaya, K., Noritake, M., and Matsuoka, T. (2009) The insulin ball. *Lancet*. **373**, 184
144. Shiba, M., and Kitazawa, T. (2016) Progressive insulin-derived amyloidosis in a patient with type 2 diabetes. *Case Reports Plast Surg Hand Surg*. **3**, 73–76
145. Fotinopoulou, A., Tsachaki, M., Vlavaki, M., Poulopoulos, A., Rostagno, A., Frangione, B., Ghiso, J., and Efthimiopoulos, S. (2005) BRI2 interacts with amyloid precursor protein (APP) and regulates amyloid beta (Abeta) production. *J Biol Chem*. **280**, 30768–30772
146. Kim, J., Miller, V. M., Levites, Y., West, K. J., Zwizinski, C. W., Moore, B. D., Troendle, F. J., Bann, M., Verbeeck, C., Price, R. W., Smithson, L., Sonoda, L., Wagg, K., Rangachari, V., Zou, F., Younkin, S. G., Graff-Radford, N., Dickson, D., Rosenberry, T., and Golde, T. E. (2008) BRI2 (ITM2b) inhibits Abeta deposition in vivo. *J Neurosci*. **28**, 6030–6036
147. Fleischer, A., Ayllon, V., and Rebollo, A. (2002) ITM2BS regulates apoptosis by inducing loss of mitochondrial membrane potential. *Eur J Immunol*. **32**, 3498–3505
148. Fleischer, A., and Rebollo, A. (2004) Induction of p53-independent apoptosis by the BH3-only protein ITM2Bs. *FEBS Lett*. **557**, 283–287
149. Kim, S. H., Wang, R., Gordon, D. J., Bass, J., Steiner, D. F., Lynn, D. G., Thinakaran, G.,

Meredith, S. C., and Sisodia, S. S. (1999) Furin mediates enhanced production of fibrillogenic ABri peptides in familial British dementia. *Nat Neurosci*. **2**, 984–988

150. Vidal, R., Revesz, T., Rostagno, A., Kim, E., Holton, J. L., Bek, T., Bojsen-Møller, M., Braendgaard, H., Plant, G., Ghiso, J., and Frangione, B. (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
151. Vidal, R., Frangione, B., Rostagno, A., Mead, S., Révész, T., Plant, G., and Ghiso, J. (1999) A stop-codon mutation in the BRI gene associated with familial British dementia. *Nature*. **399**, 776–781
152. Alehashemi, S., Dasari, S., de Jesus, A. A., Cowen, E. W., Lee, C.-C. R., Goldbach-Mansky, R., and McPhail, E. D. (2022) Anakinra-Associated Amyloidosis. *JAMA Dermatol*. **158**, 1454–1457
153. Westermark, P., Andersson, A., and Westermark, G. T. (2011) Islet amyloid polypeptide, islet amyloid, and diabetes mellitus. *Physiol Rev*. **91**, 795–826
154. Opie, E. L. (1901) ON THE RELATION OF CHRONIC INTERSTITIAL PANCREATITIS TO THE ISLANDS OF LANGERHANS AND TO DIABETES MELUTUS. *J Exp Med*. **5**, 397–428
155. Ahronheim, J. H. (1943) The Nature of the Hyaline Material in the Pancreatic Islands in Diabetes Mellitus. *Am J Pathol*. **19**, 873–882
156. Westermark, P., Wernstedt, C., Wilander, E., and Sletten, K. (1986) A novel peptide in the calcitonin gene related peptide family as an amyloid fibril protein in the endocrine pancreas. *Biochem Biophys Res Commun*. **140**, 827–831
157. Westermark, P., Wernstedt, C., Wilander, E., Hayden, D. W., O'Brien, T. D., and Johnson, K. H. (1987) Amyloid fibrils in human insulinoma and islets of Langerhans of the diabetic cat are derived from a neuropeptide-like protein also present in normal islet cells.

Proc Natl Acad Sci U S A. **84**, 3881–3885

158. Cooper, G. J., Willis, A. C., Clark, A., Turner, R. C., Sim, R. B., and Reid, K. B. (1987) Purification and characterization of a peptide from amyloid-rich pancreases of type 2 diabetic patients. *Proc Natl Acad Sci U S A*. **84**, 8628–8632

159. Gallardo, R., Iadanza, M. G., Xu, Y., Heath, G. R., Foster, R., Radford, S. E., and Ranson, N. A. (2020) Fibril structures of diabetes-related amylin variants reveal a basis for surface-templated assembly. *Nat Struct Mol Biol*. **27**, 1048–1056

160. Cao, Q., Boyer, D. R., Sawaya, M. R., Ge, P., and Eisenberg, D. S. (2020) Cryo-EM structure and inhibitor design of human IAPP (amylin) fibrils. *Nat Struct Mol Biol*. **27**, 653–659

161. Poa, N. R., Cooper, G. J. S., and Edgar, P. F. (2003) Amylin gene promoter mutations predispose to Type 2 diabetes in New Zealand Maori. *Diabetologia*. **46**, 574–578

162. Irvine, A. D., and McLean, W. H. (1999) Human keratin diseases: the increasing spectrum of disease and subtlety of the phenotype-genotype correlation. *Br J Dermatol*. **140**, 815–828

163. Kobayashi, H., and Hashimoto, K. (1983) Amyloidogenesis in organ-limited cutaneous amyloidosis: an antigenic identity between epidermal keratin and skin amyloid. *J Invest Dermatol*. **80**, 66–72

164. Huilgol, S. C., Ramnarain, N., Carrington, P., Leigh, I. M., and Black, M. M. (1998) Cytokeratins in primary cutaneous amyloidosis. *Australas J Dermatol*. **39**, 81–85

165. Chang, Y. T., Liu, H. N., Wang, W. J., Lee, D. D., and Tsai, S. F. (2004) A study of cytokeratin profiles in localized cutaneous amyloids. *Arch Dermatol Res*. **296**, 83–88

166. Betz, R. C., Planko, L., Eigelshoven, S., Hanneken, S., Pasternack, S. M., Bussow, H., Van Den Bogaert, K., Wenzel, J., Braun-Falco, M., Rutten, A., Rogers, M. A., Ruzicka, T., Nöthen, M. M., Magin, T. M., and Kruse, R. (2006) Loss-of-function mutations in the

keratin 5 gene lead to Dowling-Degos disease. *Am J Hum Genet*. **78**, 510–519

167. Chiang, Y.-Y., Chao, S.-C., Chen, W.-Y., Lee, W.-R., and Wang, K.-H. (2008) Weber-Cockayne type of epidermolysis bullosa simplex associated with a novel mutation in keratin 5 and amyloid deposits. *Br J Dermatol*. **159**, 1370–1372
168. Lim, Y., and Ku, N.-O. (2021) Revealing the Roles of Keratin 8/18-Associated Signaling Proteins Involved in the Development of Hepatocellular Carcinoma. *Int J Mol Sci*. **22**, 6401
169. Murray, K. A., Hu, C. J., Seidler, P., Hughes, M. P., Salwinski, L., Sawaya, M., Pan, H., and Eisenberg, D. S. (2022) Identifying amyloid-related diseases by mapping mutations in low-complexity protein domains to pathologies. *Nat. Struct. Mol. Biol.*
170. Ku, N. O., Gish, R., Wright, T. L., and Omary, M. B. (2001) Keratin 8 mutations in patients with cryptogenic liver disease. *N Engl J Med*. **344**, 1580–1587
171. Kamińska, A., Enguita, F. J., and Stępień, E. Ł. (2018) Lactadherin: An unappreciated haemostasis regulator and potential therapeutic agent. *Vascul Pharmacol*. **101**, 21–28
172. Häggqvist, B., Näslund, J., Sletten, K., Westermark, G. T., Mucchiano, G., Tjernberg, L. O., Nordstedt, C., Engström, U., and Westermark, P. (1999) Medin: an integral fragment of aortic smooth muscle cell-produced lactadherin forms the most common human amyloid. *Proc Natl Acad Sci U S A*. **96**, 8669–8674
173. Mucchiano, G., Cornwell, G. G., and Westermark, P. (1992) Senile aortic amyloid. Evidence for two distinct forms of localized deposits. *Am J Pathol*. **140**, 871–877
174. Giansanti, F., Panella, G., Leboffe, L., and Antonini, G. (2016) Lactoferrin from Milk: Nutraceutical and Pharmacological Properties. *Pharmaceuticals (Basel)*. **9**, 61
175. Klintworth, G. K., Valnickova, Z., Kielar, R. A., Baratz, K. H., Campbell, R. J., and Enghild, J. J. (1997) Familial subepithelial corneal amyloidosis--a lactoferrin-related amyloidosis. *Invest Ophthalmol Vis Sci*. **38**, 2756–2763

176. Ando, Y., Nakamura, M., Kai, H., Katsuragi, S., Terazaki, H., Nozawa, T., Okuda, T., Misumi, S., Matsunaga, N., Hata, K., Tajiri, T., Shoji, S., Yamashita, T., Haraoka, K., Obayashi, K., Matsumoto, K., Ando, M., and Uchino, M. (2002) A novel localized amyloidosis associated with lactoferrin in the cornea. *Lab Invest*. **82**, 757–766

