Enzymatic degradation of RNA causes widespread protein aggregation in cell and tissue lysates

Johan Aarum[1], Claudia P Cabrera[2], Tania A Jones[1], Shiron Rajendran[1], Rocco Adiutori[1], Gavin Giovannoni[1], Michael R Barnes[2], Andrea Malaspina[1,] Denise Sheer[1]

[1]Blizard Institute, Queen Mary University of London, Barts and The London School of Medicine and Dentistry, 4 Newark Street, London, UK

[2]William Harvey Research Institute, Queen Mary University of London, Barts and The London NIHR Cardiovascular Biomedical Research Centre, Charterhouse Square, London, UK

Running title: RNA prevents protein aggregation

Character count: 95,685

Correspondence to:

Dr Johan Aarum
Blizard Institute,
4 Newark Street,
London E1 2AT
Email: j.aarum@qmul.ac.uk
Tel: +44 20 7882-2597

Prof Denise Sheer
Blizard Institute,
4 Newark Street,
London E1 2AT
Email: d.sheer@qmul.ac.uk
Tel: +44 20 7882-2595

ABSTRACT

Most proteins in cell and tissue lysates are soluble. Here, we show that many of these proteins, including several that are implicated in neurodegenerative diseases, are maintained in a soluble and functional state by association with endogenous RNA, as degradation of RNA invariably leads to protein aggregation. We identify the importance of nucleic acid structure, with single-stranded pyrimidine-rich bulges or loops surrounded by double-stranded regions being particularly efficient in this role, revealing an apparent one-to-one protein-nucleic acid stoichiometry. The relationship of these findings to pathological protein aggregation is suggested by our discovery that protein aggregates isolated from brain tissue from Amyotrophic Lateral Sclerosis patients can be rendered soluble after refolding by both RNA and synthetic oligonucleotides. Together, these findings open new avenues for understanding the mechanism behind protein aggregation and shed light on how certain proteins remain soluble.

KEYWORDS

motor neurone disease / protein precipitation / ribonuclease / neurofilament / phase-transition

INTRODUCTION

Under physiological conditions, the majority of cellular proteins exist as soluble and folded entities. After protein synthesis, a complex machinery of cellular chaperones facilitates correct protein folding. Misfolded proteins are then identified and either corrected or, if the correction has failed, removed through autophagy or by the proteasome (Buchberger *et al*, 2010; Rubinsztein, 2006). Failure to resolve the misfolded state of proteins can lead to pathological accumulation and deposition of insoluble protein aggregates, as seen in many neurodegenerative diseases. However, aggregation *per se* is not necessarily a pathological phenomenon as various aspects of this process are also part of normal cellular physiology (David *et al*, 2010; Kaganovich *et al*, 2008; Wallace *et al*, 2015; Walther *et al*, 2015). For example, several proteins, in particular RNA-binding proteins, are recruited into semi-soluble, non-membrane encapsulated organelles such as stress granules and p-bodies (Brangwynne, 2011). These, and similar structures, differ from pathological aggregates in that individual components are dynamically exchanged with the surrounding environment (Andersen *et al*, 2005; Decker & Parker, 2012; Kedersha *et al*, 2005; Spector & Lamond, 2011). Some of the key factors in these transitions are just beginning to emerge and include, for example, the presence of low-complexity regions and/or unstructured regions in participating proteins (Aguzzi & Altmeyer, 2016; Bergeron-Sandoval *et al*, 2016; Kato *et al*, 2012; Wu & Fuxreiter, 2016). RNA also appears to play an important role in the formation of these structures. Indeed, RNA itself undergoes phase transition (Jain & Vale, 2017) and has been shown to promote phase transitions of several proteins (Lin *et al*, 2015; Molliex *et al*, 2015; Zhang *et al*, 2015), including Tau (Zhang *et al*, 2017), and also to inhibit protein aggregation, most notably of Fused in Sarcoma, FUS and TDP-43 (Burke *et al*, 2015; Mann *et al*, 2019; Shelkovnikova *et al*, 2014). Recently, Maharana *et al*., showed that for several prion-like proteins, including FUS, these disparate effects can be explained by the ratio of

protein to RNA; where excess of RNA promotes solubility and decreased amount of RNA induces phase-transition (Maharana *et al*, 2018). Similar results have also been observed for p53 and the prion protein (Kovachev *et al*, 2017; Kovachev *et al*, 2019).

Most neurodegenerative diseases are associated with aggregation of several proteins (Bai *et al*, 2013; Xia *et al*, 2008) and, curiously, the same proteins are frequently found aggregated across multiple diseases. For example, alpha-synuclein can be found aggregated in AD (Hamilton, 2000; Mandal *et al*, 2006), and TDP-43, a hallmark of Amyotrophic Lateral Sclerosis (ALS), is also found aggregated in many cases of AD and PD (Higashi *et al*, 2007; Nakashima-Yasuda *et al*, 2007). This indicates that a common factor might control the soluble state of these proteins. Altered RNA metabolism is a recurring theme in several neurodegenerative disorders (e.g. (Conlon & Manley, 2017; Liu *et al*, 2017; Ramaswami *et al*, 2013). In ALS and Frontotemporal Dementia, this is supported by both disease-causing genetic mutations in several RNA-binding proteins as well as by the generation of potentially toxic RNA species (Fratta *et al*, 2012; Haeusler *et al*, 2014; Kim *et al*, 2013; Kwiatkowski *et al*, 2009; Sreedharan *et al*, 2008; Vance *et al*, 2009). However, the link between altered RNA metabolism and protein aggregation *per se* is largely unexplored. Here, we show that RNA itself is critical in maintaining solubility of several disease-associated and other aggregation-prone proteins in cell and tissue lysates.

RESULTS

Enzymatic degradation of RNA causes protein aggregation

Within a few minutes after adding RNase to a clear cell lysate, it becomes opaque (Expanded View movie 1), suggesting precipitation of material. To investigate if the precipitate contained proteins, we treated lysates prepared from human neurons and mouse brain cortex with a mixture of RNase A and RNase T1, and identified any aggregated proteins by gel electrophoresis (the experimental outline is shown in Fig. 1a). RNase treatment caused a concentration-dependent aggregation of proteins from both samples (Figs. 1b and Expanded View Fig. 1a) and was accompanied by a decrease in both the amount and size distribution of RNA in the lysate (Expanded View Fig. 1b).

The specificity of the ribonuclease was not important for the overall efficiency, as the single-stranded ribonucleases RNase A, T1 and 1f showed similar efficiency to the RNase A/T1 mixture (Fig. 1c). RNase V1, which is specific for double-stranded RNA, also caused protein aggregation but only if EDTA was omitted from the buffer and replaced with Mg^{2+}, (Fig. 1c). RNase V requires Mg^{2+} for its activity. However, DNase I failed to cause protein aggregation in the presence of either EDTA or Mg^{2+} (Fig. 1c). No proteins above background (i.e. in samples without added nucleases) were aggregated when increasing amounts of a ribonuclease inhibitor were added to the lysate together with RNase A (Expanded View Fig. 1c), or when enzymatically or chemically (NaOH) degraded RNA was added to the lysate (Expanded View Fig. 1d).

Mass spectrometry identification of aggregated proteins

We used tandem mass spectrometry (MS) to identify the proteins which are aggregated by degradation of RNA in human neuronal cell lysates. More than 1300 aggregated proteins

were found to be common to two biological replicates (Expanded View Spreadsheet 1). Gene ontology analysis of the dataset indicates an over-representation of cytosolic proteins (FDR q=6.02x10^{-168}), proteins involved in protein binding (FDR q=9.60x10^{-178}), RNA binding (FDR q=2.09x10^{-153}), ATP binding (FDR q=2.43^{-57}), and translational initiation (FDR q=4.94x10^{-57}) (Expanded View Figs. 2a-c and Expanded View Spreadsheet 1). Recently, unstructured, low-complexity regions in several RNA-binding proteins have been shown to mediate protein phase transition (Kato *et al*, 2012). However, both low complexity (LC) and unstructured (US) regions are significantly under-represented in our dataset (Expanded View Figs. 2d and e).

The aggregated proteins include several proteins that are associated with neurodegenerative disease or other proteinopathies, such as huntingtin (HTT), TDP-43, Gelsolin, Lysozyme, the heterogeneous nuclear ribonucleoproteins A2B1 (HNRNPA2B1), HNRNPA1, and in one of the MS replicates, the valosin-containing protein, VCP (DiFiglia *et al*, 1997; Haltia *et al*, 1990; Hirabayashi *et al*, 2001; Kim *et al*, 2013; Kimonis *et al*, 2008; Neumann *et al*, 2006; Pepys *et al*, 1993; Watts *et al*, 2004). We used western blot to validate the mass spectrometry data and to investigate the solubility of other aggregation-prone proteins associated with proteinopathies. HTT, neurofilament heavy chain (NF-H), Tau (MAPT), FUS, TDP-43, HNRNPA1, HNRNPD, RPL7, and actin (ACTB, found aggregated in Hirano bodies in several neurodegenerative diseases (Hirano, 1994)), were selectively aggregated upon RNase A/T1 treatment of human neuronal lysates (Fig. 2a). However, the solubility of poly A binding protein, PABP, an abundant RNA-binding protein not identified in aggregates by mass spectrometry, was unaffected by RNase treatment (Fig. 2a). Similar results were obtained using tissue lysate prepared from mouse cortex (Expanded View Fig. 3). We also detected an approximately 40 kDa Amyloid beta (Aβ)–immunoreactive band in the pellet of RNase-treated lysate (Fig. 2a), possibly representing Aβ oligomers (McLean *et*

al, 1999; Walsh *et al*, 1997) formed from Aβ generated in intracellular vesicles (Rajendran *et al*, 2006). Since the molecular weight of this band is larger than expected, we also examined the aggregation of Aβ in cell lysates prepared from HEK293T cells expressing Aβ fused to GFP. This fusion protein aggregated upon degradation of RNA, while no aggregation was observed for GFP itself (Fig. 2b).

Inhibiting the added RNase activity with an RNase inhibitor abolished the aggregation of HTT, NF-H, and TDP-43 (Fig. 2c).

RNA is required for maintaining the non-aggregated state

We next denatured proteins aggregated by enzymatic degradation of RNA in 6M guanidine hydrochloride (GuHCl) and then attempted to renature them in the presence or absence of total RNA without any prior size fractionation (experimental outline in Fig. 3a). After removal of GuHCl, proteins remained soluble in the presence of RNA in an RNA concentration-dependent manner, while the majority of proteins without RNA re-aggregated (Figs. 3b and c). Remarkably, enzymatic degradation of RNA from the soluble fraction (Sup 1) after renaturing in the presence of RNA caused the proteins to re-aggregate (Fig. 3c, Pel 2). The same principles were also observed for individual proteins (Fig. 3d).

