Dipeptide Repeat derived from C9orf72 Hexanucleotide Expansions Forms Amyloids or natively unfolded structures in vitro

Laurent Brasseur, Audrey Coens, Jehan Waeytens, Ronald Melki, Luc Bousset

Dipeptide Repeat derived from C9orf72 Hexanucleotide Expansions Forms Amyloids or natively unfolded structures *in vitro*

Laurent Brasseur[1], Audrey Coens[1], Jehan Waeytens[2,3], Ronald Melki[1], Luc Bousset[1]

[1]Institut de Biologie François Jacob, Molecular Imaging Research Center (MIRCen), Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA), Direction de la Recherche Fondamentale (DRF), Laboratoire des Maladies Neurodégénératives, Centre National de la Recherche Scientifique (CNRS), Paris, Fontenay-aux-Roses, F-92265, France

[2] Laboratoire de Chimie Physique, CNRS UMR 8000Université Paris-Sud Orsay France

[3] Structure et Fonction des Membranes Biologiques, Université libre de Bruxelles, Bruxelles, Belgique

Author email adresses:

Laurent Brasseur (lbrasseu@gmail.com)

Audrey Coens (aud_coens@msn.com)

Jehan Waeytens (jehan.waeytens@ulb.ac.be)

Ronald Melki (ronald.melki@cnrs.fr)

Luc Bousset (luc.bousset@cnrs.fr)

Corresponding Authors :

Luc Bousset, Institut de Biologie François Jacob, Molecular Imaging Research Center (MIRCen), Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA), Direction de la Recherche Fondamentale (DRF), Laboratoire des Maladies Neurodégénératives, Centre National de la Recherche Scientifique (CNRS), Paris, Fontenay-aux-Roses, F-92265, France. luc.bousset@cnrs.fr

Ronald Melki, Institut de Biologie François Jacob, Molecular Imaging Research Center (MIRCen), Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA), Direction de la Recherche Fondamentale (DRF), Laboratoire des Maladies Neurodégénératives, Centre National de la Recherche Scientifique (CNRS), Paris, Fontenay-aux-Roses, F-92265, France. ronald.melki@cnrs.fr

Abstract

The abnormal repetition of the hexanucleotide GGGGCC within the C9orf72 gene is the most common genetic cause of both Amyotrophic Lateral Sclerosis (ALS) and Frontotemporal Dementia (FTD). Different hypothesis have been proposed to explain the pathogenicity of this mutation. Among them, the production of aberrant proteins called Dipeptide Repeat Proteins (DPR) from the repeated sequence. Those proteins are of interest, as they are toxic and form insoluble deposits in patient brains. In this study, we characterize the structural features of three different DPR encoded by the hexanucleotide repeat GGGGCC, namely poly-GA, poly-GP and poly-PA. We show that DPR are natively unstructured proteins but that only poly-GA forms *in vitro* fibrillary aggregates. Poly-GA fibrils are of amyloid nature as revealed by their high content in beta sheets. They neither bind Thioflavin T nor Primuline, the commonly used amyloid fluorescent dyes. Remarkably, not all of the poly-GA primary structure was part of fibrils amyloid core.

Introduction

Amyotrophic Lateral Sclerosis (ALS) is a progressive neurodegenerative disease that affects both the central and peripheral nervous system leading to impaired motor functions and a progressive paralysis of the skeletal muscles. Fronto-temporal dementia (FTD), the second cause of dementia after Alzheimer's disease, is also a neurodegenerative disease diagnosed by cognitive and behavioral symptoms. These pathologies share numerous clinical and pathological aspects and are now considered as the two extremities of the same pathological spectrum, rather than distinct entities.

Mutation within the open reading frame 72 of chromosome 9 (C9orf72) is a genetic factor shared by both pathologies [1,2]. The mutation consists of hexanucleotide (GGGGCC) repetition from 25 to 1000 time in the promotor or intron 1 sequence. It is the most common genetic cause for these diseases that account for 16% of familial ALS and 20% of familial FTD patients. It is also reported in 6 - 8 % of ALS sporadic cases [3].

The role of this mutation in the etiology of ALS/DFT is poorly understood. So far, three mechanisms have been proposed to explain it: (1) Haploinsuffisency of the gene C9orf72, possibly encoding for a GTP Exchange Factor [4–7], (2) Toxicity of the hexanucleotide sens and antisens RNA that sequester RNA-binding proteins and other factors into insoluble RNA foci [8–10], (3) Unconventional translation of repeated hexanucleotide into toxic proteins called Dipeptide Repeat Proteins (DPR) [11–16]. These proteins have been shown to be produced independently of any start codon by a rare mechanism called repeat associated non-ATG (RAN) translation [11,13–15]. As GGGGCC and CCGGGG RNA are translated in all alternate reading frames: poly-GA (Glycine-Alanine), poly-GP (Glycine-Proline) and poly-GR (Glycine-Arginine) for sense transcripts or poly-PA (Proline-Alanine), poly-PR (Proline-Arginine) and also poly-GP for antisense transcripts [11,12,14]. Antibodies directed against several DPR stain neuronal inclusions that are p62-positive and TDP-43-negative. Poly-GA seems to be the most abundant DPR in the inclusions [13]. Nonetheless, the distribution of DPR aggregates

does not completely correlate with lesions observed in patient brains [12,13,15,17,18]. It remains possible that small amounts of highly toxic DPR may be sufficient to kill cells. Another hypothesis is that DPR could trigger TDP-43 aggregation, a major pathological mechanism occurring in ALS/FTD [19–23].

So far, very few studies have been done to characterize the biochemical properties of DPR [24,25]. Here, we show that poly-GA, poly-GP and poly-PA DPR are mostly unstructured proteins. Only poly-GA form fibrillary aggregates in-vitro. We show, using FTIR spectroscopy that GAx34 fibrillar assemblies are of amyloid nature albeit they neither bind Thioflavin T nor Primuline, the fluorescent dyes that are considered as markers of amyloid aggregates. Interestingly, not all of the poly-GA repeated sequence was involved in fibrils amyloid core.