177. Araki-Sasaki, K., Ando, Y., Nakamura, M., Kitagawa, K., Ikemizu, S., Kawaji, T., Yamashita, T., Ueda, M., Hirano, K., Yamada, M., Matsumoto, K., Kinoshita, S., and Tanihara, H. (2005) Lactoferrin Glu561Asp facilitates secondary amyloidosis in the cornea. *Br J Ophthalmol*. **89**, 684–688

178. Baugh, K. A., Desai, S., Van Buren Nd, G., Fisher, W. E., Farinas, C. A., and Dhingra, S. (2021) Lactoferrin amyloid presenting as a mural nodule in a pancreatic cystic lesion prompting pancreatoduodenectomy: a case report. *BMC Gastroenterol*. **21**, 66

179. Ichimata, S., Aoyagi, D., Yoshinaga, T., Katoh, N., Kametani, F., Yazaki, M., Uehara, T., and Shiozawa, S. (2019) A case of spheroid-type localized lactoferrin amyloidosis in the bronchus. *Pathol Int*. **69**, 235–240

180. Tsutsumi, Y., Serizawa, A., and Hori, S. (1996) Localized amyloidosis of the seminal vesicle: identification of lactoferrin immunoreactivity in the amyloid material. *Pathol Int*. **46**, 491–497

181. Tsujikawa, M., Kurahashi, H., Tanaka, T., Nishida, K., Shimomura, Y., Tano, Y., and Nakamura, Y. (1999) Identification of the gene responsible for gelatinous drop-like corneal dystrophy. *Nat Genet*. **21**, 420–423

182. Nilsson, M. R., and Dobson, C. M. (2003) In vitro characterization of lactoferrin aggregation and amyloid formation. *Biochemistry*. **42**, 375–382

183. Slowik, V., and Apte, U. (2017) Leukocyte Cell-Derived Chemotaxin-2: It's Role in Pathophysiology and Future in Clinical Medicine. *Clin Transl Sci*. **10**, 249–259

184. Benson, M. D., James, S., Scott, K., Liepnieks, J. J., and Kluve-Beckerman, B. (2008)

Leukocyte chemotactic factor 2: A novel renal amyloid protein. *Kidney Int*. **74**, 218–222

185. Mann, B. K., Bhandohal, J. S., Cobos, E., Chitturi, C., and Eppanapally, S. (2022) LECT-2 amyloidosis: what do we know? *J Investig Med*. **70**, 348–353

186. Larsen, C. P., Kossmann, R. J., Beggs, M. L., Solomon, A., and Walker, P. D. (2014) Clinical, morphologic, and genetic features of renal leukocyte chemotactic factor 2 amyloidosis. *Kidney Int*. **86**, 378–382

187. Murphy, C. L., Wang, S., Kestler, D., Larsen, C., Benson, D., Weiss, D. T., and Solomon, A. (2010) Leukocyte chemotactic factor 2 (LECT2)-associated renal amyloidosis: a case series. *Am J Kidney Dis*. **56**, 1100–1107

188. Mereuta, O. M., Theis, J. D., Vrana, J. A., Law, M. E., Grogg, K. L., Dasari, S., Chandan, V. S., Wu, T.-T., Jimenez-Zepeda, V. H., Fonseca, R., Dispenzieri, A., Kurtin, P. J., and Dogan, A. (2014) Leukocyte cell-derived chemotaxin 2 (LECT2)-associated amyloidosis is a frequent cause of hepatic amyloidosis in the United States. *Blood*. **123**, 1479–1482

189. Rezk, T., Gilbertson, J. A., Rowczenio, D., Bass, P., Lachmann, H. J., Wechalekar, A. D., Fontana, M., Mahmood, S., Sachchithanantham, S., Whelan, C. J., Wong, J., Rendell, N., Taylor, G. W., Hawkins, P. N., and Gillmore, J. D. (2018) Diagnosis, pathogenesis and outcome in leucocyte chemotactic factor 2 (ALECT2) amyloidosis. *Nephrol Dial Transplant*. **33**, 241–247

190. Ortega Junco, E., Sánchez González, C., Serrano Pardo, R., Lamana Dominguez, A., Santos Sánchez, B., Sanz Sainz, M., Saharaui Catala, Y., and Sánchez Tomero, J. A. (2018) LECT2-associated renal amyloidosis (ALECT2): A case report. *Nefrologia (Engl Ed)*. **38**, 558–560

191. Ha, J.-H., Tu, H.-C., Wilkens, S., and Loh, S. N. (2021) Loss of bound zinc facilitates amyloid fibril formation of leukocyte-cell-derived chemotaxin 2 (LECT2). *J Biol Chem*. **296**, 100446

192. Richards, L. S., Flores, M. D., Zink, S., Schibrowsky, N. A., Sawaya, M. R., and Rodriguez, J. A. (2023) Cryo-EM Structure of a Human LECT2 Amyloid Fibril Reveals a Network of Polar Ladders at its Core. *bioRxiv*. 10.1101/2023.02.08.527771

193. Ferraboschi, P., Ciceri, S., and Grisenti, P. (2021) Applications of Lysozyme, an Innate Immune Defense Factor, as an Alternative Antibiotic. *Antibiotics (Basel)*. **10**, 1534

194. Pepys, M. B., Hawkins, P. N., Booth, D. R., Vigushin, D. M., Tennent, G. A., Soutar, A. K., Totty, N., Nguyen, O., Blake, C. C., and Terry, C. J. (1993) Human lysozyme gene mutations cause hereditary systemic amyloidosis. *Nature*. **362**, 553–557

195. Moura, A., Nocerino, P., Gilbertson, J. A., Rendell, N. B., Mangione, P. P., Verona, G., Rowczenio, D., Gillmore, J. D., Taylor, G. W., Bellotti, V., and Canetti, D. (2020) Lysozyme amyloid: evidence for the W64R variant by proteomics in the absence of the wild type protein. *Amyloid*. **27**, 206–207

196. Morozova-Roche, L. A., Zurdo, J., Spencer, A., Noppe, W., Receveur, V., Archer, D. B., Joniau, M., and Dobson, C. M. (2000) Amyloid fibril formation and seeding by wild-type human lysozyme and its disease-related mutational variants. *J Struct Biol*. **130**, 339–351

197. Nasr, S. H., Dasari, S., Mills, J. R., Theis, J. D., Zimmermann, M. T., Fonseca, R., Vrana, J. A., Lester, S. J., McLaughlin, B. M., Gillespie, R., Highsmith, W. E., Lee, J. J., Dispenzieri, A., and Kurtin, P. J. (2017) Hereditary Lysozyme Amyloidosis Variant p.Leu102Ser Associates with Unique Phenotype. *J Am Soc Nephrol*. **28**, 431–438

198. Legname, G. (2017) Elucidating the function of the prion protein. *PLoS Pathog*. **13**, e1006458

199. Prusiner, S. B. (1998) Prions. *Proc Natl Acad Sci U S A*. **95**, 13363–13383

200. Kim, M.-O., Takada, L. T., Wong, K., Forner, S. A., and Geschwind, M. D. (2018) Genetic PrP Prion Diseases. *Cold Spring Harb Perspect Biol*. **10**, a033134

201. Prusiner, S. B. (1982) Novel proteinaceous infectious particles cause scrapie. *Science*.

216, 136–144

202. Prusiner, S. B., McKinley, M. P., Bowman, K. A., Bolton, D. C., Bendheim, P. E., Groth, D. F., and Glenner, G. G. (1983) Scrapie prions aggregate to form amyloid-like birefringent rods. *Cell*. **35**, 349–358

203. DeArmond, S. J., McKinley, M. P., Barry, R. A., Braunfeld, M. B., McColloch, J. R., and Prusiner, S. B. (1985) Identification of prion amyloid filaments in scrapie-infected brain. *Cell*. **41**, 221–235

204. Kitamoto, T., Tateishi, J., Tashima, T., Takeshita, I., Barry, R. A., DeArmond, S. J., and Prusiner, S. B. (1986) Amyloid plaques in Creutzfeldt-Jakob disease stain with prion protein antibodies. *Ann Neurol*. **20**, 204–208

205. Minikel, E. V., Vallabh, S. M., Lek, M., Estrada, K., Samocha, K. E., Sathirapongsasuti, J. F., McLean, C. Y., Tung, J. Y., Yu, L. P. C., Gambetti, P., Blevins, J., Zhang, S., Cohen, Y., Chen, W., Yamada, M., Hamaguchi, T., Sanjo, N., Mizusawa, H., Nakamura, Y., Kitamoto, T., Collins, S. J., Boyd, A., Will, R. G., Knight, R., Ponto, C., Zerr, I., Kraus, T. F. J., Eigenbrod, S., Giese, A., Calero, M., de Pedro-Cuesta, J., Haïk, S., Laplanche, J.-L., Bouaziz-Amar, E., Brandel, J.-P., Capellari, S., Parchi, P., Poleggi, A., Ladogana, A., O'Donnell-Luria, A. H., Karczewski, K. J., Marshall, J. L., Boehnke, M., Laakso, M., Mohlke, K. L., Kähler, A., Chambert, K., McCarroll, S., Sullivan, P. F., Hultman, C. M., Purcell, S. M., Sklar, P., van der Lee, S. J., Rozemuller, A., Jansen, C., Hofman, A., Kraaij, R., van Rooij, J. G. J., Ikram, M. A., Uitterlinden, A. G., van Duijn, C. M., Exome Aggregation Consortium (ExAC), Daly, M. J., and MacArthur, D. G. (2016) Quantifying prion disease penetrance using large population control cohorts. *Sci Transl Med*. **8**, 322ra9