Other nucleic acids, including total *E.coli* RNA and human genomic DNA, efficiently prevented protein aggregation while neither yeast tRNA nor heparin at the same amounts, did (Expanded View Fig. 4a). The effect of DNA is surprising since our data show that proteins in cell lysate are maintained in a soluble state by intact RNA and not by DNA (Fig. 1c).

It is possible that the degradation of RNA causes the aggregation of a few proteins that then co-sequester and aggregate with many other proteins. To investigate this hypothesis, we therefore purified recombinant TDP-43, which forms inclusion bodies in *E.coli* (Capitini *et*

al, 2014; Furukawa *et al*, 2011), under denaturing conditions (6 M guanidine hydrochloride) and assessed the proportions of soluble and insoluble TDP-43 when renatured in the presence or absence of RNA. If co-sequestering were true, then the presence or absence of RNA should not affect the solubility of TDP-43. However, we only obtained soluble TDP-43 when renatured with RNA (Expanded View Fig. 4b). Therefore, at least for TDP-43, the aggregation depends solely on the availability of RNA.

RNA is required for protein activity

Correct folding is required for the proper function of proteins. To investigate if the renatured proteins were functional and not just solubilised, we assessed the ATP-hydrolysing activity of ATP-binding proteins, which represent a large proportion of the aggregated proteins (18%, 236/1312). As the ability to bind ATP depends on the presence of conserved structural motifs (Walker *et al*, 1982), we first investigated if the ability to bind ATP was restored in the presence or absence of RNA. Proteins from human neurons bound to ATP only in the presence of RNA (Expanded View Fig. 4c), and LC-MS/MS characterization of these showed a clear enrichment of ATP binding proteins (58/143, Expanded View Spreadsheet 2). RNA-renatured proteins hydrolysed 100-times more ATP than proteins without RNA (Fig. 3e), and concomitant degradation of RNA by the addition of RNase A/T1 hampered this activity (Fig. 3e). Similar results were obtained with aggregated proteins from Jurkat cells (Expanded View Fig. 4d).

Identification of RNA sequences that are associated with soluble proteins

To identify transcripts associated with soluble proteins in human neuronal lysates we performed native RNA immunoprecipitation (RNA-IP) for Aβ, NF-H and Tau, followed by

sequencing. The majority of associated transcripts were derived from protein coding genes (764 different transcripts associated with Aβ, 1056 with NF-H, and 1119 with Tau). This IP experiment indicates which transcripts are associated with the proteins in cell lysate but give little information as to which regions, e.g. UTRs or coding regions, are involved in the interactions. To increase the mapping resolution, we therefore performed RNA-immunoprecipitation on proteins renatured with pre-fragmented total Jurkat RNA (~100 nucleotides).

Binding peaks were mainly observed in exons (33-53% of all binding sites), in introns (15-21%), and in 3'-UTRs (6-10%), respectively (Fig. 4a, GEO GSE99127). As in the native RNA-IP experiments, the vast majority of the binding sites derived from exons and introns came from protein-coding transcripts (Fig. 4a). Approximately a third of all coding transcripts in the native RNA-IP experiments were also represented by binding peaks in the renatured samples (data not shown). These experiments indicate that soluble Aβ, NF-H, and Tau are associated with a variety of transcripts, particularly with the coding regions of these. To further deduce the principles governing this phenomenon, we next turned to DNA, which, because of the similar renaturation capacity to RNA (Expanded View Fig. 4a), represents a convenient and easily manipulated model system.

First, to identify DNA sequences capable of renaturing aggregated proteins, we performed two sets of isolation and sequencing experiments using proteins renatured with pre-fragmented genomic DNA (approximately 300 bp). In the first set, we aimed to isolate all soluble proteins (~1300) and their associated DNA fragments by capturing them on nitrocellulose membranes. These samples are hereafter referred to as "Total". In the second set of experiments we isolated Aβ and its associated DNA fragments by immunoprecipitation. For both samples the majority (Aβ ~ 86%, Total ~ 90%) of protein-associated fragments were from repetitive DNA regions, in particular Short INterspersed

Elements (SINEs) (including *Alu* repeats, 63 and 49%, respectively) and from regions containing simple repeats (11 and 33%, Fig. 4b, GEO GSE99127).

We next tried to identify common sequence motifs in these datasets. We found a pyrimidine-rich motif (Motif 1, M1, Figs. 4c and Expanded View Fig. 5a) in the "Total" DNA dataset that was similar to motifs found in both the RNA and the DNA samples associated with renatured and immunoprecipitated Aβ (Expanded View Fig. 5b).

Complementary strands are required for efficient protein renaturation

We tested the M1 motif in a 4-repeat form, i.e. M1x4, and a control oligonucleotide, Motif 2, M2, randomly generated to have a similar proportion of G/C, (Expanded View Fig. 5a), for their capacity to renature the aggregated proteins.

As single-stranded (ss) DNA sequences, none of the oligonucleotides supported overall renaturation (Fig. 5a). However, when both the forward and the reverse (For/Rev) oligonucleotides were used together they efficiently supported renaturation in a concentration-dependent manner (Figs. 5a and Expanded View Fig. 6), with a theoretical stoichiometry of ~1:1 (based on the assumption that the average molecular weight of a protein is 50 kDa) The M1x4 oligonucleotide was twice as potent in renaturing the proteins than the M2 control, indicating a sequence preference but not a strict sequence requirement. Similar results were observed for Actin, HTT, TDP-43 and RPL7 (Fig. 5b). Nucleolin (NCL), however, was equally renatured by all configurations (Fig. 5b). Thus, although the majority of proteins in our mixture have a preference for complementary oligonucleotides, individual differences do exist.

Both single- and double-stranded regions are needed for efficient protein renaturation

We next investigated how the size of the DNA oligonucleotides affects their renaturation capacity by decreasing the number of motif repeats. We found that only oligonucleotides with at least three motif repeats (e.g. Mx3) could efficiently renature the proteins (Fig. 5c), indicating that there could be a size requirement. However, experiments on single stranded (ss) segments formed after our standard nucleic acid heat denaturation and rapid cooling step (see Materials and Methods) suggested that there might also be a structural requirement.

To test this possibility, we performed a pre-annealing step in the presence of 100 mM NaCl, which promotes the normal double-stranded form. Only oligonucleotides prepared without NaCl (non-annealed) were able to renature the proteins (Fig. 5d). The amount of NaCl used in the pre-annealing reaction had no effect on protein solubility (1 mM final concentration, data not shown).

Efficient nucleic acid-mediated protein renaturation requires looped, pyrimidine-rich single-stranded regions

Since the M1 oligonucleotide is pyrimidine-rich, we next tested whether there is a preference for pyrimidines in the ss region by examining the renaturing efficiency of oligonucleotides where the ss regions were either composed of 28 consecutive pyrimidines or purines (see Fig. 6a). When assessed for their renaturing capacity, oligonucleotides with a ss region of pyrimidines were superior to the same oligonucleotides containing a ss region of purines (Fig. 6b).

We then examined a series of oligonucleotides, all containing similar double (ds) and single-stranded regions but positioned at different locations along the DNA (see second panel

of Fig. 6c for their theoretical structures). Of these, the most efficient configuration was the ds/ss/ds forms, in particular the 3x-loop or 3x-bulge oligonucleotides (Fig. 6c).

Synthetic DNA oligonucleotides can replicate the effect of endogenous RNA in cell lysates

We next examined if these oligonucleotides could prevent protein aggregation caused by enzymatic RNA degradation in cell lysates. Both single-stranded, and in particular, complementary DNA oligonucleotides almost completely prevented protein aggregation following RNase A/T1 treatment (Figs. 6d and Expanded View Fig. 7).

The unexpected finding that the single-stranded oligonucleotides were also capable of preventing protein aggregation, albeit less than complementary DNA oligonucleotides (Fig. 6d), indicates that they may complement with endogenous RNA in the lysate, forming structural RNA/DNA hybrids. However, RNase H treatment of the lysate had little or no effect (Expanded View Fig. 7), suggesting that any hybrids present are protected.

Renaturation of protein aggregates from Amyotrophic Lateral Sclerosis patient brain

Finally, we isolated aggregated proteins from brain tissue samples of two patients with ALS and investigated whether NF-H within these could be renatured with nucleic acids. Total RNA and the M1x4 oligonucleotides efficiently renatured NF-H from the ALS samples, while samples treated with vehicle show pronounced re-aggregation (Fig. 6e, top and second panel). No NF-H signal was observed in the Jurkat cell samples (used as control) as these cells do not express NFs. These were instead positive for Actin (Fig. 6e). Ribonuclease treatment of the soluble fractions (Sup 1) of proteins renatured with RNA caused re-aggregation of NF-H and Actin, but had no effect on the proteins renatured with the M1x4

DNA oligonucleotides, which remained soluble even after 16 hour at 37°C (Fig. 6e, middle panel). However, NF-H and actin in these DNA-containing fractions were efficiently re-aggregated after a further treatment step with Benzonase (Fig. 6e, lower panel).

DISCUSSION

Our study provides novel insights into how a wide range of proteins are maintained in a soluble state *in vitro*. We show that this heterogeneous set of proteins is prevented from aggregation by association with RNA, and crucially, that the presence of intact RNA is required to maintain them in a soluble state in cell and tissue lysates (Figures 1-3). We also provide evidence that RNA is required for the functionality of some proteins, as exemplified by the pronounced difference in ATP hydrolysing activity between proteins renatured in the presence or absence of RNA (Fig. 3e and Expanded View Fig. 4c). Finally, we provide a tantalizing link between these observations and human disease by demonstrating the efficient renaturation of insoluble proteins isolated from human ALS brain tissue.

Many proteins are known to bind RNA, for example the well characterised hnRNP family of proteins involved in RNA splicing (Swanson & Dreyfuss, 1988). Enzymatic RNA degradation causes the aggregation of a consistent set of proteins from neuronal cell lysates. Although there is a significant overrepresentation of RNA-binding proteins amongst them (354 of 1312 proteins), the vast majority lack conventional RNA-binding domains (Expanded View Fig. 2b). This raises at least two possible explanations for the variety of aggregated proteins. First, either most of the proteins associate with RNA through non-conventional RNA-binding regions, or alternatively, the degradation of RNA initiates the aggregation of a few RNA-associated proteins (i.e. seed proteins) which then sequester a large number of other proteins, irrespective of their association with RNA. While a combination of these two

scenarios is likely, and both are supported by the literature (Castello *et al*, 2012; Olzscha *et al*, 2011), our data favour the first possibility as the major driving force of protein aggregation mediated by enzymatic RNA degradation. This is because *in vitro* renaturing of the proteins, in the presence of the correct nucleic acids, efficiently solubilises the vast majority of them (Figs 3, 5 and 6). Should, however, a small set of RNA-binding seed proteins cause the aggregation of the majority, only these would be expected to be solubilised by RNA, while most of the other proteins should be unaffected (if the RNA-binding seed proteins are not also efficient chaperones).