Results

The DPR proteins were expressed in *E. coli* and purified by affinity for a 6xhistidine Tag on a Cobalt column (see Materials and Methods section). We produced four different DPR proteins of distinct composition: PAx50, GAx50, GAx34 and GPx24. The different DPR proteins were expressed as fusion proteins with different tags for purification and immunodetection purposes (Fig S1).

The oligomeric state of the DPR proteins, was determined by assessment of their apparent molecular weights by size-exclusion chromatography (SEC). DPR proteins freshly solubilized in HFIP were dried and the resulting film was resuspended in PBS buffer. All DPR were found soluble after this treatment, except GAx50 protein that forms immediately fibrillary aggregates. GAx34 protein in PBS displays two species with apparent molecular weights of 21 and 6,4 kDa (Fig 1A, 1G). GAx34 protein presence in each peak was confirmed by Western-blot analysis using an anti-V5 antibody (Fig 1D). GPx24 protein shows three species (Fig 1B) all of which contain GPx24 protein as confirmed by Western blot analysis (Fig 1E). The principal specie has an apparent molecular weight of 30 kDa, whereas that of the second is 63 kDa. The third species corresponds to aggregated GPx24 protein with an apparent molecular weight over 2000 kDa (Fig 1G). The species corresponding to peak 3 represents monomeric GPx24 protein. Peaks 1 and 2 correspond to multimeric GPx24 protein. PAx50 protein presents an elution profile similar that of GPx24 protein with the major species exhibiting an apparent molecular weight of 105 and 32 kDa and a third species corresponding to aggregated PAx50 protein (Fig 1C). PAx50 protein species with apparent molecular weight 105 and 32 kDa were recognized by an anti-FLAG antibody (FIG 1F). Altogether, our results shows that all three DPR we produced are of mostly monomeric in nature with varying amounts of low and high molecular weight species corresponding to small or larger multimeric species (Fig S1).

We next measured the secondary structure content of the DPR proteins by Circular Dichroïsm (CD). The different DPR proteins were prepared as for SEC (Fig 2A). GAx34, GPx24

and PAx50 proteins possess random coil structures (Fig 2B). Indeed, deconvolution of the CD spectra indicates that over 78% of the polypeptide consists of random coil (Fig 2B).

We next investigated the capacity of the DPR we produced to assemble into higher molecular weight species. We explored various assembly conditions e.g. buffers, temperature and agitation. In all explored conditions, GPx24 and PAx50 proteins remained mostly soluble with a fraction ranging from 10 to 20% of the proteins forming aggregates of amorphous nature (data not shown). In contrast, GAx34 and GAx50 proteins assembled into fibrillar aggregates that bundle together in purification buffer or PBS at 4°C, without shaking, as observed by transmission electron microscopy after negative staining with Uranyl acetate (Fig 3A and 3B). While GAx50 protein assembled within minutes to hours into fibrils, GAx34 protein (100µM) assembly into fibrils was slow. Assembly of the poly-GA proteins was followed using a sedimentation assay and SDS-PAGE analysis (Fig 3C). GAx34 protein (100 µM) assembly into fibrils was complete after 30 days (Fig 3A). Interestingly, GAx34 fibrils do not bind the fluorescent dyes Thioflavin T or Primuline that are known to bind in a quantitative manner to amyloid fibrils such as those α-synuclein forms (Fig 3B).

We next performed FTIR measurements to determine whether fibrillar GAx34 protein is of amyloid nature. Fibrillar GAx34 protein FTIR spectrum clearly shows a shoulder at 1620 cm^{-1}, indicative of the amyloid nature of the aggregates (Fig 4A). Fourier deconvolution of the spectra indicates that 42 % of GAx34 amino acids residues are involved in the β-sheet structures within the fibrillar form of the protein. The 58% remaining amino acid residues are disordered (Fig 4B). The peak centered at 1628 cm^{-1} indicates that GAx34 fibrils have cross-β-sheet content and are therefore of amyloid nature. Surprisingly, our data clearly indicate that only 40 out of the 68 GA residues are involved into the amyloid core of fibrillar GAx34.

Discussion

DPR have been proposed to be key contributors in C9orf72+ pathologies. Little is known about their biochemical properties and the deleterious effect of these proteins is still under debate. Here we have characterized the properties of three DPR proteins of different lengths: GAx34, GPx24 and PAx50. We have shown that, in PBS, these proteins have apparent molecular weights compatible with monomeric or small sized multimeric forms. The three DPR proteins exhibit the characteristics of disordered proteins as assessed by CD measurements.

We did not succeed to assemble GPx24 and PAx50 into fibrils under any of the numerous experimental conditions we tested. Amorphous aggregates were obtained instead. In contrast, GAx34 and GAx50 proteins readily assembled into fibrils at 4°C. This is in agreement with a previous study where synthetic peptides GAx3 and 6 were demonstrated to exhibit the highest propensity to aggregate [24]. This is of particular interest given that most DPR inclusions in patient brains are GA positive [12,13]. FTIR measurements reveal that the fibrillar GAx34 proteins we generated are of amyloid nature, in agreement with previous observations made with DPR with little number of repeats (25). Altogether, the present findings and previous results suggest that GA DPR assemblies of amyloid nature may play a role in neuronal degeneration in ALS/FTD.

The FTIR measurements we performed indicate that only part of GA repeat is involved in the amyloid core of GAx34 fibrils. Whether this observation is relevant to pathology remains to be established given that 400 and over GA repeats are associated to ALS/FTD in normal population. Still, our result might mean that GA aggregates elongate in a way where monomers stack with offset. The repeated nature of their sequence can allow such phenomenon, contrary to other aggregation-prone proteins such as Tau or α-Synuclein [27–31]. This could also account for the poly-GA assemblies with variable width and bifurcation reported previously [26].

Contrary to previous observations, we neither observed Thioflavin T nor Primuline binding to poly-GA fibrils [24,25]. These fluorescent dyes are often used as markers of amyloids assemblies, despite some controversies [32]. The inability of DPR to bind canonical amyloid dyes may be due to the lack of key amino acids within primary structure, in agreement with previous studies showing a requirement for Tyrosine and Leucine for Thioflavin T binding [33,34].