206. Wang, L.-Q., Zhao, K., Yuan, H.-Y., Li, X.-N., Dang, H.-B., Ma, Y., Wang, Q., Wang, C., Sun, Y., Chen, J., Li, D., Zhang, D., Yin, P., Liu, C., and Liang, Y. Genetic prion disease–

related mutation E196K displays a novel amyloid fibril structure revealed by cryo-EM. *Sci Adv*. **7**, eabg9676

207. Hadži, S., Ondračka, A., Jerala, R., and Hafner-Bratkovič, I. (2015) Pathological mutations H187R and E196K facilitate subdomain separation and prion protein conversion by destabilization of the native structure. *FASEB J*. **29**, 882–893

208. Tagliavini, F., Prelli, F., Ghiso, J., Bugiani, O., Serban, D., Prusiner, S. B., Farlow, M. R., Ghetti, B., and Frangione, B. (1991) Amyloid protein of Gerstmann-Sträussler-Scheinker disease (Indiana kindred) is an 11 kd fragment of prion protein with an N-terminal glycine at codon 58. *EMBO J*. **10**, 513–519

209. Ghetti, B., Piccardo, P., and Zanusso, G. (2018) Dominantly inherited prion protein cerebral amyloidoses - a modern view of Gerstmann-Sträussler-Scheinker. *Handb Clin Neurol*. **153**, 243–269

210. Roeber, S., Krebs, B., Neumann, M., Windl, O., Zerr, I., Grasbon-Frodl, E.-M., and Kretzschmar, H. A. (2005) Creutzfeldt-Jakob disease in a patient with an R208H mutation of the prion protein gene (PRNP) and a 17-kDa prion protein fragment. *Acta Neuropathol*. **109**, 443–448

211. Hallinan, G. I., Ozcan, K. A., Hoq, M. R., Cracco, L., Vago, F. S., Bharath, S. R., Li, D., Jacobsen, M., Doud, E. H., Mosley, A. L., Fernandez, A., Garringer, H. J., Jiang, W., Ghetti, B., and Vidal, R. (2022) Cryo-EM structures of prion protein filaments from Gerstmann-Sträussler-Scheinker disease. *Acta Neuropathol*. **144**, 509–520

212. Raposo, G., and Marks, M. S. (2007) Melanosomes--dark organelles enlighten endosomal membrane transport. *Nat Rev Mol Cell Biol*. **8**, 786–797

213. Lahola-Chomiak, A. A., Footz, T., Nguyen-Phuoc, K., Neil, G. J., Fan, B., Allen, K. F., Greenfield, D. S., Parrish, R. K., Linkroum, K., Pasquale, L. R., Leonhardt, R. M., Ritch, R., Javadiyan, S., Craig, J. E., Allison, W. T., Lehmann, O. J., Walter, M. A., and Wiggs,

J. L. (2019) Non-Synonymous variants in premelanosome protein (PMEL) cause ocular pigment dispersion and pigmentary glaucoma. *Hum Mol Genet*. **28**, 1298–1311

214. Watt, B., Tenza, D., Lemmon, M. A., Kerje, S., Raposo, G., Andersson, L., and Marks, M. S. (2011) Mutations in or near the transmembrane domain alter PMEL amyloid formation from functional to pathogenic. *PLoS Genet*. **7**, e1002286
215. Goedert, M. (2005) Tau gene mutations and their effects. *Mov Disord*. **20 Suppl 12**, S45-52
216. Zhang, X., Lin, Y., Eschmann, N. A., Zhou, H., Rauch, J. N., Hernandez, I., Guzman, E., Kosik, K. S., and Han, S. (2017) RNA stores tau reversibly in complex coacervates. *PLoS Biol*. **15**, e2002183
217. Goedert, M., Eisenberg, D. S., and Crowther, R. A. (2017) Propagation of Tau Aggregates and Neurodegeneration. *Annu Rev Neurosci*. **40**, 189–210
218. Wolfe, M. S. (2009) Tau mutations in neurodegenerative diseases. *J Biol Chem*. **284**, 6021–6025
219. Buée, L., Bussière, T., Buée-Scherrer, V., Delacourte, A., and Hof, P. R. (2000) Tau protein isoforms, phosphorylation and role in neurodegenerative disorders. *Brain Res Brain Res Rev*. **33**, 95–130
220. Tacik, P., DeTure, M., Hinkle, K. M., Lin, W.-L., Sanchez-Contreras, M., Carlomagno, Y., Pedraza, O., Rademakers, R., Ross, O. A., Wszolek, Z. K., and Dickson, D. W. (2015) A Novel Tau Mutation in Exon 12, p.Q336H, Causes Hereditary Pick Disease. *J Neuropathol Exp Neurol*. **74**, 1042–1052
221. Grundke-Iqbal, I., Iqbal, K., Quinlan, M., Tung, Y. C., Zaidi, M. S., and Wisniewski, H. M. (1986) Microtubule-associated protein tau. A component of Alzheimer paired helical filaments. *J Biol Chem*. **261**, 6084–6089
222. Spina, S., Farlow, M. R., Unverzagt, F. W., Kareken, D. A., Murrell, J. R., Fraser, G.,

Epperson, F., Crowther, R. A., Spillantini, M. G., Goedert, M., and Ghetti, B. (2008) The tauopathy associated with mutation +3 in intron 10 of Tau: characterization of the MSTD family. *Brain*. **131**, 72–89

223. Spillantini, M. G., Murrell, J. R., Goedert, M., Farlow, M. R., Klug, A., and Ghetti, B. (1998) Mutation in the tau gene in familial multiple system tauopathy with presenile dementia. *Proc Natl Acad Sci U S A*. **95**, 7737–7741

224. Varani, L., Hasegawa, M., Spillantini, M. G., Smith, M. J., Murrell, J. R., Ghetti, B., Klug, A., Goedert, M., and Varani, G. (1999) Structure of tau exon 10 splicing regulatory element RNA and destabilization by mutations of frontotemporal dementia and parkinsonism linked to chromosome 17. *Proc Natl Acad Sci U S A*. **96**, 8229–8234

225. Hasegawa, M., Smith, M. J., Iijima, M., Tabira, T., and Goedert, M. (1999) FTDP-17 mutations N279K and S305N in tau produce increased splicing of exon 10. *FEBS Lett*. **443**, 93–96

226. Hutton, M., Lendon, C. L., Rizzu, P., Baker, M., Froelich, S., Houlden, H., Pickering-Brown, S., Chakraverty, S., Isaacs, A., Grover, A., Hackett, J., Adamson, J., Lincoln, S., Dickson, D., Davies, P., Petersen, R. C., Stevens, M., de Graaff, E., Wauters, E., van Baren, J., Hillebrand, M., Joosse, M., Kwon, J. M., Nowotny, P., Che, L. K., Norton, J., Morris, J. C., Reed, L. A., Trojanowski, J., Basun, H., Lannfelt, L., Neystat, M., Fahn, S., Dark, F., Tannenberg, T., Dodd, P. R., Hayward, N., Kwok, J. B., Schofield, P. R., Andreadis, A., Snowden, J., Craufurd, D., Neary, D., Owen, F., Oostra, B. A., Hardy, J., Goate, A., van Swieten, J., Mann, D., Lynch, T., and Heutink, P. (1998) Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature*. **393**, 702–705

227. Ando, K., Ferlini, L., Suain, V., Yilmaz, Z., Mansour, S., Le Ber, I., Bouchard, C., Leroy, K., Durr, A., Clot, F., Sarazin, M., Bier, J.-C., and Brion, J.-P. (2020) de novo MAPT

mutation G335A causes severe brain atrophy, 3R and 4R PHF-tau pathology and early onset frontotemporal dementia. *Acta Neuropathol Commun*. **8**, 94

228. Pickering-Brown, S. M., Baker, M., Nonaka, T., Ikeda, K., Sharma, S., Mackenzie, J., Simpson, S. A., Moore, J. W., Snowden, J. S., de Silva, R., Revesz, T., Hasegawa, M., Hutton, M., and Mann, D. M. A. (2004) Frontotemporal dementia with Pick-type histology associated with Q336R mutation in the tau gene. *Brain*. **127**, 1415–1426
229. Jeganathan, S., von Bergen, M., Mandelkow, E.-M., and Mandelkow, E. (2008) The natively unfolded character of tau and its aggregation to Alzheimer-like paired helical filaments. *Biochemistry*. **47**, 10526–10539
230. Nacharaju, P., Lewis, J., Easson, C., Yen, S., Hackett, J., Hutton, M., and Yen, S. H. (1999) Accelerated filament formation from tau protein with specific FTDP-17 missense mutations. *FEBS Lett*. **447**, 195–199
231. Goedert, M., Jakes, R., and Crowther, R. A. (1999) Effects of frontotemporal dementia FTDP-17 mutations on heparin-induced assembly of tau filaments. *FEBS Lett*. **450**, 306–311
232. Gamblin, T. C., King, M. E., Dawson, H., Vitek, M. P., Kuret, J., Berry, R. W., and Binder, L. I. (2000) In vitro polymerization of tau protein monitored by laser light scattering: method and application to the study of FTDP-17 mutants. *Biochemistry*. **39**, 6136–6144
233. Barghorn, S., Zheng-Fischhöfer, Q., Ackmann, M., Biernat, J., von Bergen, M., Mandelkow, E. M., and Mandelkow, E. (2000) Structure, microtubule interactions, and paired helical filament aggregation by tau mutants of frontotemporal dementias. *Biochemistry*. **39**, 11714–11721
234. von Bergen, M., Barghorn, S., Li, L., Marx, A., Biernat, J., Mandelkow, E. M., and Mandelkow, E. (2001) Mutations of tau protein in frontotemporal dementia promote aggregation of paired helical filaments by enhancing local beta-structure. *J Biol Chem*.