Fundamentally, and based on the generalisation that the average molecular weight of a protein is 50 kDa, we find that in order to achieve efficient refolding, the nucleic acids and the proteins need to be present in equimolar quantities [i.e. 30 µg of proteins (~600 pmol) is efficiently renatured by 600 pmol of oligonucleotides, see e.g. Fig. 5]. Such a stoichiometry suggests that each protein needs to be associated with one oligonucleotide to maintain solubility. These figures are in line with what has been observed for other aggregate-prone proteins such as the prion protein and FUS (Kovachev *et al*, 2019; Maharana *et al*, 2018). If seed-driven co-precipitation was a major force, a much lower oligo-to-protein ratio would be expected, as only the few seed proteins needs to be associated with the nucleic acid.

The cellular proteome contains many polyanion-binding proteins which are critical for proper function of cells and organisms (reviewed in (Jones *et al*, 2004)). With respect to protein stability and aggregation, the effect of polyanions varies, sometimes in opposing ways, for different proteins, or even between different polyanions on the same protein. As an example of these varying and conflicting effects, aggregation and pathological conversion of the prion protein is stimulated by the presence of RNA and possibly by endogenous proteoglycans, but is inhibited by heparin and exogenous forms of proteoglycans (Caughey & Raymond, 1993; Deleault *et al*, 2003; Vieira *et al*, 2014; Wong *et al*, 2001). Similarly, FUS

aggregation is either inhibited or promoted by the presence of presumably different species of RNA (Lin *et al*, 2015; Shelkovnikova *et al*, 2014). Tau aggregation on the other hand has been reported to be stimulated by RNA (Kampers *et al*, 1996; Zhang *et al*, 2017), while the aggregation of Aβ has been reported to be inhibited (Mathura *et al*, 2005; Takahashi *et al*, 2009), a finding also confirmed here. Thus, it is clear that polyanions, at least *in vitro*, have diverse and sometimes paradoxical effects on protein structure and solubility, possibly depending on the protein / polyanion concentration (Kovachev *et al*, 2017; Kovachev *et al*, 2019; Maharana *et al*, 2018).

Consistent with these findings, our data show a varying capability of different nucleic acids to promote protein renaturation and maintain solubility, and strikingly, this effect is not directly determined by sequence, but primarily by structure (Figs. 5 and 6). This is in line with recent findings that highly structured RNAs had more interactions with proteins than RNAs with low structural complexity (Sanchez de Groot *et al*, 2019). In our study, the most efficient renaturing and aggregate-preventing nucleic acids consist of pyrimidine-rich loops or bulges interspersed by stretches of double-stranded regions, e.g. nucleic acids with a ds/ss/ds configuration (Figs. 6b-c). In addition, the M1 oligonucleotides prevent protein aggregation when both strands are provided, while the same amount (in moles) of ss oligonucleotides (i.e. twice the amount of the ds oligo) fails to do so (Figs. 5a and b). Similarly, when the two strands are pre-annealed, i.e. seemingly forming a perfect ds strand, the renaturing effect is all but lost (Fig. 5d). These results also suggest that the anionic charge provided by the nucleic acids (mainly by the phosphate backbone) cannot be its sole contribution, as the same amount of charge is provided by the two different configurations. This notion is further supported by the much higher efficiency observed for oligonucleotides with a high proportion of pyrimidines in the single-stranded region as compared to their purine-rich counterparts (Fig. 6). Indeed, the preference for pyrimidines over purines suggests

that the bases themselves, and not solely the negative sugar backbone *per se*, are important for the solubilising effect. Such structural dominance may also explain some of the conflicting effects observed with polyanions on protein solubility e.g. (Lin *et al*, 2015; Shelkovnikova *et al*, 2014). More generally, and as also suggested by others (Zhang *et al*, 2015), such disparate effects point to the intriguing possibility that some nucleic acids may not only prevent protein aggregation but may actually promote/induce the same.

The structure- and base-dependent principles summarised above explain the majority of the observed effects, but we do see deviations from these rules for individual proteins. For example, while nucleolin (NCL) clearly requires nucleic acids for efficient renaturation, similarly efficient renaturation is achieved with either single-stranded or looped oligonucleotides (Fig. 5b). Thus, it is likely that other structures, besides the highly efficient looped/bulge configuration described here, can efficiently renature individual or groups of proteins. Indeed, bacterial 23S ribosomal RNA and nucleic acid homopolymers (e.g. single-stranded poly Ts) have recently been shown to assist renaturation/refolding of a small number of proteins, an effect suggested to derive from the nucleic acids functioning as non-protein molecular chaperones (Chattopadhyay *et al*, 1996; Docter *et al*, 2016). Although our data support the general conclusions of these studies, we suggest that the solubilising nucleic acids are integral parts of these complexes with roles beyond traditional chaperones, whose effects are usually energy dependent and mediated through transient interactions of importance mainly for initial folding. This view is backed by our finding that ATP-hydrolysing proteins require the continuous presence of RNA for activity *in vitro*, and similarly, enzymatic degradation of RNA from proteins renatured in the presence of RNA causes them to re-aggregate. Of note, all our renaturing experiments were performed in the absence of traditional energy sources such as ATP.

The main feature in solubilising genomic DNA is interspersed repeat elements, in particular SINE elements and low complexity regions. This is true both for individual aggregation-prone proteins (Aβ) as well as for the majority of the proteins, i.e. Total (Fig. 4b). While some of these may represent true associations, it is plausible that we are observing a pure *in vitro* effect. The argument for this is that similar, but divergent, repeats are likely to readily form the ds-ss-ds structures found here to be highly efficient in protein renaturation.

What is the nature of the protein-RNA associations? Following on from our arguments above, our data support a form of interaction where the permissive nucleic acid (RNA) is an integral but non-covalent part of certain proteins or protein complexes. To be biologically relevant, these arrangements are likely to be dynamic, with proteins and RNA being changeable, i.e. the same RNA may associate with different proteins and *vice versa*. Indeed, such interchangeability is supported by our *in vitro* sequencing and oligonucleotide characterisation studies as individual proteins are found associated with different nucleic acids, and different proteins are found associated with the same nucleic acids (Fig. 4). Similar promiscuity is likely to exist also in cells (Maharana *et al*, 2018), as it is unlikely that the solubility of a particular protein is maintained only by any one particular RNA.

The protein-RNA associations are influenced by nucleic acid structure, as discussed above, but are also likely to be determined by post-translational modifications of the proteins. Indeed, the capacity of several RNA-binding proteins to undergo phase-transition *in vitro* is inhibited by arginine methylation, e.g. FUS (Hofweber *et al*, 2018; Qamar *et al*, 2018) or serine phosphorylation, e.g. TDP-43 (Wang *et al*, 2018). More specifically, certain arginine/serine rich proteins (e.g. SRSF Protein Kinase 1) require either association with RNA or serine phosphorylation to be soluble in cell lysate (Nikolakaki *et al*, 2008). However, in addition to such post-translational modification it is likely that associations with other

proteins or macromolecules could replace the requirement for RNA. In support of this observation, for some of the proteins we examine here, like HTT and RPL7, RNA degradation causes substantial aggregation with little or no soluble protein left in the supernatant (Figs. 2a and c). For others, e.g. NF-H, FUS and TDP-43, only a proportion appears to aggregate with varying amounts still detected in the soluble fraction, even after prolonged RNA degradation.

It is clear that RNA has several important biological roles besides its conventional function as a template for translation and assisting in protein synthesis. These include acting as post-transcriptional regulators (e.g. miRNA), as structural scaffolds (e.g. ncRNA and rRNA), as cofactors (primers for DNA replication) and as enzymatic entities, i.e. ribozymes. The data presented suggest that RNA has an additional role in maintaining proteins in cell and tissue lysate in a soluble and presumably functional form.

Finally, our data indicate that alterations to RNA could contribute to pathological protein aggregation, providing a mechanistic rationale for the observed aggregation of the same proteins across multiple diseases.

MATERIALS AND METHODS

Cells and ALS samples

Neurons were differentiated from human neural stem cells (hNP1 cells) by withdrawal of basic FGF for 6 days, as described (Jones *et al*, 2011). The majority (>95%) of the cells differentiate into Map2- and β III-tubulin -positive cells within 6 days (Jones *et al*, 2011).

Jurkat T cells and HEK 293T (CRL-3216, ATCC) were maintained in RPMI (21875-034, Life Technologies) supplemented with 10% FCS (Life Technologies) and 1 x Penicillin-Streptomycin (15070063, Life Technologies). All cells were maintained at 37°C in 5% CO_2.

Cortices from day 16-21 C56BL mice were dissected at room temperature, rolled on filter paper to remove most of the meninges and immediately frozen on dry ice and stored at -80^{o}C until use.

The ALS brain samples were obtained from The Netherlands Brain Bank (NBB), Netherlands Institute for Neuroscience, Amsterdam (open access: www.brainbank.nl) under ethical permission 2009/148. All material was collected from donors for or from whom a written informed consent for a brain autopsy and the use of the material and clinical information for research purposes had been obtained by the NBB. Samples were obtained from the precentral gyrus of two ALS donors, both male, 62 and 71 years of age, with an ALS diagnosis confirmed neuropathologically by the detection of TDP-43 inclusions, particularly in the spinal cord. The study of proteins from human ALS tissue has been granted to AM by the London - City & East Research Ethics Committee, reference number: 09/H0703/27.

Enzymes and reagents

RNase T1 (AM2280), RNase V1 (AM2275), RNase A/T1 cocktail (EN0551), DNase I (2222), and Yeast t-RNA (15401-011) were from Thermo Fisher Scientific. RNase A (R4642), Sodium acetate (S7899), and EtOH were from Sigma. RNase 1f (M0243) were from NEB. Heparin (07980) was from Stemcell Technologies. Guanidine hydrochloride (BP178-1) was from Fisher Scientific.