As indicated above, we were not able to generate fibrillar poly-GP and poly-PA. Those DPR are nonetheless detected within neuronal inclusions [12–14]. Thus, they either form under conditions we did not explore or co-assemble with other DPR or aggregate into amorphous structures such as those we generated that are relocated in neurons, as many other proteins, to inclusions containing filamentous poly-GA. The ability of poly-GP and poly-PA at different concentrations to co-assemble into fibrils with poly-GA at different molar ratio was assessed. Amorphous aggregates together with fibrils were obtained as when poly-GP or poly-PA amorphous assemblies are diluted in the presence of fibrillar poly-GA (not shown). This suggests that poly-GP and poly-PA do not co-assemble into fibrils with poly-GA.

Materials and methods

Dipeptide Repeat Proteins (DPR) cloning and purification

DNA sequences encoding a V5 tag followed by 34 repeats of GA or 24 repeats of GP were subclonned in a bacterial expression vector containing an N-teminal 6xHis-Tag and a TEV protease cleavage site (pETM-11 vector, EMBL). Plasmids pAG416-Gal encoding for PAx50 and GAx50 were obtained from Addgene. Proteins were expressed in *E. coli* BL21, and purified on 5 mL Talon column (Clontech®) loaded with Cobalt. The proteins were eluted with a linear gradient of 12 ml from buffer A (20mM Tris pH 7.5, 250mM NaCl, 5mM Imidazole, 1mM βMercaptoethanol, Glycerol 10%, 0.1 mM PMSF) to buffer B (20mM Tris pH 7.5, 250mN NaCl, 250mM Imidazole, 1mM βMercaptoethanol, Glycerol 10%, 0.1 mM PMSF). Eluted fraction were analysed by SDS PAGE, and proteins were quantified spectrophotometrically using a extinction coefficient of $\varepsilon_{GP/GA}$ = 2980 $M^{-}1cm^{-1}$ and ε_{PA} = 4470 $M^{-}1cm^{-1}$. Proteins were assembled into fibrils at 4°C without shaking for at least 30 days.

Size-Exclusion Chromatography (SEC) and Western-Blot analysis

DPR species of different sizes were separated by SEC on a Superose 6 (10/300) column (GE Healthcare) equilibrated in 0.1X PBS pH 7.4 at a flow rate of 0.4 ml/min. Elution was monitored by measuring absorbance at 280 nm. During elution, fractions were collected for Western-Blot analysis using antibodies directed against V5 tag for GAx34 or GPx24 and FLAG tag for PAx50. The Superose 6 column was calibrated using Dextran blue (2000 kDa), Thyroglobulin (669 kDa), Alcohol Deshydrogenase (150 kDa), BSA (66 kDa), and carbonic anhydrase (29 kDa) standards (Sigma).

Circular dichroism

CD spectra of GAx34, GPx24 and PAx50 were recorded at 20°C in a Jasco J810 dichrograph using 0.1 mm pathlength quartz cuvettes (Hellma) containing 20 µL of the protein solutions. Proteins were solubilized in HFIP, dried under an N2 stream before being dissolved in PBS buffer and analyzed. All spectra were normalized to the mean residue weight ellipticity (θ_{MRW}) [deg cm² /dmol] using the equation $\theta(\lambda)_{MRW} = \theta(\lambda)_{deg}/10cnd$ where $\theta(\lambda)_{deg}$ is the recorded spectra in degrees, d is the path length of the cuvette in centimeters, n is the number of amino acid residues, and c is the sample concentration in moles per liter.

Electron microscopy

The morphology of DPR assemblies was examined by transmission electron microscopy (TEM) in a Jeol 1400 transmission electron microscope, following adsorption onto carbon-coated 200 mesh grids and negative staining with 1% uranyl acetate. The images were recorded with a Gatan Orius CCD camera (Gatan Inc., Pleasanton, CA, USA).

Fibrillization kinetics by SDS-PAGE analysis

Protein fibrillization was followed by disappearence over time of monomeric DPR (100 µM) from the supernatant of aliquots withdrawn at different time points from the assembling reaction at 4°C after sedimentation at 100,000 g for 30 min using SDS-PAGE following Coomasie Blue staining. Quantification was done with ImageLab. Kinetic points are the mean of three different experiments ± SEM.

Fourier transform infrared spectroscopy

GA fibrils were centrifuged for 30 min at 100,000 g at 4 °C then extensively washed through three cycles of resuspension and sedimentation with D2O. The spectra were recorded on a Bruker Vertex 70 Fourier transform infrared spectroscopy (FTIR) spectrometer equipped with

a liquid nitrogen-cooled MCT detector. The background consisted of D_2O and water vapour. A total of 100 interferograms were collected with a resolution of 1 cm^{-1}. The sample were loaded on a single reflection ATR chamber (PikeTech). All the spectra were baseline-corrected, smoothed and normalized prior to further data processing. The amide I (1,600–1,700 cm^{-1}) band of the spectra was fitted using a Gaussian species model centered at 1628, 1648,5, 1667 and 1680 cm^{-1}. During the fitting procedure the peak height was free, whereas the width at half-height was maintained at <20 cm^{-1}.

Thioflavin T and Primuline fluorescence in presence of GA and α-Synuclein fibrils

The fibrillary samples were spun at 100,000g for 30 min at 4°C for GA and at 37°C for α-Synuclein. The pellets were resuspended in PBS at 150 µM and then diluted at different concentrations (100, 50, 25 and 10 µM). Samples were incubated with Thioflavin T or Primuline (10 µM). Fluorescence was recorded with a Cary Eclipse spectrofluorimeter (Varian Medical Systems Inc.) using excitation and emission wavelengths set at 440 and 480 nm for Thioflavin T and 400 and 480 nm for Primuline.