276, 48165–48174

235. Alonso, A., Zaidi, T., Novak, M., Grundke-Iqbal, I., and Iqbal, K. (2001) Hyperphosphorylation induces self-assembly of tau into tangles of paired helical filaments/straight filaments. *Proc Natl Acad Sci U S A*. **98**, 6923–6928
236. Alonso, A. C., Grundke-Iqbal, I., and Iqbal, K. (1996) Alzheimer's disease hyperphosphorylated tau sequesters normal tau into tangles of filaments and disassembles microtubules. *Nat Med*. **2**, 783–787
237. Cohen, T. J., Guo, J. L., Hurtado, D. E., Kwong, L. K., Mills, I. P., Trojanowski, J. Q., and Lee, V. M. Y. (2011) The acetylation of tau inhibits its function and promotes pathological tau aggregation. *Nat Commun*. **2**, 252
238. Li, L., Nguyen, B., Mullapudi, V., Saelices, L., and Joachimiak, L. A. (2023) Disease-associated patterns of acetylation stabilize tau fibril formation. *bioRxiv*. 10.1101/2023.01.10.523459
239. Zhu, S., Xiang, C., Charlesworth, O., Bennett, S., Zhang, S., Zhou, M., Kujan, O., and Xu, J. (2022) The versatile roles of odontogenic ameloblast-associated protein in odontogenesis, junctional epithelium regeneration and periodontal disease. *Front Physiol*. **13**, 1003931
240. Murphy, C. L., Kestler, D. P., Foster, J. S., Wang, S., Macy, S. D., Kennel, S. J., Carlson, E. R., Hudson, J., Weiss, D. T., and Solomon, A. (2008) Odontogenic ameloblast-associated protein nature of the amyloid found in calcifying epithelial odontogenic tumors and unerupted tooth follicles. *Amyloid*. **15**, 89–95
241. Solomon, A., Murphy, C. L., Weaver, K., Weiss, D. T., Hrncic, R., Eulitz, M., Donnell, R. L., Sletten, K., Westermark, G., and Westermark, P. (2003) Calcifying epithelial odontogenic (Pindborg) tumor-associated amyloid consists of a novel human protein. *J Lab Clin Med*. **142**, 348–355

242. Khan, M., Jose, A., and Sharma, S. (2023) Physiology, Parathyroid Hormone. in *StatPearls*, StatPearls Publishing, Treasure Island (FL), [online] http://www.ncbi.nlm.nih.gov/books/NBK499940/ (Accessed May 25, 2023)

243. Colombat, M., Barres, B., Renaud, C., Ribes, D., Pericard, S., Camus, M., Anesia, R., van Acker, N., Chauveau, D., Burlet-Schiltz, O., Brousset, P., and Valleix, S. (2021) Mass spectrometry-based proteomic analysis of parathyroid adenomas reveals PTH as a new human hormone-derived amyloid fibril protein. *Amyloid*. **28**, 153–157

244. Kedar, I., Ravid, M., and Sohar, E. (1976) In vitro synthesis of “amyloid”fibrils from insulin, calcitonin and parathormone. *Isr J Med Sci*. **12**, 1137–1140

245. Gopalswamy, M., Kumar, A., Adler, J., Baumann, M., Henze, M., Kumar, S. T., Fändrich, M., Scheidt, H. A., Huster, D., and Balbach, J. (2015) Structural characterization of amyloid fibrils from the human parathyroid hormone. *Biochim Biophys Acta*. **1854**, 249–257

246. Wahle, E. (1991) A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. *Cell*. **66**, 759–768

247. Calado, A., Tomé, F. M., Brais, B., Rouleau, G. A., Kühn, U., Wahle, E., and Carmo-Fonseca, M. (2000) Nuclear inclusions in oculopharyngeal muscular dystrophy consist of poly(A) binding protein 2 aggregates which sequester poly(A) RNA. *Hum Mol Genet*. **9**, 2321–2328

248. Scheuermann, T., Schulz, B., Blume, A., Wahle, E., Rudolph, R., and Schwarz, E. (2003) Trinucleotide expansions leading to an extended poly-L-alanine segment in the poly (A) binding protein PABPN1 cause fibril formation. *Protein Sci*. **12**, 2685–2692

249. Robinson, D. O., Wills, A. J., Hammans, S. R., Read, S. P., and Sillibourne, J. (2006) Oculopharyngeal muscular dystrophy: a point mutation which mimics the effect of the PABPN1 gene triplet repeat expansion mutation. *J Med Genet*. **43**, e23

250. Al-Chalabi, M., Bass, A. N., and Alsalman, I. (2023) Physiology, Prolactin. in *StatPearls*, StatPearls Publishing, Treasure Island (FL), [online] http://www.ncbi.nlm.nih.gov/books/NBK507829/ (Accessed May 25, 2023)

251. Westermark, P., Eriksson, L., Engström, U., Eneström, S., and Sletten, K. (1997) Prolactin-derived amyloid in the aging pituitary gland. *Am J Pathol*. **150**, 67–73

252. Hinton, D. R., Polk, R. K., Linse, K. D., Weiss, M. H., Kovacs, K., and Garner, J. A. (1997) Characterization of spherical amyloid protein from a prolactin-producing pituitary adenoma. *Acta Neuropathol*. **93**, 43–49

253. Johnson, A., Bhattacharya, N., Hanna, M., Pennington, J. G., Schuh, A. L., Wang, L., Otegui, M. S., Stagg, S. M., and Audhya, A. (2015) TFG clusters COPII-coated transport carriers and promotes early secretory pathway organization. *EMBO J*. **34**, 811–827

254. Tsai, P.-C., Huang, Y.-H., Guo, Y.-C., Wu, H.-T., Lin, K.-P., Tsai, Y.-S., Liao, Y.-C., Liu, Y.-T., Liu, T.-T., Kao, L.-S., Yet, S.-F., Fann, M.-J., Soong, B.-W., and Lee, Y.-C. (2014) A novel TFG mutation causes Charcot-Marie-Tooth disease type 2 and impairs TFG function. *Neurology*. **83**, 903–912

255. Ishiura, H., Sako, W., Yoshida, M., Kawarai, T., Tanabe, O., Goto, J., Takahashi, Y., Date, H., Mitsui, J., Ahsan, B., Ichikawa, Y., Iwata, A., Yoshino, H., Izumi, Y., Fujita, K., Maeda, K., Goto, S., Koizumi, H., Morigaki, R., Ikemura, M., Yamauchi, N., Murayama, S., Nicholson, G. A., Ito, H., Sobue, G., Nakagawa, M., Kaji, R., and Tsuji, S. (2012) The TRK-fused gene is mutated in hereditary motor and sensory neuropathy with proximal dominant involvement. *Am J Hum Genet*. **91**, 320–329

256. Rosenberg, G. M., Murray, K. A., Salwinski, L., Hughes, M. P., Abskharon, R., and Eisenberg, D. S. (2022) Bioinformatic identification of previously unrecognized amyloidogenic proteins. *J Biol Chem*. **298**, 101920

257. Sehlmeyer, K., Ruwisch, J., Roldan, N., and Lopez-Rodriguez, E. (2020) Alveolar

Dynamics and Beyond - The Importance of Surfactant Protein C and Cholesterol in Lung Homeostasis and Fibrosis. *Front Physiol*. **11**, 386

258. Gustafsson, M., Thyberg, J., Näslund, J., Eliasson, E., and Johansson, J. (1999) Amyloid fibril formation by pulmonary surfactant protein C. *FEBS Lett*. **464**, 138–142

259. Szyperski, T., Vandenbussche, G., Curstedt, T., Ruysschaert, J. M., Wüthrich, K., and Johansson, J. (1998) Pulmonary surfactant-associated polypeptide C in a mixed organic solvent transforms from a monomeric alpha-helical state into insoluble beta-sheet aggregates. *Protein Sci*. **7**, 2533–2540

260. Lagier-Tourenne, C., Polymenidou, M., and Cleveland, D. W. (2010) TDP-43 and FUS/TLS: emerging roles in RNA processing and neurodegeneration. *Hum Mol Genet*. **19**, R46-64

261. Zinszner, H., Sok, J., Immanuel, D., Yin, Y., and Ron, D. (1997) TLS (FUS) binds RNA in vivo and engages in nucleo-cytoplasmic shuttling. *J Cell Sci*. **110 (Pt 15)**, 1741–1750