Preparation of cell free lysates from neurons and mouse cortex

Differentiated neural stem cells were detached by trypsin (0.5%, Life Technologies) and collected in RPMI medium with 10% FCS (Life Technologies). Cells were pelleted by centrifugation and washed twice in ice-cold PBS (14190-094) before being lysed in four cell-pellet volumes of either Lysis Buffer 1 [20 mM Tris-HCl pH 7.5, 150 mM NaCl, 3 mM EDTA, 1% Triton X-100, 0.5% Na-Deoxycholate, 1X protease inhibitors cocktail (Roche), 1 mM DTT] or Lysis Buffer 2 [20 mM Tris-HCl pH 7.5, 150 mM NaCl, 1.5 mM $MgCl_2$, 1% Triton X-100, 0.5% Na-Doxycholate, 1X EDTA-free protease inhibitors cocktail (Roche), 1 mM DTT]. Most experiments were performed in Lysis Buffer 1, except when DNase I or RNase V1 treatment was performed (Fig. 1d), in which case Lysis Buffer 2 was used. Lysed cells were sonicated (Bioruptor, Diagenode) at maximum setting for 5 seconds on ice and centrifuged at 21.000 x g for 30 min at 4^{o}C. The supernatant was filtered through a 0.1 µm syringe filter (Santa Cruz, sc-358809) into new tubes and the protein concentration determined with the BCA kit (Thermo Fisher) according to the manufacturer's instructions. Lysates were diluted in Lysis Buffer-1 or -2 to 2-4 µg/µl and treated as described below.

Mouse cortical tissue was thawed on ice and disrupted in cold PBS using a 1 ml pipette tip. Disrupted tissue was washed 3 times in PBS before being lysed in Lysis Buffer 1 and prepared as described above for human neurons.

Ribonuclease treatment and isolation of aggregated proteins

Typically, 150-400 µg cell lysates at 2-4µg/µl were mixed with indicated amounts of ribonucleases, DNase I, or Vehicle (50% Glycerol in 20 mM Tris-HCl pH 7.5) and incubated at 37°C for one hour, shaking at 1,200 rpm for 5 seconds every two minutes. Samples were then centrifuged at 21.000 x g for 15 min at +4°C and the supernatants removed and saved for analysis. The pellets were washed twice in 500 µl RIPA buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.5% Na-deoxycholate, 0.1% SDS, 1% Triton X-100) at room temperature and dissolved in 20 mM Tris-HCl pH 7.5, 2% SDS, 8M Urea by sonication (Bioruptor, Diagenode) at maximum setting for 5 min at room temperature. Samples for SDS-PAGE analysis were mixed with 4X LDS Loading Buffer (Life Technologies) supplemented with DTT (Sigma) to 100 mM final concentration and heated for 10 min at 70°C before being loaded on SDS-PAGE gels (Life Technologies).

Video recording of RNase-treated cell lysate

Jurkat T-cell lysate at 10 mg/ml, prepared as described above, was divided into two quartz cuvettes at room temperature with or without 5 µl RNase A/T1 or Vehicle and mixed. Recording was done with a Canon digital camera and started immediately (time 0) taking 60 frames/second for a total of 30 minutes.

Immobilisation of RNase A

100 µg RNase A at 1 µg/µl was coupled to Tosyl activated magnetic beads (Life Technologies) for 20 hours at 37°C according to the manufacturer's instructions. After quenching and washing, the coupled RNase A was re-suspended in 0.1% BSA in PBS and

kept at +4°C until use. Approximately 50% activity remained after coupling, as determined on yeast tRNA using the RiboGreen kit (Life Technologies).

Inhibition of RNase A and addition of pre-hydrolysed RNA

RNase A inhibition: 200 µg lysate was mixed with 0.1 µl RNase A (~3 mg/ml) and increasing concentrations of RNasin (Promega), as indicated. Hydrolysis of RNA: 40 µg of total RNA in TE-buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA) was incubated with 10 µl immobilised RNase A for 1 hour at 37^{o}C. RNase A was removed by magnetic separation and the hydrolysed RNA mixed with 120 U RNasin and kept on ice until used. Alternatively, 40 µg total RNA in 0.1 M NaOH was incubated at 85^{o}C for 1 hour and then adjusted to pH 7.5 with 1 M Tris-HCl pH 7.0. RNase A digested and NaOH hydrolysed RNA was then added to 200 µg of neuronal lysate, prepared as outlined above, and incubated at 37^{o}C for one hour. Aggregated and soluble proteins were collected as before and analysed by SDS-PAGE.

Cloning and use of TDP-43 and Aβ

All PCR reactions were performed with Q5 polymerase (New England Biolabs) according to the manufacturer's instructions. Human Aβ 1-40 was PCR amplified from full length APP (Origen, #RC209575) using Abeta_XhoI_F and Abeta_BamHI_R primers (Expanded View Table 1), purified and then cleaved with Xho I and Bam HI (both NEB). After further purification, the fragment were ligated into the Xho I and Bam HI sites of pEGFP-N3 (Clontech), creating Aβ fused in frame to the N-terminus of GFP. HEK293 cells, plated at a density of 0.2x10^{6} cells/well in a 24 well plate, were transfected with Aβ-GFP or empty vector using Fugene HD (Promega). For each well we used 0.6 µg DNA and 2 µl Fugene HD in a total volume of 30 µl OptiMEM (Life Technologies). Cells were harvested 48 hours after

transfection and washed in PBS and then either stored at -80°C or used directly. Thawed or fresh cells were lysed in 80 µl Lysis Buffer 1 as described above and treated with RNase A/T1 for one hour at 37°C. Aggregated proteins were collected by centrifugation and samples processed and analysed by SDS-PAGE as described above.

Human TDP-43 was PCR amplified with TARDBP BspHI and TARDBP Not I primers (Expanded View Table 1) from cDNA, prepared from Jurkat RNA using Superscript II (Thermo Fisher Scientific) according to the manufactures instructions. Purified product were cleaved with Nco I and Not I and cloned into pA4D5-8mRFP (Markiv *et al*, 2011), creating TDP-43 with an C-terminal His tag. Ligated plasmids were transformed into BL21 (DE3) cells (C2527, NEB) and correct clones verified by Sanger sequencing. A single colony was grown overnight at 37°C in LB in the presence of 2.5 % glucose and ampicillin (50 µg/ml). We found that the presence of glucose, which further suppresses the T7 promoter in pA4D5-8mRFP, was critical to allow the expression of TDP-43. Overnight culture was diluted 20 times in fresh LB medium with glucose (2.5 %) and ampicillin (50 µg/ml) and grown at 37°C until an OD_{600} of approximately 0.8. The bacteria were then pelleted by centrifugation and re-suspended in fresh LB supplemented with 1 mM IPTG (GEN-S-02122-5, Generon) to induce expression, and left shaking at 250 rpm at RT for 2 hours. Recombinant TDP-43 was purified using the Ni-NTA SpinColumn purification kit (31014, Qiagen), essentially according to the manufacturer's instructions. Briefly, bacteria were lysed in Buffer 1 (6 M GuHCl, 0.1 M NaPhosphate, 10 mM Tris-HCl pH 8.0) by sonication, and insoluble debris cleared by centrifugation (21.000 x g, 30 min). Spin column-captured proteins were washed twice in Buffer 1 at pH 6.3, and trice in Buffer 1 at pH 4.5 and then eluted in 6 M GuHCl, pH 2.0. Purity was determined by separating a TCA-precipitated (to remove GuHCl) sample on a SDS-PAGE gel followed by coomassie staining, and were judged to be <80 % (data not shown). 5 µg recombinant TDP-43 was used for renaturation, in the presence or absence of

various amounts of total Jurkat RNA, using the procedure for "renaturation through dialysis" described below. To quench released Ni^{2+} in the protein samples before renaturation, EDTA was added to a final concentration of 10 mM before the addition of DTT (see procedure below).

Nucleic acid-mediated renaturation by dialysis

Proteins were isolated from neuronal or Jurkat cell lysate by RNase A/T1 treatment and centrifugation. Pelleted proteins were dissolved in 50 µl of denaturation buffer (20 mM Tris-HCl pH 7.5, 6 M Guanidine hydrochloride, 1 % Triton X-100, 20 mM DTT) and sonicated for 5 min at room temperature. The protein concentration was determined with the BCA kit (Thermo Fisher) and diluted to 0.4 µg/µl in denaturation buffer. 30-100 µg of solublised proteins was mixed with 0.5X, in µg, of RNA, DNA, or heparin (all in TE buffer). All nucleic acids were heat-denatured at 96°C for 3 min and then rapidly cooled on ice before addition to the denatured proteins. The nucleic acid / protein mixture was transferred to dialysis tubes (see below) equipped with a 6-8.000 kDa cut-off membrane (Spectrum Lab). Dialysis was performed against 600 ml PBS buffer at 4°C overnight, after which the PBS was replaced with fresh PBS (400 ml) and the container placed in a water bath and kept at 37°C for 1 h. The dialysed samples were transferred to 1.5 ml tubes and the volume adjusted to 100-200 µl with PBS. 7.5-10% of this was taken as Input. Aggregated proteins (Pel 1) were pelleted by centrifugation at 21.000 x g for 10 min at +4°C, washed twice in RIPA buffer and process for SDS-PAGE as before. 7.5-10% of the supernatants were saved (Sup 1) and the remaining supernatant was either divided into two new tubes supplemented with 0.5 µl vehicle or 0.5 µl RNase A/T1 or the whole sample placed in one tube and treated with 0.5 µl RNase A/T1. All samples were incubated at 37°C for one hour and centrifuged as before. Pelleted proteins (Pel 2) were washed as before and dissolved in SDS/Urea and sonicated. Equal volumes of each

fraction were separated on SDS-PAGE gels and then either stained with coomassie or transferred to membranes for western blot analysis.

Dialysis tubes were prepared by drilling a 3 mm hole in the lid of a 1.5 ml microcentriruge tube (Crystal Clear, StarLab). The tube was then cut 1 cm from the top and a new intact lid inserted at the bottom. After sample addition, the tube was sealed with a dialysis membrane and capped with the drilled lid. This creates a dialysis tube where one end is in contact with the surrounding solution, separated by the membrane. Tubes were placed in the dialysis solution with the holed side facing down.

Nucleic-acid mediated renaturation through two-step dilution

Proteins aggregated by enyzmatic RNA degradation were isolated and denatured in denaturation buffer (20 mM Tris-HCl pH 7.5, 6 M Guanidine hydrochloride, 1 % Triton X-100, 20 mM DTT) as described above. The protein concentration was determined with the BCA kit (Thermo Fisher) and diluted to 5.55 µg/µl in denaturation buffer. Typically, 30 µg of denatured proteins, on ice, was mixed (by vortexing) with indicated amounts of nucleic acids, prepared in 1 X TE buffer, to achieve a 1:10 dilution, e.g. 5 µl of denatured proteins + 45 µl of diluted nucleic acids. All nucleic acids were heat-denatured (96^{o}C, 3 min) and cooled on ice before use, except when pre-annealed oligonucleotides were used. Pre-annealed oligonucleotides were prepared in TE buffer supplemented with 100 mM NaCl and placed in a PCR machine at 96^{o}C for 3 min followed by cooling to room temperature for approximately 45 minutes. All protein / nucleic acid mixtures were incubated for 5 min on ice and then diluted 10X with cold Renaturing buffer (10 mM Tris-HCl pH 7.4, 30 mM NaCl), e.g. 50 µl original mixture + 450 µl of Renaturing buffer. After incubation for 10 min on ice, samples were placed at 37^{o}C for 1 hour, shaking at 1200 rpm for 3 seconds every second minute. Aggregated and soluble proteins were separated by centrifugation, 21,000 x g for 30 min.