Acknowledgments

This work was supported by grants from the Agence Nationale de la Recherche Scientifique, European Commission Joint Programme on Neurodegenerative Diseases (JPND-TransPathND). The present work has benefited from Imagerie-Gif core facility, supported by l'Agence Nationale de la Recherche (ANR-11-EQPX-0029/Morphoscope, ANR-10-INBS-04/FranceBioImaging; ANR-11-IDEX-0003-02/ Saclay Plant Sciences) for access to Electron Microscopes. LB and RM were supported by Equipe FRM (Fondation pour la Recherche Médicale) 2016 (DEQ2016033489).

Legends

Figure 1. Apparent molecular weights of solubilized GAx34, GPx24 and PAx50.

(A – C) Size-exclusion chromatograms on Superose 6 (10/300) column of GAx34, GPx24 and PAx50 (150 µM). Samples were solubilized in HFIP and filtred through 0,2 µm pore filters and loaded onto the column. Peptides elution was followed by absorbancy at 280 nm. Column calibration is shown for Dextran blue (>2000 kDa), thyroglobulin (670 kDa), β-amylase (200 kDa), BSA (66 kDa) and carbonic anhydrase (29 kDa). All spectra were normalized to the highest value. **(D – F)** Western-Blot analysis of colected fractions using antibodies directed against V5 tag for GAx34 or GPx24 and FLAG tag for PAx50. **(G)** Apparent molecular weight of GAx34, GPx24 and PAx50 peaks based on their Ve/V_0 ratio (see supplemental 1).

Figure 2. Secondary structure determination of solubilized GAx34, GPx24 and PAx50.

A Circular dichroism spectra of GAx34 (red line), GPx24 (blue line) and PAx50 (green line). DPR were all solubilized in HFIP before resuspension in PBS and analysis. All spectra were normalized to the mean residue weight ellipticity (θ_{MRW}) [deg cm² /dmol] **B** Secondary structure estimation of solubilized GAx34, GPx24 and PAx50 obtained by deconvolution of the CD spectroscopic measurements. Deconvolution was performed on Dichroweb server using CONTIN algorithm (reference set 7).

Figure 3. DPR made of poly-GA form fibrillar assemblies that do not bind Thioflavin T nor Primuline

(A, B) Negatively stained TEM of GAx34 and GAx50 fibrillar assemblies obtained by incubation at 4°C without shaking. Scale bar, 200 nm **(C)** poly-GAx34 aggregation followed by disappearence over time of monomeric protein (100 µM) from the SDS-PAGE gel. Aliquots of GAx34 were taken during assembly and analyzed by SDS-PAGE followed by Coomasie Blue staining (inset). Quantification was done by comassie signal integration using ImageLab.

Quantification is shown as the average and SEM of three independant experiments. **(D)** Thioflavin T and Primuline (10 µM) binding assay with increasing concnetration of GAx34 fibrils (●) and control amyloid fibrils made of alpha synuclein aSyn (X). Fluorescence intensity was recorded using excitation and emission wavelengths set at 440 and 480 nm for Thioflavin T and 400 and 480 nm for Primuline. Quantification are shown as the average and SEM of three independant experiments.

Figure 4. Deconvoluted FT-IR spectrum of GA fibrils

(A) GAx34 fibrils deconvoluted FT-IR spectrum. GAx34 fibrils were washed 3 times in D_2O with repeated cycles of spinning/resuspension. FTIR spectra (thick line), Fourier deconvolution (thin lines) and curve fit (dashed line) data are presented. FT-IR spectrum was fitted using a Gaussian species model centered at 1628, 1680 (ß-sheets), 1649 (disordered and α helix) and 1667 (loops) cm^{-1}. **(B)** Secondary structure estimation fo GAx34 fibrils obtained by deconvolution of FTIR spectrum. The percent of the Area and the associated number of peptide bonds are listed for each structure assignment.

Supplemental Figure 1. Sequence DPR use in this study

The His tag and TEV cleavage site are common to all DPR studied. An V5 TAG is present upstream to GAx34 and GPx24, while a Myc and FLAG tag are present downstream to GAx50 and PAx50 sequences.

Supplemental Figure 2. Calibration of Superose 6 (10/300) column

A Size-exclusion chromatograms of molecular weights markers (1 : Blue Dextran, 2 : Thryoglobulin, 3 : Alcohol deshydrogenase, 4 : BSA, 5 : Carbonic deshydrogenase). **B** Calibration curve derived from molecular weight standards. The ratio Ve/V_0 relate to the elution volume of the proteins divide by the dead volume (elution volume of Blue Dextran). Molecular weights are presented on a log scale. Exponential regression's equation is presented on the graph.

References

1. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, et al. Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. Neuron. 2011. doi:10.1016/j.neuron.2011.09.011

2. Renton AE, Majounie E, Waite A, Simón-Sánchez J, Rollinson S, Gibbs JR, et al. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron. 2011. doi:10.1016/j.neuron.2011.09.010

3. Marogianni C, Rikos D, Provatas A, Dadouli K, Ntellas P, Tsitsi P, et al. The role of C9orf72 in neurodegenerative disorders: a systematic review, an updated meta-analysis, and the creation of an online database. Neurobiol Aging. 2019. doi:10.1016/j.neurobiolaging.2019.04.012

4. Levine TP, Daniels RD, Gatta AT, Wong LH, Hayes MJ. The product of C9orf72, a gene strongly implicated in neurodegeneration, is structurally related to DENN Rab-GEFs. Bioinformatics. 2013. doi:10.1093/bioinformatics/bts725

5. Ciura S, Lattante S, Le Ber I, Latouche M, Tostivint H, Brice A, et al. Loss of function of C9orf72 causes motor deficits in a zebrafish model of amyotrophic lateral sclerosis. Ann Neurol. 2013. doi:10.1002/ana.23946

6. Therrien M, Rouleau GA, Dion PA, Parker JA. Deletion of C9ORF72 results in motor neuron degeneration and stress sensitivity in C. elegans. PLoS One. 2013. doi:10.1371/journal.pone.0083450

7. Haeusler AR, Donnelly CJ, Periz G, Simko EAJ, Shaw PG, Kim MS, et al. C9orf72 nucleotide repeat structures initiate molecular cascades of disease. Nature. 2014. doi:10.1038/nature13124