262. Kato, M., Han, T. W., Xie, S., Shi, K., Du, X., Wu, L. C., Mirzaei, H., Goldsmith, E. J., Longgood, J., Pei, J., Grishin, N. V., Frantz, D. E., Schneider, J. W., Chen, S., Li, L., Sawaya, M. R., Eisenberg, D., Tycko, R., and McKnight, S. L. (2012) Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell*. **149**, 753–767

263. Murray, D. T., Kato, M., Lin, Y., Thurber, K. R., Hung, I., McKnight, S. L., and Tycko, R. (2017) Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. *Cell*. **171**, 615-627.e16

264. Kwiatkowski, T. J., Bosco, D. A., Leclerc, A. L., Tamrazian, E., Vanderburg, C. R., Russ, C., Davis, A., Gilchrist, J., Kasarskis, E. J., Munsat, T., Valdmanis, P., Rouleau, G. A., Hosler, B. A., Cortelli, P., de Jong, P. J., Yoshinaga, Y., Haines, J. L., Pericak-Vance, M. A., Yan, J., Ticozzi, N., Siddique, T., McKenna-Yasek, D., Sapp, P. C., Horvitz, H. R.,

Landers, J. E., and Brown, R. H. (2009) Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science*. **323**, 1205–1208

265. Sun, Z., Diaz, Z., Fang, X., Hart, M. P., Chesi, A., Shorter, J., and Gitler, A. D. (2011) Molecular determinants and genetic modifiers of aggregation and toxicity for the ALS disease protein FUS/TLS. *PLoS Biol*. **9**, e1000614
266. Patel, A., Lee, H. O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M. Y., Stoynov, S., Mahamid, J., Saha, S., Franzmann, T. M., Pozniakovski, A., Poser, I., Maghelli, N., Royer, L. A., Weigert, M., Myers, E. W., Grill, S., Drechsel, D., Hyman, A. A., and Alberti, S. (2015) A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell*. **162**, 1066–1077
267. Sun, Y., Zhang, S., Hu, J., Tao, Y., Xia, W., Gu, J., Li, Y., Cao, Q., Li, D., and Liu, C. (2022) Molecular structure of an amyloid fibril formed by FUS low-complexity domain. *iScience*. **25**, 103701
268. Nomura, T., Watanabe, S., Kaneko, K., Yamanaka, K., Nukina, N., and Furukawa, Y. (2014) Intranuclear aggregation of mutant FUS/TLS as a molecular pathomechanism of amyotrophic lateral sclerosis. *J Biol Chem*. **289**, 1192–1202
269. Vogl, T., Gharibyan, A. L., and Morozova-Roche, L. A. (2012) Pro-inflammatory S100A8 and S100A9 proteins: self-assembly into multifunctional native and amyloid complexes. *Int J Mol Sci*. **13**, 2893–2917
270. Yanamandra, K., Alexeyev, O., Zamotin, V., Srivastava, V., Shchukarev, A., Brorsson, A.-C., Tartaglia, G. G., Vogl, T., Kayed, R., Wingsle, G., Olsson, J., Dobson, C. M., Bergh, A., Elgh, F., and Morozova-Roche, L. A. (2009) Amyloid formation by the pro-inflammatory S100A8/A9 proteins in the ageing prostate. *PLoS One*. **4**, e5562
271. Sakaguchi, D., Miyado, K., Iwamoto, T., Okada, H., Yoshida, K., Kang, W., Suzuki, M., Yoshida, M., and Kawano, N. (2020) Human Semenogelin 1 Promotes Sperm Survival in

the Mouse Female Reproductive Tract. *Int J Mol Sci*. **21**, 3961

272. Linke, R. P., Joswig, R., Murphy, C. L., Wang, S., Zhou, H., Gross, U., Rocken, C., Westermark, P., Weiss, D. T., and Solomon, A. (2005) Senile seminal vesicle amyloid is derived from semenogelin I. *J Lab Clin Med*. **145**, 187–193
273. Eklund, K. K., Niemi, K., and Kovanen, P. T. (2012) Immune functions of serum amyloid A. *Crit Rev Immunol*. **32**, 335–348
274. Liberta, F., Rennegarbe, M., Rösler, R., Bijzet, J., Wiese, S., Hazenberg, B. P. C., and Fändrich, M. (2019) Morphological and primary structural consistency of fibrils from different AA patients (common variant). *Amyloid*. **26**, 164–170
275. Sikora, J., Kmochová, T., Mušálková, D., Pohludka, M., Přikryl, P., Hartmannová, H., Hodaňová, K., Trešlová, H., Nosková, L., Mrázová, L., Stránecký, V., Lunová, M., Jirsa, M., Honsová, E., Dasari, S., McPhail, E. D., Leung, N., Živná, M., Bleyer, A. J., Rychlík, I., Ryšavá, R., and Kmoch, S. (2022) A mutation in the SAA1 promoter causes hereditary amyloid A amyloidosis. *Kidney Int*. **101**, 349–359
276. Levin, M., Franklin, E. C., Frangione, B., and Pras, M. (1972) The amino acid sequence of a major nonimmunoglobulin component of some amyloid fibrils. *J Clin Invest*. **51**, 2773–2776
277. Linke, R. P., Sipe, J. D., Pollock, P. S., Ignaczak, T. F., and Glenner, G. G. (1975) Isolation of a low-molecular-weight serum component antigenically related to an amyloid fibril protein of unknown origin. *Proc Natl Acad Sci U S A*. **72**, 1473–1476
278. Rosenthal, C. J., Franklin, E. C., Frangione, B., and Greenspan, J. (1976) Isolation and partial characterization of SAA-an amyloid-related protein from human serum. *J Immunol*. **116**, 1415–1418
279. O'Toole, T. J., and Sharma, S. (2023) Physiology, Somatostatin. in *StatPearls*, StatPearls Publishing, Treasure Island (FL), [online] http://www.ncbi.nlm.nih.gov/books/NBK538327/

(Accessed May 26, 2023)

280. van Grondelle, W., Iglesias, C. L., Coll, E., Artzner, F., Paternostre, M., Lacombe, F., Cardus, M., Martinez, G., Montes, M., Cherif-Cheikh, R., and Valéry, C. (2007) Spontaneous fibrillation of the native neuropeptide hormone Somatostatin-14. *J Struct Biol*. **160**, 211–223

281. Ichimata, S., Katoh, N., Abe, R., Yoshinaga, T., Kametani, F., Yazaki, M., Kusama, Y., Sano, K., Uehara, T., and Sekijima, Y. (2022) Somatostatin-derived amyloid deposition associated with duodenal neuroendocrine tumour (NET): a report of novel localised amyloidosis associated with NET. *Amyloid*. **29**, 64–65

282. Van Treeck, B. J., Dasari, S., Kurtin, P. J., Theis, J. D., Nasr, S. H., Zhang, L., Yasir, S., Graham, R. P., McPhail, E. D., and Said, S. (2022) Somatostatin-derived amyloidosis: a novel type of amyloidosis associated with well-differentiated somatostatin-producing neuroendocrine tumours. *Amyloid*. **29**, 58–63

283. Wang, Y., Branicky, R., Noë, A., and Hekimi, S. (2018) Superoxide dismutases: Dual roles in controlling ROS damage and regulating ROS signaling. *J Cell Biol*. **217**, 1915–1928

284. Chattopadhyay, M., and Valentine, J. S. (2009) Aggregation of copper-zinc superoxide dismutase in familial and sporadic ALS. *Antioxid Redox Signal*. **11**, 1603–1614

285. Gruzman, A., Wood, W. L., Alpert, E., Prasad, M. D., Miller, R. G., Rothstein, J. D., Bowser, R., Hamilton, R., Wood, T. D., Cleveland, D. W., Lingappa, V. R., and Liu, J. (2007) Common molecular signature in SOD1 for both sporadic and familial amyotrophic lateral sclerosis. *Proc Natl Acad Sci U S A*. **104**, 12524–12529

286. Basso, M., Massignan, T., Samengo, G., Cheroni, C., De Biasi, S., Salmona, M., Bendotti, C., and Bonetto, V. (2006) Insoluble mutant SOD1 is partly oligoubiquitinated in amyotrophic lateral sclerosis mice. *J Biol Chem*. **281**, 33325–33335

287. Wang, J., Slunt, H., Gonzales, V., Fromholt, D., Coonfield, M., Copeland, N. G., Jenkins, N. A., and Borchelt, D. R. (2003) Copper-binding-site-null SOD1 causes ALS in transgenic mice: aggregates of non-native SOD1 delineate a common feature. *Hum Mol Genet*. **12**, 2753–2764

288. Wang, L.-Q., Ma, Y., Yuan, H.-Y., Zhao, K., Zhang, M.-Y., Wang, Q., Huang, X., Xu, W.-C., Dai, B., Chen, J., Li, D., Zhang, D., Wang, Z., Zou, L., Yin, P., Liu, C., and Liang, Y. (2022) Cryo-EM structure of an amyloid fibril formed by full-length human SOD1 reveals its conformational conversion. *Nat Commun*. **13**, 3491

289. Chattopadhyay, M., Durazo, A., Sohn, S. H., Strong, C. D., Gralla, E. B., Whitelegge, J. P., and Valentine, J. S. (2008) Initiation and elongation in fibrillation of ALS-linked superoxide dismutase. *Proc Natl Acad Sci U S A*. **105**, 18663–18668