Aggregated proteins were solubilised in 20 mM Tris-HCl pH 7.5, 2% SDS, 8M Urea by sonication (Bioruptor, Diagenode) at maximum setting for 5 min at room temperature and either used for SDS-PAGE or the amount of protein determined using the BCA assay. When determining the amount of protein, the whole pellet was used in a typical 225 µl reaction, e.g. 25 µl sample + 200 µl BCA reagent. In the figures, the amount of aggregation is usually expressed a percentage of the amount of aggregation observed in the vehicle control, i.e. in samples treated with TE buffer only.

SDS-PAGE and western blot analysis

Heated samples were separated on 4-12% Bis-Tris gels (Life Technologies) in MOPS or MES buffer and either transferred to 0.2 µm nitrocellulose or 0.45 µm PVDF membranes (both GE Healthcare) for 2 hour at 45V on ice or, alternatively, used directly for coomassie staining (ProtoBlue, National Diagnostics) according to the manufacturer's protocol. In figures with Input and Supernatants, these represent 10%, typically 30 µg of protein, while for the Pellets the full amount was loaded. After transfer, membranes for Western blot were blocked for one hour at room temperature in 5% milk in TBS-T (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% Tween-20) and incubated with primary antibodies in the same solution or TBS-T/5% BSA overnight at +4^{o}C. Membranes were then washed 4 x 5 min in TBS-T and incubated for 1 hour at RT with HRP-conjugated secondary antibodies diluted in 5% milk/TBS-T. Membranes were then washed as before and incubated for 5 min in ECL Prime (GE Healthcare) before being exposed to films (Thermo Fisher). Primary antibodies used were: TDP-43 (New England Biolabs, NEB, #G400), HTT (NEB, #D7F7), FUS (SantaCruz, #sc-47711), MAPT (NEB, #Tau46), NF-H (mouse, Covance, #SMI-32R), NF-H (rabbit, Sigma, # N4142), Aβ 6E10 (Covance, #SIG-39320), ACTB (Sigma, #A2228), RPL7

(Abcam, #ab72550), PABP (Abcam, #ab21060), GFP (Abcam. #ab1218) HNRNPA1 (Proteintech, # 11176-1-AP), HNRNPD (Proteintech, #12770-1-AP). All primary antibodies were used at 1:1,000 dilution, except ACTB (1:4,000), NF-H (1:4,000), GFP (1:2,000), FUS (1:100), and RPL7 (1:2,000). As secondary antibodies, we used Donkey anti-Rabbit HRP (#NA934V) or Sheep anti-Mouse HRP (#NXA931), both from GE Healthcare, diluted 1:50,000 in 5% milk-TBS-T.

Isolation and renaturation of brain protein aggregates

Brain tissues (~200 mg frozen) were homogenised for approximately 30 seconds in 2 ml 0.8 M NaCl, 1% Triton X-100, 0.1 m EDTA, 0.01 M Tris-HCl pH 7.4, 1 mM DTT using a TissueRuptor (Qiagen). The homogenate was then centrifuged 5 min at 2,500 x g at 4°C and the supernatant transferred to a new tube on ice followed by further homogenisation using a 27G needle and syringe. The samples were then centrifuged 21,000 x g at 4°C for 30 min and the supernatant collected. SDS and Na-deoxycholate were added to the supernatant to 0.1% and 0.5% final concentrations, respectively and incubated for 10 minutes at RT. 500 µl of each sample was overlaid 800 µl of sucrose cushion (1M sucrose, 0.8 M NaCl, 1 % TX-100, 0.5 % Na- doxycholate, 0.2 % SDS, 50 mM Tris-HCl pH 7.8) and centrifuged at 167,000 x g for 2 hours at 4°C. The supernatant were decanted and discarded and the pellet re-suspended in PBS supplemented with NaCl to 500 mM and transferred to new tubes. Pellets were washed twice in PBS/0.5 M NaCl by centrifugation at 21,000 x g for 30 min and then dissolved in 6 M GuHCl, 1 % Triton X-100, 20 mM Tris-HCl by sonication. Brain and Jurkat samples were diluted to 2 µg / µl and 30 µg used for renaturation through dilution as described above, using total Jurkat RNA or the DNA versions of the M1x4 oligonucleotides (Table S1). Renatured samples were incubated for 1 hour at 37°C, shaking at 1,200 rpm every 5th second. Aggregated proteins (Pellet 1) were collected by centrifugation (21,000 x g,

1 hour, 4°C) and saved for analysis. The supernatants were supplemented with RNase A/T1 and incubated shaking 1,200 rpm at 37°C overnight. Aggregated proteins (Pellet 2) were collected by centrifugation as above and the supernatant supplemented with $MgCl_2$, 1 mM final concentration, and Benzonase (sc-391121B, Santa Cruz) and incubated at 37°C for 1 hour, shaking as above. Aggregated proteins (Pellet 3) were collected as above. All pellets (1-3) were dissolved in 2% SDS/8M Urea by sonication and analysed by western blot as described.

ATP-binding and hydrolysis

100 µg proteins aggregated by enzymatic degradation of RNA in human neuronal or Jurkat cell lysates of were renatured with 50 µg of total RNA or Ve (TE buffer) as described above. After dialysis, the samples were adjusted to 250 µl with PBS and centrifuged at 2,000 x g for 15 min, to remove protein aggregates that would later co-sediment with the agarose beads. Capturing of ATP binding proteins was performed on 75 µl of this mixture using 30 µl of Aminophenyl-ATP- or naked agarose beads (Jena Bioscience) according to the manufacturer's protocol. Elution was performed by two sequential 10 min incubations in 20 µl 1X LDS loading buffer (Life Technologies) supplemented with DTT (100 mM final concentration). One fourth of the eluted samples was separated on 4-12% NuPage gels (Life Technologies) and the gels stained with coomassie blue (ProtoBlue, National Diagnostics). The remaining eluate from two independent replicates was electrophoresed approximately 1 cm into a 4-12% NuPage and the top piece of the gel excised and prepared for mass spectrometry analysis as described below. To measure ATP hydrolysis, we used the ADP-Glo™ Kinase Assay (Promega) according to the manufacturer's instructions. Briefly, 5 µl of renatured proteins were mixed in a white 96-well plate (Santa Cruz Biotechnology) with ATP (100 uM final concentration) and 0.1 µl RNase A/T1 mixture or vehicle (50% Glycerol in 20

mM Tris-HCl pH 7.5), all diluted in 1X PBS, 5 mM $MgCl_2$, 2 mM DTT, in a total volume of 15 μl and incubated at room temperature for 1.5 hour. Non-hydrolysed ATP was removed by the addition of 15 μl of ADP-Glo reagent followed by incubation for 1 hour at room temperature. ADP was converted back to ATP by the addition of 30 μl Kinase Detection Reagent and the emitted light quantified after 1.5 hours incubation at room temperature using a Victor2 Multilabel plate-reader (Wallac). All samples were run in duplicate and data are presented as the mean of three independent replicates.

RNA and DNA isolation and analysis

RNA and DNA were isolated from cells or cell lysates with Trizol LS or Isol-RNA Lysis Reagent (Life Technologies and 5 PRIME, respectively), according to the manufacturer's instructions. All RNA and DNA samples was dissolved in either 0.1X (for RNA fragmentation, see below) or 1X TE buffer. RNA was analysed by 1.5% agarose gel electrophoresis and visualised with ethidium bromide.

Mass spectrometry analysis

30 μg of aggregated proteins in 1 x LDS loading buffer (Life Technologies) supplemented with 100 mM DTT were separated on 4-12% Bis-Tris gels in MOPS running buffer. After coomassie staining, each gel lane was divided into 10 equal gel-slices and cut into 1mm cubes. Gel bands were destained and reduced with 5 mM TCEP (Pierce) and alkylated with 50 mM chloracetamide (Sigma) and then digested with trypsin (Promega) for 16 hours. Samples were desalted using homemade C18 columns and then analysed using a QExactive mass spectrometer (Thermo) at the Central Proteomics Facility (University of Oxford, UK). Data were analysed using Mascot (MatrixScience) with searches performed against the

UniProt Human database. Proteins with a Mascot score greater than or equal to 60 and with two unique peptide sequences were considered to be confidently identified. Keratin hits derived from hair (cuticular) were removed from the list of ATP-precipitated proteins as they likely represent contamination introduced during handling.

Computational analysis of proteins aggregated by enzymatic RNA degradation

Proteins common to both MS samples were compiled into a list and used for further analysis. Gene ontology analysis was performed with FunRich (Pathan *et al*, 2015). Low complexity regions of 30 or more consecutive amino acids were identified with SEG (Wootton & Federhen, 1993) using the following parameters: [30 amino acid length][3.2 low complexity trigger][3.55 high extension complexity]. Unstructured regions were identified with DisEMBL (Linding *et al*, 2003) using the default settings with the following changes: amino acid window of 30, join 2, threshold 1.75. For statistical evaluation, the results were compared to those obtained by permutation analyses. A total of 1,000 permutations per analysis were performed. The permutations consisted of random sets of proteins (n=1,603), drawn from the complete set of human proteins (http://www.uniprot.org/downloads, accessed on 07/2013) and analysed using SEG and DisEMBL with the same settings as above. The cumulative distributions of the proportion of low complexity and unstructured regions were compared to the results obtained from the proteins aggregated by enzymatic degradation of RNA. Statistical analysis for each dataset was done using the two-sample Kalmogrov-Smirnov test (Conover, 1999) and corrected for multiple testing using Bonferroni correction.

Nucleic acid immunoprecipitation and sequencing

RNA and DNA fragmentation. Nucleic acids were isolated as described above. Total RNA (~400 ug), prepared in 0.1X TE buffer, at 0.5 µg/µl, were chemically fragmented in 50 mM

Tris-acetate pH 8.1, 100 mM CH_3CO_2K, 30 mM $Mg(CH3COO)_2$ by incubating the samples at 96°C for 12 min. The reaction was stopped by transferring the samples to ice and the addition of 0.5 M EDTA to final concentration of 45 mM. Fragmented RNA samples were then mixed with 1/10 volume of Na-acetate (3M, pH 5, Sigma) and 2.5 volumes of EtOH and incubated for 1 hour at -80°C. Precipitated RNA were collected by centrifugation: 30 min at 12,000 x g at +4°C, and washed once in 75% EtOH before being air-dried and dissolved in TE. Recovered RNA was spectrophotometrically quantified (NanoDrop, Thermo Fisher Scientific) and kept on ice until used. The average size of fragmented RNA was estimated from an aliquot analysed on a Bioanalyzer chip (RNA9000, Agilent Technologies), according to the manufacturer's instructions. Approximately 70% of the initial amount of RNA was recovered after fragmentation (data not shown).