8. Donnelly CJ, Zhang PW, Pham JT, Heusler AR, Mistry NA, Vidensky S, et al. RNA Toxicity from the ALS/FTD C9ORF72 Expansion Is Mitigated by Antisense Intervention. Neuron. 2013. doi:10.1016/j.neuron.2013.10.015

9. Lee YB, Chen HJ, Peres JN, Gomez-Deza J, Attig J, Štalekar M, et al. Hexanucleotide repeats in ALS/FTD form length-dependent RNA Foci, sequester RNA binding proteins, and are neurotoxic. Cell Rep. 2013. doi:10.1016/j.celrep.2013.10.049

10. Cooper-Knock J, Walsh MJ, Higginbottom A, Highley JR, Dickman MJ, Edbauer D, et al. Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. Brain. 2014. doi:10.1093/brain/awu120

11. Gendron TF, Bieniek KF, Zhang YJ, Jansen-West K, Ash PEA, Caulfield T, et al. Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. Acta Neuropathol. 2013. doi:10.1007/s00401-013-1192-8

12. Mori K, Arzberger T, Grässer FA, Gijselinck I, May S, Rentzsch K, et al. Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. Acta Neuropathol. 2013. doi:10.1007/s00401-013-1189-3

13. Mori K, Weng SM, Arzberger T, May S, Rentzsch K, Kremmer E, et al. The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. Science (80-). 2013. doi:10.1126/science.1232927

14. Zu T, Liu Y, Bañez-Coronel M, Reid T, Pletnikova O, Lewis J, et al. RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. Proc Natl Acad Sci U S A. 2013. doi:10.1073/pnas.1315438110

15. Ash PEA, Bieniek KF, Gendron TF, Caulfield T, Lin WL, DeJesus-Hernandez M, et al. Unconventional Translation of C9ORF72 GGGGCC Expansion Generates Insoluble

Polypeptides Specific to c9FTD/ALS. Neuron. 2013. doi:10.1016/j.neuron.2013.02.004

16. Mann DMA, Rollinson S, Robinson A, Bennion Callister J, Thompson JC, Snowden JS, et al. Dipeptide repeat proteins are present in the p62 positive inclusions in patients with frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9ORF72. Acta Neuropathol Commun. 2014. doi:10.1186/20515960168

17. Gomez-Deza J, Lee YB, Troakes C, Nolan M, Al-Sarraj S, Gallo JM, et al. Dipeptide repeat protein inclusions are rare in the spinal cord and almost absent from motor neurons in C9ORF72 mutant amyotrophic lateral sclerosis and are unlikely to cause their degeneration. Acta Neuropathol Commun. 2015. doi:10.1186/s40478-015-0218-y

18. Davidson YS, Barker H, Robinson AC, Thompson JC, Harris J, Troakes C, et al. Brain distribution of dipeptide repeat proteins in frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9ORF72. Acta Neuropathol Commun. 2014. doi:10.1186/2051-5960-2-70

19. Chew J, Gendron TF, Prudencio M, Sasaguri H, Zhang YJ, Castanedes-Casey M, et al. C9ORF72 repeat expansions in mice cause TDP-43 pathology, neuronal loss, and behavioral deficits. Science (80-). 2015. doi:10.1126/science.aaa9344

20. Khosravi B, Hartmann H, May S, Möhl C, Ederle H, Michaelsen M, et al. Cytoplasmic poly-GA aggregates impair nuclear import of TDP-43 in C9orf72 ALS/FTLD. Hum Mol Genet. 2017. doi:10.1093/hmg/ddw432

21. Solomon DA, Stepto A, Au WH, Adachi Y, Diaper DC, Hall R, et al. A feedback loop between dipeptide-repeat protein, TDP-43 and karyopherin-α mediates C9orf72-related neurodegeneration. Brain. 2018. doi:10.1093/brain/awy241

22. Baborie A, Griffiths TD, Jaros E, Perry R, McKeith IG, Burn DJ, et al. Accumulation of dipeptide repeat proteins predates that of TDP-43 in frontotemporal lobar degeneration associated with hexanucleotide repeat expansions in C9ORF72 gene. Neuropathol Appl

Neurobiol. 2015. doi:10.1111/nan.12178

23. Nonaka T, Masuda-Suzukake M, Hosokawa M, Shimozawa A, Hirai S, Okado H, et al. C9ORF72 dipeptide repeat poly-GA inclusions promote intracellular aggregation of phosphorylated TDP-43. Hum Mol Genet. 2018. doi:10.1093/hmg/ddy174

24. Flores BN, Dulchavsky ME, Krans A, Sawaya MR, Paulson HL, Todd PK, et al. Distinct c9orf72-associated dipeptide repeat structures correlate with neuronal toxicity. PLoS One. 2016. doi:10.1371/journal.pone.0165084

25. Chang YJ, Jeng US, Chiang YL, Hwang IS, Chen YR. The glycine-alanine dipeptide repeat from C9 or f72 hexanucleotide expansions forms toxic amyloids possessing cell-to-cell transmission properties. J Biol Chem. 2016. doi:10.1074/jbc.M115.694273

26. Guo Q, Lehmer C, Martínez-Sánchez A, Rudack T, Beck F, Hartmann H, et al. In Situ Structure of Neuronal C9orf72 Poly-GA Aggregates Reveals Proteasome Recruitment. Cell. 2018. doi:10.1016/j.cell.2017.12.030

27. Fitzpatrick AWP, Falcon B, He S, Murzin AG, Murshudov G, Garringer HJ, et al. Cryo-EM structures of tau filaments from Alzheimer's disease. Nature. 2017. doi:10.1038/nature23002

28. Falcon B, Zhang W, Schweighauser M, Murzin AG, Vidal R, Garringer HJ, et al. Tau filaments from multiple cases of sporadic and inherited Alzheimer's disease adopt a common fold. Acta Neuropathol. 2018. doi:10.1007/s00401-018-1914-z

29. Zhang W, Falcon B, Murzin AG, Fan J, Crowther RA, Goedert M, et al. Heparin-induced tau filaments are polymorphic and differ from those in alzheimer's and pick's diseases. Elife. 2019. doi:10.7554/eLife.43584