290. Tiwari, A., Xu, Z., and Hayward, L. J. (2005) Aberrantly increased hydrophobicity shared by mutants of Cu,Zn-superoxide dismutase in familial amyotrophic lateral sclerosis. *J Biol Chem*. **280**, 29771–29779

291. Johnson, B. S., Snead, D., Lee, J. J., McCaffery, J. M., Shorter, J., and Gitler, A. D. (2009) TDP-43 is intrinsically aggregation-prone, and amyotrophic lateral sclerosis-linked mutations accelerate aggregation and increase toxicity. *J Biol Chem*. **284**, 20329–20339

292. Jiang, L.-L., Zhao, J., Yin, X.-F., He, W.-T., Yang, H., Che, M.-X., and Hu, H.-Y. (2016) Two mutations G335D and Q343R within the amyloidogenic core region of TDP-43 influence its aggregation and inclusion formation. *Sci Rep*. **6**, 23928

293. Neumann, M., Sampathu, D. M., Kwong, L. K., Truax, A. C., Micsenyi, M. C., Chou, T. T., Bruce, J., Schuck, T., Grossman, M., Clark, C. M., McCluskey, L. F., Miller, B. L., Masliah, E., Mackenzie, I. R., Feldman, H., Feiden, W., Kretzschmar, H. A., Trojanowski, J. Q., and Lee, V. M.-Y. (2006) Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science*. **314**, 130–133

294. Kwong, L. K., Uryu, K., Trojanowski, J. Q., and Lee, V. M.-Y. (2008) TDP-43 proteinopathies: neurodegenerative protein misfolding diseases without amyloidosis. *Neurosignals*. **16**, 41–51

295. Jiang, L.-L., Che, M.-X., Zhao, J., Zhou, C.-J., Xie, M.-Y., Li, H.-Y., He, J.-H., and Hu, H.-Y. (2013) Structural transformation of the amyloidogenic core region of TDP-43 protein initiates its aggregation and cytoplasmic inclusion. *J Biol Chem*. **288**, 19614–19624

296. Bigio, E. H., Wu, J. Y., Deng, H.-X., Bit-Ivan, E. N., Mao, Q., Ganti, R., Peterson, M., Siddique, N., Geula, C., Siddique, T., and Mesulam, M. (2013) Inclusions in frontotemporal lobar degeneration with TDP-43 proteinopathy (FTLD-TDP) and amyotrophic lateral sclerosis (ALS), but not FTLD with FUS proteinopathy (FTLD-FUS), have properties of amyloid. *Acta Neuropathol*. **125**, 463–465

297. Lin, W.-L., and Dickson, D. W. (2008) Ultrastructural localization of TDP-43 in filamentous neuronal inclusions in various neurodegenerative diseases. *Acta Neuropathol*. **116**, 205–213

298. Nonaka, T., Masuda-Suzukake, M., Arai, T., Hasegawa, Y., Akatsu, H., Obi, T., Yoshida, M., Murayama, S., Mann, D. M. A., Akiyama, H., and Hasegawa, M. (2013) Prion-like properties of pathological TDP-43 aggregates from diseased brains. *Cell Rep*. **4**, 124–134

299. Arseni, D., Hasegawa, M., Murzin, A. G., Kametani, F., Arai, M., Yoshida, M., and Ryskeldi-Falcon, B. (2022) Structure of pathological TDP-43 filaments from ALS with FTLD. *Nature*. **601**, 139–143

300. Guo, W., Chen, Y., Zhou, X., Kar, A., Ray, P., Chen, X., Rao, E. J., Yang, M., Ye, H., Zhu, L., Liu, J., Xu, M., Yang, Y., Wang, C., Zhang, D., Bigio, E. H., Mesulam, M., Shen, Y., Xu, Q., Fushimi, K., and Wu, J. Y. (2011) An ALS-associated mutation affecting TDP-43 enhances protein aggregation, fibril formation and neurotoxicity. *Nat Struct Mol Biol*. **18**, 822–830

301. Barmada, S. J., Skibinski, G., Korb, E., Rao, E. J., Wu, J. Y., and Finkbeiner, S. (2010) Cytoplasmic mislocalization of TDP-43 is toxic to neurons and enhanced by a mutation associated with familial amyotrophic lateral sclerosis. *J Neurosci*. **30**, 639–649

302. Luquin, N., Yu, B., Saunderson, R. B., Trent, R. J., and Pamphlett, R. (2009) Genetic variants in the promoter of TARDBP in sporadic amyotrophic lateral sclerosis. *Neuromuscul Disord*. **19**, 696–700

303. Holbert, S., Denghien, I., Kiechle, T., Rosenblatt, A., Wellington, C., Hayden, M. R., Margolis, R. L., Ross, C. A., Dausset, J., Ferrante, R. J., and Néri, C. (2001) The Gln-Ala repeat transcriptional activator CA150 interacts with huntingtin: neuropathologic and genetic evidence for a role in Huntington's disease pathogenesis. *Proc Natl Acad Sci U S A*. **98**, 1811–1816

304. Ferguson, N., Becker, J., Tidow, H., Tremmel, S., Sharpe, T. D., Krause, G., Flinders, J., Petrovich, M., Berriman, J., Oschkinat, H., and Fersht, A. R. (2006) General structural motifs of amyloid protofilaments. *Proc Natl Acad Sci U S A*. **103**, 16248–16253

305. Ferguson, N., Berriman, J., Petrovich, M., Sharpe, T. D., Finch, J. T., and Fersht, A. R. (2003) Rapid amyloid fiber formation from the fast-folding WW domain FBP28. *Proc Natl Acad Sci U S A*. **100**, 9814–9819

306. Dyrlund, T. F., Poulsen, E. T., Scavenius, C., Nikolajsen, C. L., Thøgersen, I. B., Vorum, H., and Enghild, J. J. (2012) Human cornea proteome: identification and quantitation of the proteins of the three main layers including epithelium, stroma, and endothelium. *J Proteome Res*. **11**, 4231–4239

307. Klintworth, G. K. (2009) Corneal dystrophies. *Orphanet J Rare Dis*. **4**, 7

308. Venkatraman, A., Dutta, B., Murugan, E., Piliang, H., Lakshminaryanan, R., Sook Yee, A. C., Pervushin, K. V., Sze, S. K., and Mehta, J. S. (2017) Proteomic Analysis of Amyloid Corneal Aggregates from TGFBI-H626R Lattice Corneal Dystrophy Patient Implicates

Serine-Protease HTRA1 in Mutation-Specific Pathogenesis of TGFBIp. *J Proteome Res*. **16**, 2899–2913

309. Munier, F. L., Korvatska, E., Djemaï, A., Le Paslier, D., Zografos, L., Pescia, G., and Schorderet, D. F. (1997) Kerato-epithelin mutations in four 5q31-linked corneal dystrophies. *Nat Genet*. **15**, 247–251
310. Korvatska, E., Henry, H., Mashima, Y., Yamada, M., Bachmann, C., Munier, F. L., and Schorderet, D. F. (2000) Amyloid and non-amyloid forms of 5q31-linked corneal dystrophy resulting from kerato-epithelin mutations at Arg-124 are associated with abnormal turnover of the protein. *J Biol Chem*. **275**, 11465–11469
311. Korvatska, E., Munier, F. L., Chaubert, P., Wang, M. X., Mashima, Y., Yamada, M., Uffer, S., Zografos, L., and Schorderet, D. F. (1999) On the role of kerato-epithelin in the pathogenesis of 5q31-linked corneal dystrophies. *Invest Ophthalmol Vis Sci*. **40**, 2213–2219
312. Sørensen, 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
313. Karring, H., Runager, K., Thøgersen, I. B., Klintworth, G. K., Højrup, P., and Enghild, J. J. (2012) Composition and proteolytic processing of corneal deposits associated with mutations in the TGFBI gene. *Exp Eye Res*. **96**, 163–170
314. Karring, H., Poulsen, E. T., Runager, K., Thøgersen, I. B., Klintworth, G. K., Højrup, P., and Enghild, J. J. (2013) Serine protease HtrA1 accumulates in corneal transforming growth factor beta induced protein (TGFBIp) amyloid deposits. *Mol Vis*. **19**, 861–876
315. Poulsen, E. T., Runager, K., Risør, M. W., Dyrlund, T. F., Scavenius, C., Karring, H., Praetorius, J., Vorum, H., Otzen, D. E., Klintworth, G. K., and Enghild, J. J. (2014)

Comparison of two phenotypically distinct lattice corneal dystrophies caused by mutations in the transforming growth factor beta induced (TGFBI) gene. *Proteomics Clin Appl*. **8**, 168–177