Genomic DNA, in TE buffer at 3 µg/µl, was fragmented by several rounds of sonication (Bioruptor, Daigenode) at RT until an average fragment length of approximately 200 bp was achieved (data not shown).

Nucleic acid-mediated renaturation. Proteins aggregated by enzymatic degradation of RNA in human neuronal lysates were renatured with fragmented RNA or DNA using the dialysis procedure described above. To achieve efficient renaturation, the amount of fragmented RNA had to be increased from the normal 0.5 times to 1.5 times the amount of protein (weight/weight, data not shown). After dialysis, samples were transferred to 1.5 ml tubes and incubated at 37°C for one hour, shaking at 1200 rpm for 3 seconds every 1.5 min. Aggregated and soluble proteins were separated by centrifugation (21,000 x g, 30 min, +4°C). The soluble fraction was collected and divided into several tubes for immunoprecipitation or capturing onto nitrocellulose membranes, see below.

Native RNA-Immunoprecipitation. Human neuronal cell lysate were prepared as described above under "Preparation of cell free lysates from neurons and mouse cortex". 200 µg of lysate were incubated with 3 µg anti-NF-H (N4142, Sigma), 3 µg anti-Tau (T9450, Sigma), 3 µg anti-Aβ (4G8, SIG-39220, Covance), or 3 µg mouse IgG (Sigma) and incubated rotating for 2 hours at +4°C. Antibodies were captured by the addition of 15 µl protein A magnetic beads (10002D, Thermo Fisher Scientific) or 15 µl protein goat-anti mouse Ig magnetic (112.01D,Thermo Fisher Scientific) for Tau (both types of beads were pre-blocked in 5% BSA, 30 µg M1x4 R oligo in PBS/0.1 % Triton X-100 for 1 hour at RT), and incubated for 30 minutes at RT. Magnetic beads were collected using a magnet and washed two times in PBS, 0.5% Triton X-100, 2 times in PBS, 500 mM NaCl, 0.5% Triton X-100, and 1 time in PBS, 1000 mM NaCl, 0.5% Triton X-100 and once in TE. RNA were isolated with the GeneJET PCR Purification Kit according to the manufacturer's instructions with the following alterations: RNA was eluted from the magnetic beads by the addition of 50 µl 6 M guanidine hydrochloride. The supernatant were collected and mixed with 100 µl Binding Buffer, 100 µl water, and 300 µl 2-propanol and loaded onto the column. Co-purified DNA was removed by on-column digestion, using the RNase-Free DNase Set (Qiagen) reagents and protocol. Eluted RNA was then precipitated with Na-Acetate and EtOH, using glycogen (R0561, Thermo Fisher Scientific) as a carrier. DNase-treated samples were dissolved in 10 µl water and used for library preparation using the NEBNext Ultra II Directional RNA Library Prep Kit and Indexing primers (E7760, and E7335, respectively, both NEB) according to the manufacturer's instructions. Libraries were pooled and pair-end sequenced (2x75 bp) on the NextSeq 500 platform (Illumina). Two biological replicates of each immunoprecipitation were analysed, except for 4G8 where only one sample were analysed. Data can be accessed at GEO GSE99127.

RNA-Immunoprecipitation of renatured proteins. RNA-renatured proteins (~40 µg) were supplemented with 1 µg anti-NF-H (N4142, Sigma), 1 µg anti-Tau (T9450, Sigma), 1 µg anti-Aβ (4G8, SIG-39220, Covance), or 1 µg anti GFP (11122, Thermo Fisher Scientific) and incubated rotating for 2 hours at +4°C. Antibodies were captured and washed as described above. Samples were eluted with 20 µl 6 M guanidine thiocyanate at RT and RNA purified from the eluate with Isol (5-PRIME), according to the manufacturer's instructions using 1 µg glycogen (R0561, Thermo Fisher Scientific) as co-precipitant. RNA were dissolved in 30 µg 1x DNAs I buffer and treated with 1 µl DNAase I (2222, Thermo Fisher Scientific) for 30 min at 37°C. The sample-volumes were adjusted to 100 µl with TE and extracted with an equal volume of phenol (77617, Sigma) and precipitated with Na-Acetate and EtOH, using glycogen as a carrier. DNase-treated samples were dissolved in 10 µl water and used for sequencing library preparation using the NEBNext Directional Ultra RNA kit and Indexing primers (E7530, and E7335, respectively, both NEB) according to the manufacturer's instructions, except that no initial RNA fragmentation were performed. Libraries were pooled and pair-end sequenced (2x75 bp) on the MiSeq platform (Illumina). Two biological replicates of each immuprecipitation were analysed. Data can be accessed at GEO GSE99127.

DNA-immunopreciptation and membrane capture of renatured proteins. DNA-renatured samples were prepared as described above and aliquoted for immunoprecipitation or membrane capture. For membrane capture, approximately 40 µg renatured proteins or fragmented DNA alone (20 µg, negative control) were slowly passed through a 0.2 µm nitrocellulose membrane (1060004, GE-Healthcare) pre-blocked with 250 µg Yeast t-RNA (Thermo Fisher Scientific). The membrane were then washed three times with 10 mM Tris-HCl pH 7.5, 100 mM KCl, 0.1% Triton X-100, 2 times with 10 mM Tris-HCl pH 7.5, 500

mM KCl, 0.1% Triton X-100, and once with 1X TE. Captured proteins and nucleic acids were eluted by the addition of 2% SDS/8M Urea and the nucleic acids extracted with Isol (5-PRIME) as described above. Immunoprecipitation with Aβ 4G8 (1 μg) and mouse IgG (1 μg) were performed as described for RNA immunoprecipitation of renatured proteins above using approximately 40 μg of renatured proteins per IP. Protein A magnetic beads were pre-blocked for 1 hour at RT with 70 μg Yeast t-RNA (Thermo Fisher Scientific). Sequencing libraries were prepared using the NEBNext DNA ultra kit (E7370, NEB), using NEBNext Indexing primers (E7335, NEB), and the libraries pooled and sequenced on the MiSeq system (Illumina) using pair-end reads of 300 bp. Data can be accessed at GEO GSE99127.

Computational nucleic acid sequence analysis

Reads from both the DNA and RNA-seq experiments were trimmed using Trim Galore with default settings for paired-end reads. RNA reads were aligned to the human reference genome (hg19), using STAR (Dobin *et al*, 2013) using default settings. Potential PCR artefacts in the aligned RNA-seq files were removed by the MarkDuplicates tool in Picard-tools, using default settings. DNA sequences were aligned to the unmasked human reference genome (hg19) using Bowtie2. To identify enriched regions we employed MACS (Zhang *et al*, 2008), using the GFP (RNA), IgG (Native RNA-IP and DNA-IP), or DNA-only (membrane) samples as a negative control. We used the reads from both biological replicates as input files, except for the Native RNA-IP sample of 4G8, where only one sample passed the quality check. MACS were run with the following settings: band width = 150 for RNA or 188 for DNA, Broad region calling = off (RNA), Searching for subpeak summits = on (RNA), and call-summit = on for DNA. The *annotatePeaks* package in the HOMER software package (Heinz *et al*, 2010), run with default settings, was used to annotate the peaks files generated by MACS. For motif discovery we extracted a 300 nt fragment, centred at the peak

detected by MACS. We used the MEME suit of programs (MEME-Chip, (Machanick & Bailey, 2011)) with the following settings: *-meme-mod anr -meme-minw 5 -meme-maxw 50 -meme-nmotifs 6 -dreme-e 0.05 -centrimo-local -centrimo-score 5.0 -centrimo-ethresh 10.0*. Motif similarities were evaluated with tomtom (in the MEME suit) using default setting.

Quantification and Statistical Analysis

Statistical analyses in Figs. 3b, 3e, 5a, 5c-d, 6b-d, Expanded View Figs. 4c, and 7 was performed by one-way ANOVA analysis followed by post-hoc analysis with Bonferroni correction for multiple testing using the online tool at: http://astatsa.com/OneWay_Anova_with_TukeyHSD/. Statistical analysis for Expanded View Figs. 2d and e was performed using the two-sample Kalmogrov-Smirnov test (Conover, 1999) and corrected for multiple testing using Bonferroni correction. When mentioned, the term "independent samples" refers to biological repeats of the same experiments but using different starting material e.g. cell lysate.

Oligonucleotides

All oligonucleotides investigated in this study were purchased from IDT in a desalted form and used without any further purification. Sequences of oligonucleotides are listed in Expanded View Table 1.

Acknowledgements. We thank Dr B Thomas and the staff of the Central Proteomics Facility at the University of Oxford (www.proteomics.ox.ac.uk) for mass spectrometry and data analysis services, and Neil Atam for creating the movie. We also wish to thank Prof Michael Curtis, Prof Thomas MacDonald, Dr John Lanham, Dr Michele Hill-Perkins, Mrs. Julie Fennell and Mr. Graham Fennell for their support, and Dr Jarek Szary for discussions. We are grateful to Professor Linda Greensmith for critical review of the manuscript. This work was generously supported by a Strategic Project Grant from Barts and the London Charity, grants from Queen Mary Innovation, The Medical Research Council, The Motor Neurone Disease Association and Rosetrees Trust, a Wellcome Trust Value in People award, and a legacy from Mr. Ivan Fennell. In memory of Ivan Fennell.

Author contributions. JA conceived and planned the research, performed most of the experiments, analysed the data and wrote the paper. SR, TAJ and RA performed experiments. AM provided ALS samples, strategic planning, evaluated the data, and commented on the manuscript. CPC and MRB performed computational analysis of aggregated proteins and nucleic acids. GG supported the study and evaluated the data. DS conceived and planned the research, evaluated the data and wrote the paper together with JA.

Conflicts of interest

The authors declare that they have no conflicts of interest.