30. Guerrero-Ferreira R, Taylor NMI, Mona D, Ringler P, Lauer ME, Riek R, et al. Cryo-EM structure of alpha-synuclein fibrils. Elife. 2018. doi:10.7554/eLife.36402

31. Li B, Ge P, Murray KA, Sheth P, Zhang M, Nair G, et al. Cryo-EM of full-length α-

synuclein reveals fibril polymorphs with a common structural kernel. Nat Commun. 2018. doi:10.1038/s41467-018-05971-2

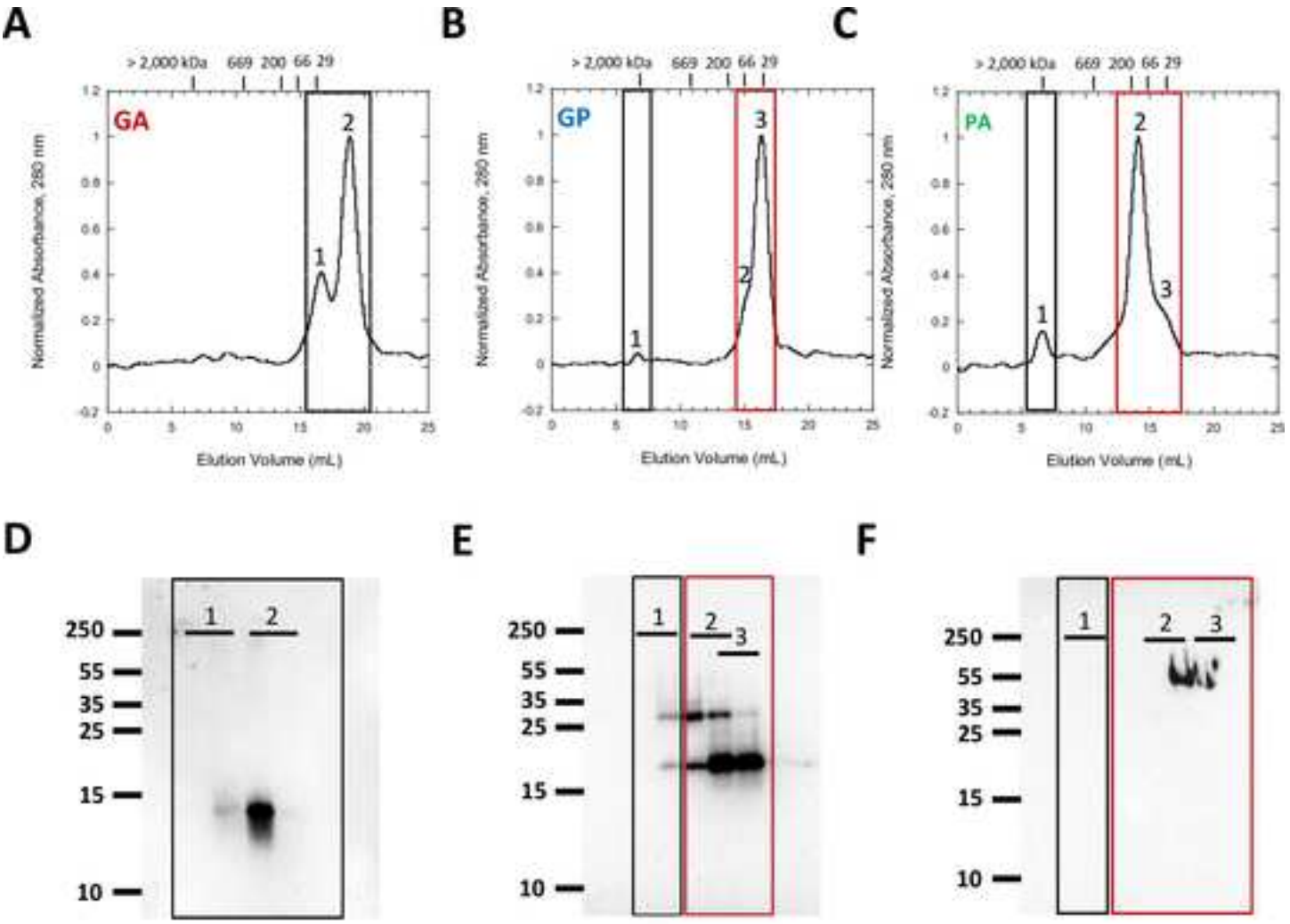
32. Sabaté R, Lascu I, Saupe SJ. On the binding of Thioflavin-T to HET-s amyloid fibrils assembled at pH 2. J Struct Biol. 2008. doi:10.1016/j.jsb.2008.02.002

33. Biancalana M, Makabe K, Koide A, Koide S. Molecular Mechanism of Thioflavin-T Binding to the Surface of β-Rich Peptide Self-Assemblies. J Mol Biol. 2009. doi:10.1016/j.jmb.2008.11.006

34. Wu C, Biancalana M, Koide S, Shea JE. Binding Modes of Thioflavin-T to the Single-Layer β-Sheet of the Peptide Self-Assembly Mimics. J Mol Biol. 2009. doi:10.1016/j.jmb.2009.09.056