316. Nicholson, A. M., Finch, N. A., Wojtas, A., Baker, M. C., Perkerson, R. B., Castanedes-Casey, M., Rousseau, L., Benussi, L., Binetti, G., Ghidoni, R., Hsiung, G.-Y. R., Mackenzie, I. R., Finger, E., Boeve, B. F., Ertekin-Taner, N., Graff-Radford, N. R., Dickson, D. W., and Rademakers, R. (2013) TMEM106B p.T185S regulates TMEM106B protein levels: implications for frontotemporal dementia. *J Neurochem*. **126**, 781–791
317. Chang, A., Xiang, X., Wang, J., Lee, C., Arakhamia, T., Simjanoska, M., Wang, C., Carlomagno, Y., Zhang, G., Dhingra, S., Thierry, M., Perneel, J., Heeman, B., Forgrave, L. M., DeTure, M., DeMarco, M. L., Cook, C. N., Rademakers, R., Dickson, D. W., Petrucelli, L., Stowell, M. H. B., Mackenzie, I. R. A., and Fitzpatrick, A. W. P. (2022) Homotypic fibrillization of TMEM106B across diverse neurodegenerative diseases. *Cell*. **185**, 1346-1355.e15
318. Schweighauser, M., Arseni, D., Bacioglu, M., Huang, M., Lövestam, S., Shi, Y., Yang, Y., Zhang, W., Kotecha, A., Garringer, H. J., Vidal, R., Hallinan, G. I., Newell, K. L., Tarutani, A., Murayama, S., Miyazaki, M., Saito, Y., Yoshida, M., Hasegawa, K., Lashley, T., Revesz, T., Kovacs, G. G., van Swieten, J., Takao, M., Hasegawa, M., Ghetti, B., Spillantini, M. G., Ryskeldi-Falcon, B., Murzin, A. G., Goedert, M., and Scheres, S. H. W. (2022) Age-dependent formation of TMEM106B amyloid filaments in human brains. *Nature*. **605**, 310–314
319. Fan, Y., Zhao, Q., Xia, W., Tao, Y., Yu, W., Chen, M., Liu, Y., Zhao, J., Shen, Y., Sun, Y., Si, C., Zhang, S., Zhang, Y., Li, W., Liu, C., Wang, J., and Li, D. (2022) Generic amyloid fibrillation of TMEM106B in patient with Parkinson's disease dementia and normal elders. *Cell Res*. **32**, 585–588

320. Jiang, Y. X., Cao, Q., Sawaya, M. R., Abskharon, R., Ge, P., DeTure, M., Dickson, D. W., Fu, J. Y., Ogorzalek Loo, R. R., Loo, J. A., and Eisenberg, D. S. (2022) Amyloid fibrils in FTLD-TDP are composed of TMEM106B and not TDP-43. *Nature*. **605**, 304–309

321. Van Deerlin, V. M., Sleiman, P. M. A., Martinez-Lage, M., Chen-Plotkin, A., Wang, L.-S., Graff-Radford, N. R., Dickson, D. W., Rademakers, R., Boeve, B. F., Grossman, M., Arnold, S. E., Mann, D. M. A., Pickering-Brown, S. M., Seelaar, H., Heutink, P., van Swieten, J. C., Murrell, J. R., Ghetti, B., Spina, S., Grafman, J., Hodges, J., Spillantini, M. G., Gilman, S., Lieberman, A. P., Kaye, J. A., Woltjer, R. L., Bigio, E. H., Mesulam, M., Al-Sarraj, S., Troakes, C., Rosenberg, R. N., White, C. L., Ferrer, I., Lladó, A., Neumann, M., Kretzschmar, H. A., Hulette, C. M., Welsh-Bohmer, K. A., Miller, B. L., Alzualde, A., Lopez de Munain, A., McKee, A. C., Gearing, M., Levey, A. I., Lah, J. J., Hardy, J., Rohrer, J. D., Lashley, T., Mackenzie, I. R. A., Feldman, H. H., Hamilton, R. L., Dekosky, S. T., van der Zee, J., Kumar-Singh, S., Van Broeckhoven, C., Mayeux, R., Vonsattel, J. P. G., Troncoso, J. C., Kril, J. J., Kwok, J. B. J., Halliday, G. M., Bird, T. D., Ince, P. G., Shaw, P. J., Cairns, N. J., Morris, J. C., McLean, C. A., DeCarli, C., Ellis, W. G., Freeman, S. H., Frosch, M. P., Growdon, J. H., Perl, D. P., Sano, M., Bennett, D. A., Schneider, J. A., Beach, T. G., Reiman, E. M., Woodruff, B. K., Cummings, J., Vinters, H. V., Miller, C. A., Chui, H. C., Alafuzoff, I., Hartikainen, P., Seilhean, D., Galasko, D., Masliah, E., Cotman, C. W., Tuñón, M. T., Martínez, M. C. C., Munoz, D. G., Carroll, S. L., Marson, D., Riederer, P. F., Bogdanovic, N., Schellenberg, G. D., Hakonarson, H., Trojanowski, J. Q., and Lee, V. M.-Y. (2010) Common variants at 7p21 are associated with frontotemporal lobar degeneration with TDP-43 inclusions. *Nat Genet*. **42**, 234–239

322. Cruchaga, C., Graff, C., Chiang, H.-H., Wang, J., Hinrichs, A. L., Spiegel, N., Bertelsen, S., Mayo, K., Norton, J. B., Morris, J. C., and Goate, A. (2011) Association of TMEM106B gene polymorphism with age at onset in granulin mutation carriers and plasma granulin

protein levels. *Arch Neurol*. **68**, 581–586

323. Richardson, S. J. (2007) Cell and molecular biology of transthyretin and thyroid hormones. *Int Rev Cytol*. **258**, 137–193

324. Adams, D., Koike, H., Slama, M., and Coelho, T. (2019) Hereditary transthyretin amyloidosis: a model of medical progress for a fatal disease. *Nat Rev Neurol*. **15**, 387–404

325. Ruberg, F. L., Grogan, M., Hanna, M., Kelly, J. W., and Maurer, M. S. (2019) Transthyretin Amyloid Cardiomyopathy: JACC State-of-the-Art Review. *J Am Coll Cardiol*. **73**, 2872–2891

326. Costa, P. P., Figueira, A. S., and Bravo, F. R. (1978) Amyloid fibril protein related to prealbumin in familial amyloidotic polyneuropathy. *Proc Natl Acad Sci U S A*. **75**, 4499–4503

327. Westermark, P., Sletten, K., Johansson, B., and Cornwell, G. G. (1990) Fibril in senile systemic amyloidosis is derived from normal transthyretin. *Proc Natl Acad Sci U S A*. **87**, 2843–2845

328. Planté-Bordeneuve, V., and Said, G. (2011) Familial amyloid polyneuropathy. *Lancet Neurol*. **10**, 1086–1097

329. Schmidt, M., Wiese, S., Adak, V., Engler, J., Agarwal, S., Fritz, G., Westermark, P., Zacharias, M., and Fändrich, M. (2019) Cryo-EM structure of a transthyretin-derived amyloid fibril from a patient with hereditary ATTR amyloidosis. *Nat Commun*. **10**, 5008

330. Steinebrei, M., Gottwald, J., Baur, J., Röcken, C., Hegenbart, U., Schönland, S., and Schmidt, M. (2022) Cryo-EM structure of an ATTRwt amyloid fibril from systemic non-hereditary transthyretin amyloidosis. *Nat Commun*. **13**, 6398

331. Ko, H. S., Uehara, T., Tsuruma, K., and Nomura, Y. (2004) Ubiquilin interacts with ubiquitylated proteins and proteasome through its ubiquitin-associated and ubiquitin-like

domains. *FEBS Lett*. **566**, 110–114

332. Sharkey, L. M., Safren, N., Pithadia, A. S., Gerson, J. E., Dulchavsky, M., Fischer, S., Patel, R., Lantis, G., Ashraf, N., Kim, J. H., Meliki, A., Minakawa, E. N., Barmada, S. J., Ivanova, M. I., and Paulson, H. L. (2018) Mutant UBQLN2 promotes toxicity by modulating intrinsic self-assembly. *Proc Natl Acad Sci U S A*. **115**, E10495–E10504

333. Mori, F., Tanji, K., Odagiri, S., Toyoshima, Y., Yoshida, M., Ikeda, T., Sasaki, H., Kakita, A., Takahashi, H., and Wakabayashi, K. (2012) Ubiquilin immunoreactivity in cytoplasmic and nuclear inclusions in synucleinopathies, polyglutamine diseases and intranuclear inclusion body disease. *Acta Neuropathol*. **124**, 149–151

334. Rutherford, N. J., Lewis, J., Clippinger, A. K., Thomas, M. A., Adamson, J., Cruz, P. E., Cannon, A., Xu, G., Golde, T. E., Shaw, G., Borchelt, D. R., and Giasson, B. I. (2013) Unbiased screen reveals ubiquilin-1 and -2 highly associated with huntingtin inclusions. *Brain Res*. **1524**, 62–73

335. Zeng, L., Wang, B., Merillat, S. A., Minakawa, E. N., Perkins, M. D., Ramani, B., Tallaksen-Greene, S. J., Costa, M. do C., Albin, R. L., and Paulson, H. L. (2015) Differential recruitment of UBQLN2 to nuclear inclusions in the polyglutamine diseases HD and SCA3. *Neurobiol Dis*. **82**, 281–288

336. Deng, H.-X., Chen, W., Hong, S.-T., Boycott, K. M., Gorrie, G. H., Siddique, N., Yang, Y., Fecto, F., Shi, Y., Zhai, H., Jiang, H., Hirano, M., Rampersaud, E., Jansen, G. H., Donkervoort, S., Bigio, E. H., Brooks, B. R., Ajroud, K., Sufit, R. L., Haines, J. L., Mugnaini, E., Pericak-Vance, M. A., and Siddique, T. (2011) Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature*. **477**, 211–215