REFERENCES

Aguzzi A, Altmeyer M (2016) Phase Separation: Linking Cellular Compartmentalization to Disease. *Trends in Cell Biology* **26:** 547-558

Andersen JS, Lam YW, Leung AK, Ong SE, Lyon CE, Lamond AI, Mann M (2005) Nucleolar proteome dynamics. *Nature* **433:** 77-83

Bai B, Hales CM, Chen P-C, Gozal Y, Dammer EB, Fritz JJ, Wang X, Xia Q, Duong DM, Street C, Cantero G, Cheng D, Jones DR, Wu Z, Li Y, Diner I, Heilman CJ, Rees HD, Wu H, Lin L, Szulwach KE, Gearing M, Mufson EJ, Bennett DA, Montine TJ, Seyfried NT, Wingo TS, Sun YE, Jin P, Hanfelt J, Willcock DM, Levey A, Lah JJ, Peng J (2013) U1 small nuclear ribonucleoprotein complex and RNA splicing alterations in Alzheimer's disease. *Proc Natl Acad Sci U S A* **110:** 16562-16567

Bergeron-Sandoval L-P, Safaee N, Michnick Stephen W (2016) Mechanisms and Consequences of Macromolecular Phase Separation. *Cell* **165:** 1067-1079

Brangwynne CP (2011) Soft active aggregates: mechanics, dynamics and self-assembly of liquid-like intracellular protein bodies. *Soft Matter* **7:** 3052-3059

Buchberger A, Bukau B, Sommer T (2010) Protein Quality Control in the Cytosol and the Endoplasmic Reticulum: Brothers in Arms. *Molecular Cell* **40:** 238-252

Burke Kathleen A, Janke Abigail M, Rhine Christy L, Fawzi Nicolas L (2015) Residue-by-Residue View of In Vitro FUS Granules that Bind the C-Terminal Domain of RNA Polymerase II. *Molecular Cell* **60:** 231-241

Capitini C, Conti S, Perni M, Guidi F, Cascella R, De Poli A, Penco A, Relini A, Cecchi C, Chiti F (2014) TDP-43 inclusion bodies formed in bacteria are structurally amorphous, non-amyloid and inherently toxic to neuroblastoma cells. *PLoS One* **9:** e86720

Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann BM, Strein C, Davey NE, Humphreys DT, Preiss T, Steinmetz LM, Krijgsveld J, Hentze MW (2012) Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. *Cell* **149:** 1393-1406

Caughey B, Raymond GJ (1993) Sulfated polyanion inhibition of scrapie-associated PrP accumulation in cultured cells. *Journal of Virology* **67:** 643-650

Chattopadhyay S, Das B, Dasgupta C (1996) Reactivation of denatured proteins by 23S ribosomal RNA: role of domain V. *Proc Natl Acad Sci U S A* **93:** 8284-8287

Conlon EG, Manley JL (2017) RNA-binding proteins in neurodegeneration: mechanisms in aggregate. *Genes & Development* **31:** 1509-1528

Conover WJ (1999) *Practical Nonparametric Statistics*, 3rd edn. New York: Wiley.

David DC, Ollikainen N, Trinidad JC, Cary MP, Burlingame AL, Kenyon C (2010) Widespread Protein Aggregation as an Inherent Part of Aging in *C. elegans*. *PLoS Biol* **8:** e1000450

Decker CJ, Parker R (2012) P-Bodies and Stress Granules: Possible Roles in the Control of Translation and mRNA Degradation. *Cold Spring Harbor Perspectives in Biology* **4:** a012286

Deleault NR, Lucassen RW, Supattapone S (2003) RNA molecules stimulate prion protein conversion. *Nature* **425:** 717-720

DiFiglia M, Sapp E, Chase KO, Davies SW, Bates GP, Vonsattel JP, Aronin N (1997) Aggregation of Huntingtin in Neuronal Intranuclear Inclusions and Dystrophic Neurites in Brain. *Science* **277:** 1990-1993

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)* **29:** 15-21

Docter BE, Horowitz S, Gray MJ, Jakob U, Bardwell JCA (2016) Do nucleic acids moonlight as molecular chaperones? *Nucleic Acids Research* **44:** 4835-4845

Fratta P, Mizielinska S, Nicoll AJ, Zloh M, Fisher EMC, Parkinson G, Isaacs AM (2012) C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. *Scientific Reports* **2:** 1016

Furukawa Y, Kaneko K, Watanabe S, Yamanaka K, Nukina N (2011) A Seeding Reaction Recapitulates Intracellular Formation of Sarkosyl-insoluble Transactivation Response Element (TAR) DNA-binding Protein-43 Inclusions. *Journal of Biological Chemistry* **286:** 18664-18672

Haeusler AR, Donnelly CJ, Periz G, Simko EAJ, Shaw PG, Kim M-S, Maragakis NJ, Troncoso JC, Pandey A, Sattler R, Rothstein JD, Wang J (2014) C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* **507:** 195-200

Haltia M, Prelli F, Ghiso J, Kiuru S, Somer H, Palo J, 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

Hamilton RL (2000) Lewy Bodies in Alzheimer's Disease: A Neuropathological Review of 145 Cases Using α-Synuclein Immunohistochemistry. *Brain Pathology* **10:** 378-384

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38:** 576-589

Higashi S, Iseki E, Yamamoto R, Minegishi M, Hino H, Fujisawa K, Togo T, Katsuse O, Uchikado H, Furukawa Y, Kosaka K, Arai H (2007) Concurrence of TDP-43, tau and α-synuclein pathology in brains of Alzheimer's disease and dementia with Lewy bodies. *Brain Research* **1184:** 284-294

Hirabayashi M, Inoue K, Tanaka K, Nakadate K, Ohsawa Y, Kamei Y, Popiel AH, Sinohara A, Iwamatsu A, Kimura Y, Uchiyama Y, Hori S, Kakizuka A (2001) VCP/p97 in abnormal protein aggregates, cytoplasmic vacuoles, and cell death, phenotypes relevant to neurodegeneration. *Cell Death Differ* **8:** 977-984

Hirano A (1994) Hirano bodies and related neuronal inclusions. *Neuropathology and Applied Neurobiology* **20:** 3-11

Hofweber M, Hutten S, Bourgeois B, Spreitzer E, Niedner-Boblenz A, Schifferer M, Ruepp M-D, Simons M, Niessing D, Madl T, Dormann D (2018) Phase Separation of FUS Is Suppressed by Its Nuclear Import Receptor and Arginine Methylation. *Cell* **173:** 706-719.e713

Jain A, Vale RD (2017) RNA phase transitions in repeat expansion disorders. *Nature* **546:** 243

Jones LS, Yazzie B, Middaugh CR (2004) Polyanions and the Proteome. *Molecular & Cellular Proteomics* **3:** 746-769

Jones TA, Ogunkolade BW, Szary J, Aarum J, Mumin MA, Patel S, Pieri CA, Sheer D (2011) Widespread expression of BORIS/CTCFL in normal and cancer cells. *PLoS One* **6:** e22399

Kaganovich D, Kopito R, Frydman J (2008) Misfolded proteins partition between two distinct quality control compartments. *Nature* **454:** 1088-1095

Kampers T, Friedhoff P, Biernat J, Mandelkow EM, Mandelkow E (1996) RNA stimulates aggregation of microtubule-associated protein tau into Alzheimer-like paired helical filaments. *FEBS Letters* **399:** 344-349

Kato M, Han Tina W, Xie S, Shi K, Du X, Wu Leeju C, Mirzaei H, Goldsmith Elizabeth J, Longgood J, Pei J, Grishin Nick V, Frantz Douglas E, Schneider Jay W, Chen S, Li L, Sawaya Michael R, Eisenberg D, Tycko R, McKnight Steven L (2012) Cell-free Formation of RNA Granules: Low Complexity Sequence Domains Form Dynamic Fibers within Hydrogels. *Cell* **149:** 753-767

Kedersha N, Stoecklin G, Ayodele M, Yacono P, Lykke-Andersen J, Fritzler MJ, Scheuner D, Kaufman RJ, Golan DE, Anderson P (2005) Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *The Journal of Cell Biology* **169:** 871-884

Kim HJ, Kim NC, Wang YD, Scarborough EA, Moore J, Diaz Z, MacLea KS, Freibaum B, Li S, Molliex A, Kanagaraj AP, Carter R, Boylan KB, Wojtas AM, Rademakers R, Pinkus JL, Greenberg SA, Trojanowski JQ, Traynor BJ, Smith BN, Topp S, Gkazi AS, Miller J, Shaw CE, Kottlors M, Kirschner J, Pestronk A, Li YR, Ford AF, Gitler AD, Benatar M, King OD, Kimonis VE, Ross ED, Weihl CC, Shorter J, Taylor JP (2013) Mutations in prion-like

domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* **495:** 467-473

Kimonis VE, Fulchiero E, Vesa J, Watts G (2008) VCP disease associated with myopathy, Paget disease of bone and frontotemporal dementia: Review of a unique disorder. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **1782:** 744-748

Kovachev PS, Banerjee D, Rangel LP, Eriksson J, Pedrote MM, Martins-Dinis MMDC, Edwards K, Cordeiro Y, Silva JL, Sanyal S (2017) Distinct modulatory role of RNA in the aggregation of the tumor suppressor protein p53 core domain. *Journal of Biological Chemistry* **292:** 9345-9357

Kovachev PS, Gomes MPB, Cordeiro Y, Ferreira NC, Valadão LPF, Ascari LM, Rangel LP, Silva JL, Sanyal S (2019) RNA modulates aggregation of the recombinant mammalian prion protein by direct interaction. *Scientific Reports* **9:** 12406

Kwiatkowski TJ, Bosco DA, LeClerc AL, Tamrazian E, Vanderburg CR, Russ C, Davis A, Gilchrist J, Kasarskis EJ, Munsat T, Valdmanis P, Rouleau GA, Hosler BA, Cortelli P, de Jong PJ, Yoshinaga Y, Haines JL, Pericak-Vance MA, Yan J, Ticozzi N, Siddique T, McKenna-Yasek D, Sapp PC, Horvitz HR, Landers JE, Brown RH (2009) Mutations in the FUS/TLS Gene on Chromosome 16 Cause Familial Amyotrophic Lateral Sclerosis. *Science* **323:** 1205-1208

Lin Y, Protter DS, Rosen MK, Parker R (2015) Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Molecular Cell* **60:** 208-219

Linding R, Jensen LJ, Diella F, Bork P, Gibson TJ, Russell RB (2003) Protein disorder prediction: implications for structural proteomics. *Structure* **11:** 1453-1459

Liu EY, Cali CP, Lee EB (2017) RNA metabolism in neurodegenerative disease. *Dis Model Mech* **10:** 509-518

Machanick P, Bailey TL (2011) MEME-ChIP: motif analysis of large DNA datasets. *Bioinformatics (Oxford, England)* **27:** 1696-1697

Maharana S, Wang J, Papadopoulos DK, Richter D, Pozniakovsky A, Poser I, Bickle M, Rizk S, Guillén-Boixet J, Franzmann TM, Jahnel M, Marrone L, Chang Y-T, Sterneckert J, Tomancak P, Hyman AA, Alberti S (2018) RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science* **360:** 918-921