Figure 1

Click here to download Figure 16-10-19 ArticleDPR-lb-FIGURES-1.tif

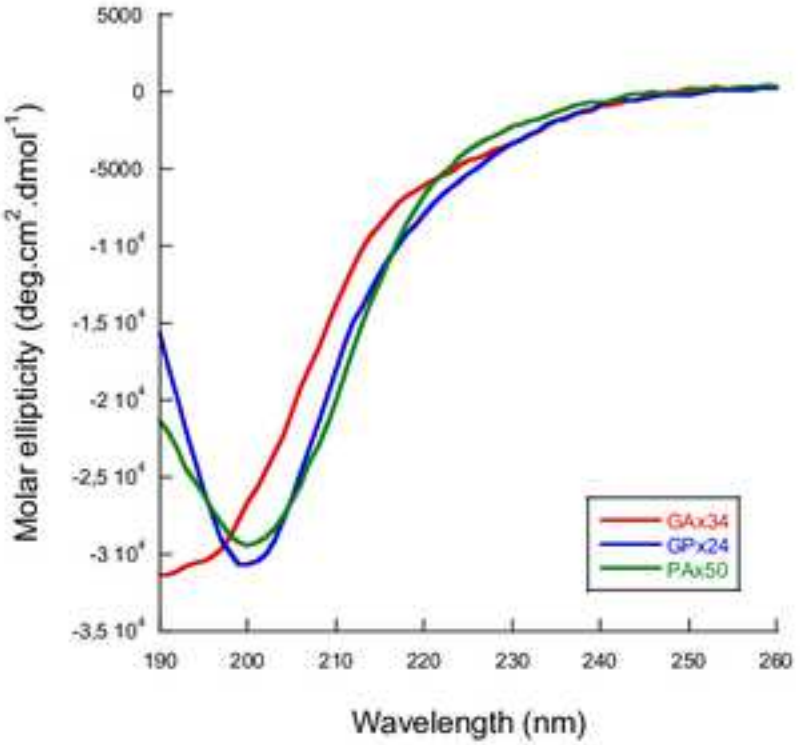


G

Protein	Peak	Apparent MW (kDa)
GAx34	1	21
	2	6.4
GPx24	1	> 2.000
	2	63
	3	30
PAx50	1	> 2.000
	2	105
	3	32

Figure 1. Apparent molecular weights of solubilized GAx34, GPx24 and PAx50.

A



B

	α-helix (%)	β-sheet (%)	Other (%)
GAx34	10	1	89
GPx24	12	2	86
PAx50	20	2	78

Figure 2. Secondary structure determination of solubilized GAx34, GPx24 and PAx50.

Figure 3

Click here to download Figure 16-10-19 ArticleDPR-lb-FIGURES-3.tif

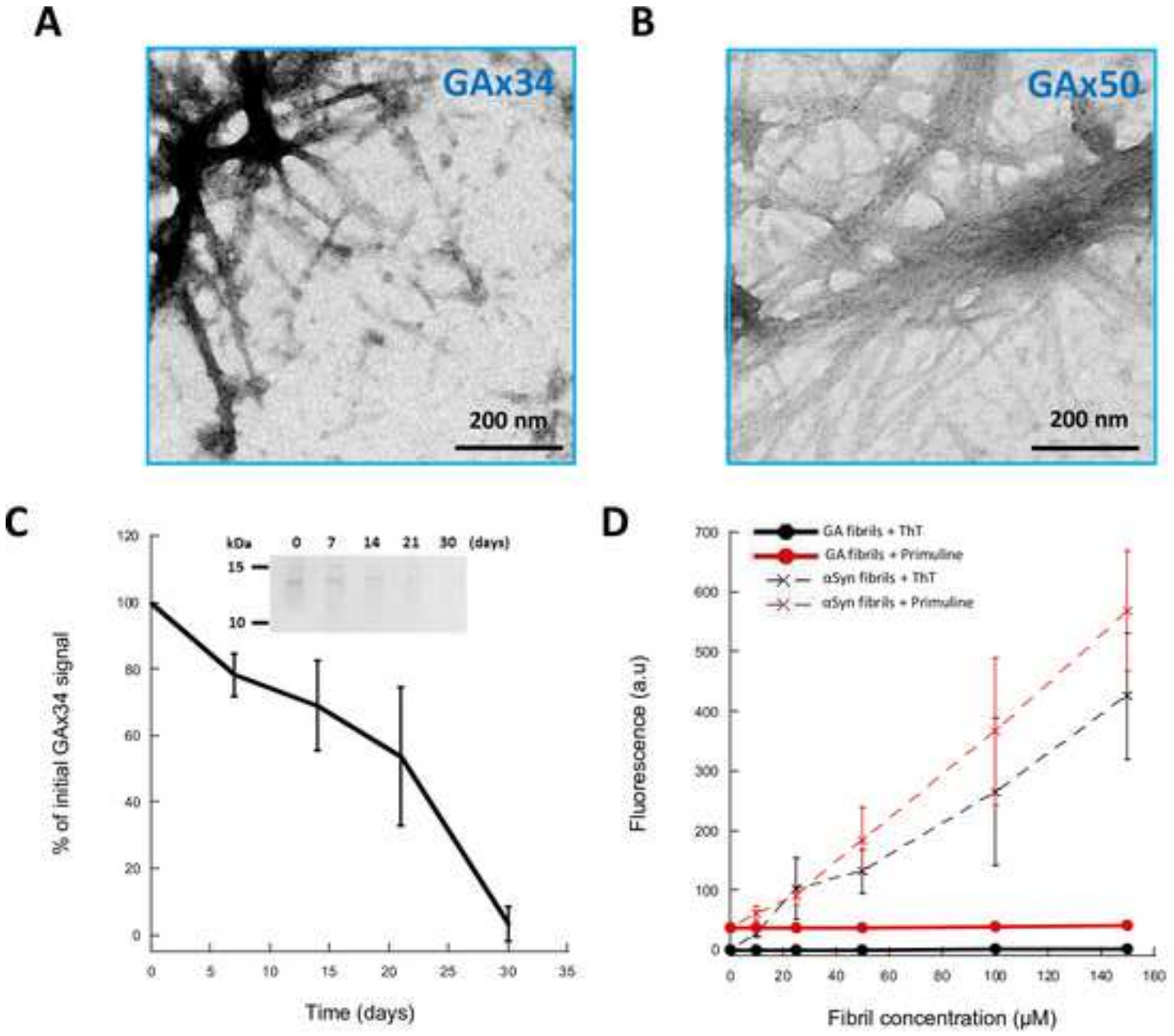
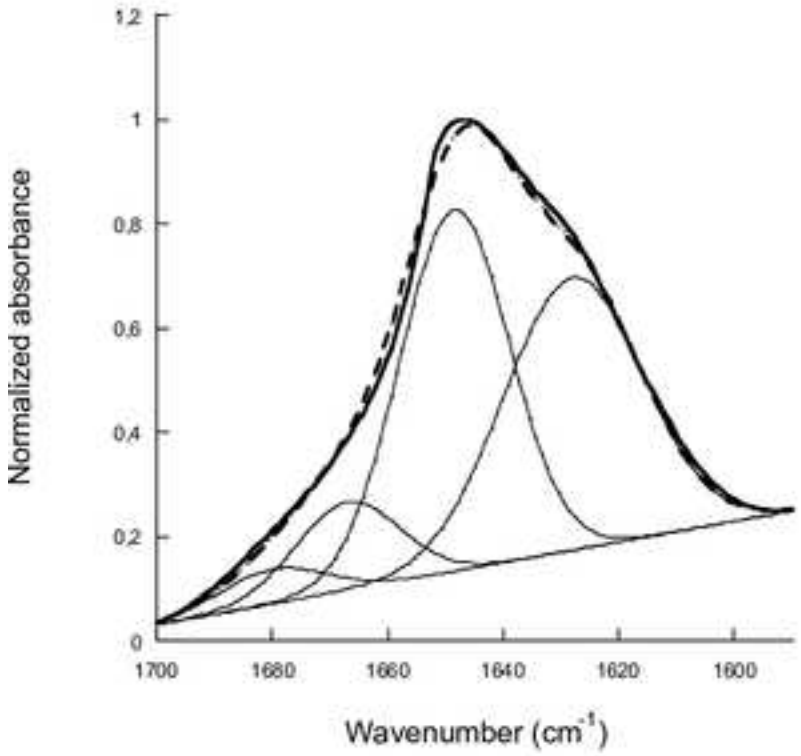