337. Dao, T. P., Kolaitis, R.-M., Kim, H. J., O'Donovan, K., Martyniak, B., Colicino, E., Hehnly, H., Taylor, J. P., and Castañeda, C. A. (2018) Ubiquitin Modulates Liquid-Liquid Phase

Separation of UBQLN2 via Disruption of Multivalent Interactions. *Mol Cell*. **69**, 965-978.e6

338. Burré, J. (2015) The Synaptic Function of α-Synuclein. *J Parkinsons Dis*. **5**, 699–713

339. Spillantini, M. G., Schmidt, M. L., Lee, V. M., Trojanowski, J. Q., Jakes, R., and Goedert, M. (1997) Alpha-synuclein in Lewy bodies. *Nature*. **388**, 839–840

340. Ghosh, D., Sahay, S., Ranjan, P., Salot, S., Mohite, G. M., Singh, P. K., Dwivedi, S., Carvalho, E., Banerjee, R., Kumar, A., and Maji, S. K. (2014) The newly discovered Parkinson's disease associated Finnish mutation (A53E) attenuates α-synuclein aggregation and membrane binding. *Biochemistry*. **53**, 6419–6421

341. Ruggeri, F. S., Flagmeier, P., Kumita, J. R., Meisl, G., Chirgadze, D. Y., Bongiovanni, M. N., Knowles, T. P. J., and Dobson, C. M. (2020) The Influence of Pathogenic Mutations in α-Synuclein on Biophysical and Structural Characteristics of Amyloid Fibrils. *ACS Nano*. **14**, 5213–5222

342. Conway, K. A., Harper, J. D., and Lansbury, P. T. (1998) Accelerated in vitro fibril formation by a mutant alpha-synuclein linked to early-onset Parkinson disease. *Nat Med*. **4**, 1318–1320

343. Narhi, L., Wood, S. J., Steavenson, S., Jiang, Y., Wu, G. M., Anafi, D., Kaufman, S. A., Martin, F., Sitney, K., Denis, P., Louis, J. C., Wypych, J., Biere, A. L., and Citron, M. (1999) Both familial Parkinson's disease mutations accelerate alpha-synuclein aggregation. *J Biol Chem*. **274**, 9843–9846

344. Serpell, L. C., Berriman, J., Jakes, R., Goedert, M., and Crowther, R. A. (2000) Fiber diffraction of synthetic alpha-synuclein filaments shows amyloid-like cross-beta conformation. *Proc Natl Acad Sci U S A*. **97**, 4897–4902

345. Boyer, D. R., Li, B., Sun, C., Fan, W., Sawaya, M. R., Jiang, L., and Eisenberg, D. S. (2019) Structures of fibrils formed by α-synuclein hereditary disease mutant H50Q reveal new polymorphs. *Nat Struct Mol Biol*. **26**, 1044–1052

346. Sun, Y., Hou, S., Zhao, K., Long, H., Liu, Z., Gao, J., Zhang, Y., Su, X.-D., Li, D., and Liu, C. (2020) Cryo-EM structure of full-length α-synuclein amyloid fibril with Parkinson's disease familial A53T mutation. *Cell Res*. **30**, 360–362

347. Zhao, K., Li, Y., Liu, Z., Long, H., Zhao, C., Luo, F., Sun, Y., Tao, Y., Su, X.-D., Li, D., Li, X., and Liu, C. (2020) Parkinson's disease associated mutation E46K of α-synuclein triggers the formation of a distinct fibril structure. *Nat Commun*. **11**, 2643

348. Boyer, D. R., Li, B., Sun, C., Fan, W., Zhou, K., Hughes, M. P., Sawaya, M. R., Jiang, L., and Eisenberg, D. S. (2020) The α-synuclein hereditary mutation E46K unlocks a more stable, pathogenic fibril structure. *Proc Natl Acad Sci U S A*. **117**, 3592–3602

349. Sun, Y., Long, H., Xia, W., Wang, K., Zhang, X., Sun, B., Cao, Q., Zhang, Y., Dai, B., Li, D., and Liu, C. (2021) The hereditary mutation G51D unlocks a distinct fibril strain transmissible to wild-type α-synuclein. *Nat Commun*. **12**, 6252

350. Araki, K., Yagi, N., Aoyama, K., Choong, C.-J., Hayakawa, H., Fujimura, H., Nagai, Y., Goto, Y., and Mochizuki, H. (2019) Parkinson's disease is a type of amyloidosis featuring accumulation of amyloid fibrils of α-synuclein. *Proc Natl Acad Sci U S A*. **116**, 17963–17969

351. Porcari, R., Proukakis, C., Waudby, C. A., Bolognesi, B., Mangione, P. P., Paton, J. F. S., Mullin, S., Cabrita, L. D., Penco, A., Relini, A., Verona, G., Vendruscolo, M., Stoppini, M., Tartaglia, G. G., Camilloni, C., Christodoulou, J., Schapira, A. H. V., and Bellotti, V. (2015) The H50Q mutation induces a 10-fold decrease in the solubility of α-synuclein. *J Biol Chem*. **290**, 2395–2404

352. Weinreb, P. H., Zhen, W., Poon, A. W., Conway, K. A., and Lansbury, P. T. (1996) NACP, a protein implicated in Alzheimer's disease and learning, is natively unfolded. *Biochemistry*. **35**, 13709–13715

353. Uversky, V. N. (2003) A protein-chameleon: conformational plasticity of alpha-synuclein,

a disordered protein involved in neurodegenerative disorders. *J Biomol Struct Dyn*. **21**, 211–234

354. Rutherford, N. J., Dhillon, J.-K. S., Riffe, C. J., Howard, J. K., Brooks, M., and Giasson, B. I. (2017) Comparison of the in vivo induction and transmission of α-synuclein pathology by mutant α-synuclein fibril seeds in transgenic mice. *Hum Mol Genet*. **26**, 4906–4915
355. Chartier-Harlin, M.-C., Kachergus, J., Roumier, C., Mouroux, V., Douay, X., Lincoln, S., Levecque, C., Larvor, L., Andrieux, J., Hulihan, M., Waucquier, N., Defebvre, L., Amouyel, P., Farrer, M., and Destée, A. (2004) Alpha-synuclein locus duplication as a cause of familial Parkinson's disease. *Lancet*. **364**, 1167–1169
356. Singleton, A. B., Farrer, M., Johnson, J., Singleton, A., Hague, S., Kachergus, J., Hulihan, M., Peuralinna, T., Dutra, A., Nussbaum, R., Lincoln, S., Crawley, A., Hanson, M., Maraganore, D., Adler, C., Cookson, M. R., Muenter, M., Baptista, M., Miller, D., Blancato, J., Hardy, J., and Gwinn-Hardy, K. (2003) alpha-Synuclein locus triplication causes Parkinson's disease. *Science*. **302**, 841
357. Robotta, M., Cattani, J., Martins, J. C., Subramaniam, V., and Drescher, M. (2017) Alpha-Synuclein Disease Mutations Are Structurally Defective and Locally Affect Membrane Binding. *J Am Chem Soc*. **139**, 4254–4257
358. Bernier, G. M. (1980) beta 2-Microglobulin: structure, function and significance. *Vox Sang*. **38**, 323–327
359. Gejyo, F., Odani, S., Yamada, T., Honma, N., Saito, H., Suzuki, Y., Nakagawa, Y., Kobayashi, H., Maruyama, Y., and Hirasawa, Y. (1986) Beta 2-microglobulin: a new form of amyloid protein associated with chronic hemodialysis. *Kidney Int*. **30**, 385–390
360. Dember, L. M., and Jaber, B. L. (2006) Dialysis-related amyloidosis: late finding or hidden epidemic? *Semin Dial*. **19**, 105–109
361. Hoshino, J., Yamagata, K., Nishi, S., Nakai, S., Masakane, I., Iseki, K., and Tsubakihara,

Y. (2016) Significance of the decreased risk of dialysis-related amyloidosis now proven by results from Japanese nationwide surveys in 1998 and 2010. *Nephrol Dial Transplant*. **31**, 595–602

362. Gejyo, F., Yamada, T., Odani, S., Nakagawa, Y., Arakawa, M., Kunitomo, T., Kataoka, H., Suzuki, M., Hirasawa, Y., and Shirahama, T. (1985) A new form of amyloid protein associated with chronic hemodialysis was identified as beta 2-microglobulin. *Biochem Biophys Res Commun*. **129**, 701–706

363. Valleix, S., Gillmore, J. D., Bridoux, F., Mangione, P. P., Dogan, A., Nedelec, B., Boimard, M., Touchard, G., Goujon, J.-M., Lacombe, C., Lozeron, P., Adams, D., Lacroix, C., Maisonobe, T., Planté-Bordeneuve, V., Vrana, J. A., Theis, J. D., Giorgetti, S., Porcari, R., Ricagno, S., Bolognesi, M., Stoppini, M., Delpech, M., Pepys, M. B., Hawkins, P. N., and Bellotti, V. (2012) Hereditary systemic amyloidosis due to Asp76Asn variant β2-microglobulin. *N Engl J Med*. **366**, 2276–2283