Figure 3. DPR made of poly-GA form fibrillar assemblies that do not bind Thioflavin T nor Primuline

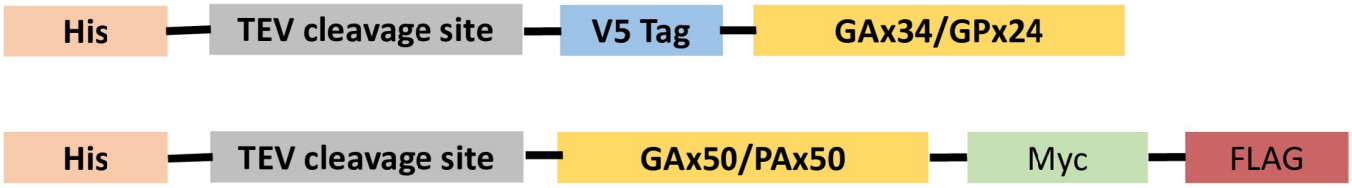
A



B

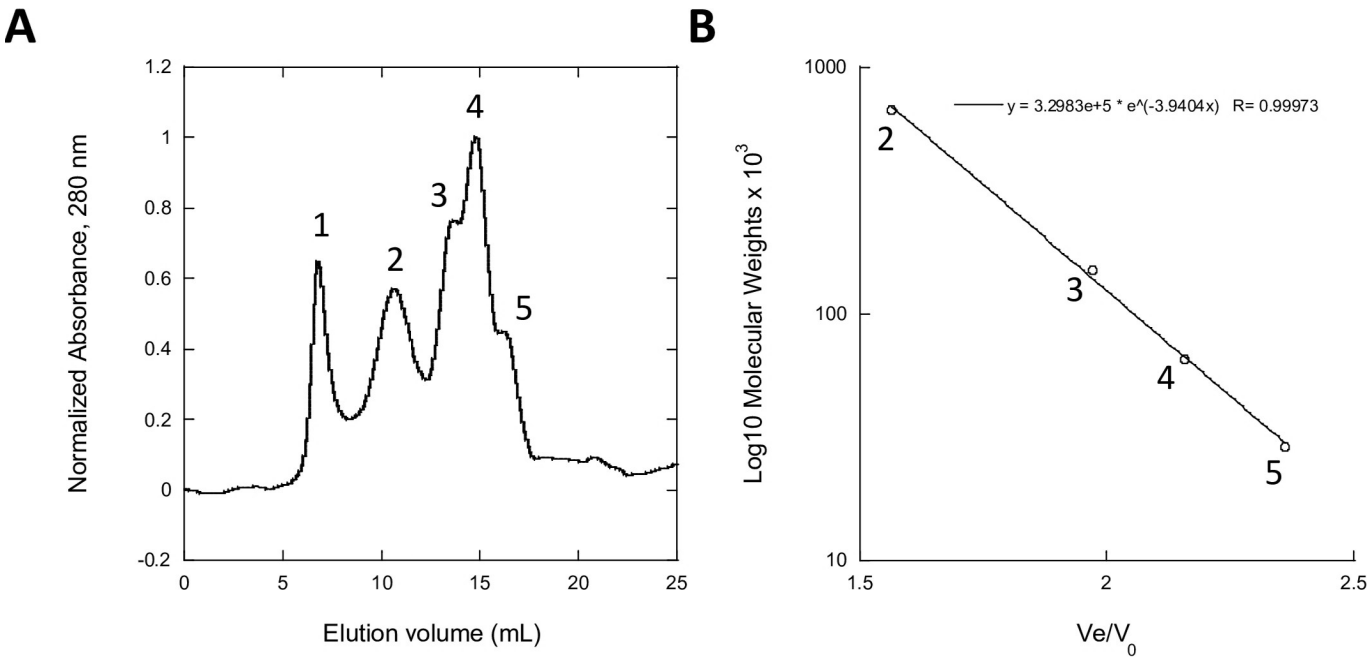
	Structure assignement	Area (%)	Number of peptide bonds
1628 cm^{-1}	β-sheets	38	40
1680 cm^{-1}	β-sheets	4	4
1667 cm^{-1}	Turns	10	11
1649 cm^{-1}	Disordered and α-helix	48	52

Figure 4. Deconvoluted FT-IR spectrum of GA fibrils



DPR	Mw (kDa)
GPx24	9,2
GAx34	9,0
GAx50	12,3
PAx50	14,3

Supplemental Figure 1. Sequence DPR use in this study



Supplemental Figure 2. Calibration of Superose 6 (10/300) column