Mandal PK, Pettegrew JW, Masliah E, Hamilton RL, Mandal R (2006) Interaction between Aβ Peptide and α Synuclein: Molecular Mechanisms in Overlapping Pathology of Alzheimer's and Parkinson's in Dementia with Lewy Body Disease. *Neurochemical Research* **31:** 1153-1162

Mann JR, Gleixner AM, Mauna JC, Gomes E, DeChellis-Marks MR, Needham PG, Copley KE, Hurtle B, Portz B, Pyles NJ, Guo L, Calder CB, Wills ZP, Pandey UB, Kofler JK, Brodsky JL, Thathiah A, Shorter J, Donnelly CJ (2019) RNA Binding Antagonizes Neurotoxic Phase Transitions of TDP-43. *Neuron* **102:** 321-338.e328

Markiv A, Beatson R, Burchell J, Durvasula RV, Kang AS (2011) Expression of recombinant multi-coloured fluorescent antibodies in gor -/trxB- E. colicytoplasm. *BMC Biotechnology* **11:** 117

Mathura VS, Paris D, Ait-Ghezala G, Quadros A, Patel NS, Kolippakkam DN, Volmar CH, Mullan MJ (2005) Model of Alzheimer's disease amyloid-beta peptide based on a RNA binding protein. *Biochem Biophys Res Commun* **332:** 585-592

McLean CA, Cherny RA, Fraser FW, Fuller SJ, Smith MJ, Beyreuther K, Bush AI, Masters CL (1999) Soluble pool of Abeta amyloid as a determinant of severity of neurodegeneration in Alzheimer's disease. *Ann Neurol* **46:** 860-866

Molliex A, Temirov J, Lee J, Coughlin M, Kanagaraj Anderson P, Kim Hong J, Mittag T, Taylor JP (2015) Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. *Cell* **163:** 123-133

Nakashima-Yasuda H, Uryu K, Robinson J, Xie SX, Hurtig H, Duda JE, Arnold SE, Siderowf A, Grossman M, Leverenz JB, Woltjer R, Lopez OL, Hamilton R, Tsuang DW, Galasko D, Masliah E, Kaye J, Clark CM, Montine TJ, Lee VM, Trojanowski JQ (2007) Co-morbidity of TDP-43 proteinopathy in Lewy body related diseases. *Acta Neuropathol* **114:** 221-229

Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT, Bruce J, Schuck T, Grossman M, Clark CM, McCluskey LF, Miller BL, Masliah E, Mackenzie IR, Feldman H, Feiden W, Kretzschmar HA, Trojanowski JQ, Lee VM (2006) Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314:** 130-133

Nikolakaki E, Drosou V, Sanidas I, Peidis P, Papamarcaki T, Iakoucheva LM, Giannakouros T (2008) RNA association or phosphorylation of the RS domain prevents aggregation of RS domain-containing proteins. *Biochimica et biophysica acta* **1780:** 214-225

Olzscha H, Schermann SM, Woerner AC, Pinkert S, Hecht MH, Tartaglia GG, Vendruscolo M, Hayer-Hartl M, Hartl FU, Vabulas RM (2011) Amyloid-like Aggregates Sequester Numerous Metastable Proteins with Essential Cellular Functions. *Cell* **144:** 67-78

Pathan M, Keerthikumar S, Ang CS, Gangoda L, Quek CY, Williamson NA, Mouradov D, Sieber OM, Simpson RJ, Salim A, Bacic A, Hill AF, Stroud DA, Ryan MT, Agbinya JI, Mariadason JM, Burgess AW, Mathivanan S (2015) FunRich: An open access standalone functional enrichment and interaction network analysis tool. *Proteomics* **15:** 2597-2601

Pepys MB, Hawkins PN, Booth DR, Vigushin DM, Tennent GA, Soutar AK, Totty N, Nguyen O, Blake CC, Terry CJ, et al. (1993) Human lysozyme gene mutations cause hereditary systemic amyloidosis. *Nature* **362:** 553-557

Qamar S, Wang G, Randle SJ, Ruggeri FS, Varela JA, Lin JQ, Phillips EC, Miyashita A, Williams D, Ströhl F, Meadows W, Ferry R, Dardov VJ, Tartaglia GG, Farrer LA, Kaminski Schierle GS, Kaminski CF, Holt CE, Fraser PE, Schmitt-Ulms G, Klenerman D, Knowles T, Vendruscolo M, St George-Hyslop P (2018) FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation-π Interactions. *Cell* **173:** 720-734.e715

Rajendran L, Honsho M, Zahn TR, Keller P, Geiger KD, Verkade P, Simons K (2006) Alzheimer's disease beta-amyloid peptides are released in association with exosomes. *Proc Natl Acad Sci U S A* **103:** 11172-11177

Ramaswami M, Taylor JP, Parker R (2013) Altered Ribostasis: RNA-Protein Granules in Degenerative Disorders. *Cell* **154:** 727-736

Rubinsztein DC (2006) The roles of intracellular protein-degradation pathways in neurodegeneration. *Nature* **443:** 780-786

Sanchez de Groot N, Armaos A, Graña-Montes R, Alriquet M, Calloni G, Vabulas RM, Tartaglia GG (2019) RNA structure drives interaction with proteins. *Nature Communications* **10:** 3246

Shelkovnikova TA, Robinson HK, Southcombe JA, Ninkina N, Buchman VL (2014) Multistep process of FUS aggregation in the cell cytoplasm involves RNA-dependent and RNA-independent mechanisms. *Hum Mol Genet* **23:** 5211-5226

Spector DL, Lamond AI (2011) Nuclear Speckles. *Cold Spring Harbor Perspectives in Biology* **3:** a000646

Sreedharan J, Blair IP, Tripathi VB, Hu X, Vance C, Rogelj B, Ackerley S, Durnall JC, Williams KL, Buratti E, Baralle F, de Belleroche J, Mitchell JD, Leigh PN, Al-Chalabi A, Miller CC, Nicholson G, Shaw CE (2008) TDP-43 Mutations in Familial and Sporadic Amyotrophic Lateral Sclerosis. *Science* **319:** 1668-1672

Swanson MS, Dreyfuss G (1988) RNA binding specificity of hnRNP proteins: a subset bind to the 3' end of introns. *EMBO J* **7:** 3519-3529

Takahashi T, Tada K, Mihara H (2009) RNA aptamers selected against amyloid β-peptide (Aβ) inhibit the aggregation of Aβ. *Molecular BioSystems* **5:** 986-991

Vance C, Rogelj B, Hortobagyi T, De Vos KJ, Nishimura AL, Sreedharan J, Hu X, Smith B, Ruddy D, Wright P, Ganesalingam J, Williams KL, Tripathi V, Al-Saraj S, Al-Chalabi A, Leigh PN, Blair IP, Nicholson G, de Belleroche J, Gallo JM, Miller CC, Shaw CE (2009) Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* **323:** 1208-1211

Vieira TCRG, Cordeiro Y, Caughey B, Silva JL (2014) Heparin binding confers prion stability and impairs its aggregation. *The FASEB Journal* **28:** 2667-2676

Walker JE, Saraste M, Runswick MJ, Gay NJ (1982) Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J* **1:** 945-951

Wallace EWJ, Kear-Scott JL, Pilipenko EV, Schwartz MH, Laskowski PR, Rojek AE, Katanski CD, Riback JA, Dion MF, Franks AM, Airoldi EM, Pan T, Budnik BA, Drummond DA (2015) Reversible, Specific, Active Aggregates of Endogenous Proteins Assemble upon Heat Stress. *Cell* **162:** 1286-1298

Walsh DM, Lomakin A, Benedek GB, Condron MM, Teplow DB (1997) Amyloid beta-protein fibrillogenesis. Detection of a protofibrillar intermediate. *J Biol Chem* **272:** 22364-22372

Walther DM, Kasturi P, Zheng M, Pinkert S, Vecchi G, Ciryam P, Morimoto RI, Dobson CM, Vendruscolo M, Mann M, Hartl FU (2015) Widespread Proteome Remodeling and Aggregation in Aging *C. elegans*. *Cell* **161:** 919-932

Wang A, Conicella AE, Schmidt HB, Martin EW, Rhoads SN, Reeb AN, Nourse A, Ramirez Montero D, Ryan VH, Rohatgi R, Shewmaker F, Naik MT, Mittag T, Ayala YM, Fawzi NL (2018) A single N-terminal phosphomimic disrupts TDP-43 polymerization, phase separation, and RNA splicing. *EMBO J* **37:** e97452

Watts GD, Wymer J, Kovach MJ, Mehta SG, Mumm S, Darvish D, Pestronk A, Whyte MP, Kimonis VE (2004) Inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia is caused by mutant valosin-containing protein. *Nat Genet* **36:** 377-381

Wong Cn, Xiong L-W, Horiuchi M, Raymond L, Wehrly K, Chesebro B, Caughey B (2001) Sulfated glycans and elevated temperature stimulate PrP(Sc)-dependent cell-free formation of protease-resistant prion protein. *EMBO J* **20:** 377-386

Wootton JC, Federhen S (1993) Statistics of local complexity in amino acid sequences and sequence databases. *Computers & Chemistry* **17:** 149-163

Wu H, Fuxreiter M (2016) The Structure and Dynamics of Higher-Order Assemblies: Amyloids, Signalosomes, and Granules. *Cell* **165:** 1055-1066

Xia Q, Liao L, Cheng D, Duong DM, Gearing M, Lah JJ, Levey AI, Peng J (2008) Proteomic identification of novel proteins associated with Lewy bodies. *Frontiers in bioscience : a journal and virtual library* **13:** 3850-3856

Zhang H, Elbaum-Garfinkle S, Langdon EM, Taylor N, Occhipinti P, Bridges AA, Brangwynne CP, Gladfelter AS (2015) RNA Controls PolyQ Protein Phase Transitions. *Mol Cell* **60:** 220-230

Zhang X, Lin Y, Eschmann NA, Zhou H, Rauch JN, Hernandez I, Guzman E, Kosik KS, Han S (2017) RNA stores tau reversibly in complex coacervates. *PLOS Biology* **15:** e2002183

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS (2008) Model-based analysis of ChIP-Seq (MACS). *Genome biology* **9:** R137

FIGURE LEGENDS

Figure 1. Enzymatic degradation of RNA causes protein aggregation. a Diagram showing the experimental design. **b** SDS-PAGE analysis of soluble (Supernatant) and insoluble proteins (Pellet) from human neurons after treatment with a mixture of RNase A and RNase T1 (A/T1), or vehicle (Ve-). **c** Protein aggregation (Pellet) after incubation with different ribonucleases or DNase I in the presence of either EDTA or Mg^{2+}. Ribonucleases used were RNase A (A), RNase T1 (T1), a mixture of RNase A and RNase T1 (A/T1), RNase 1f (1f), and RNase V1 (V1),

Figure 2. Enzymatic degradation of RNA induces aggregation of proteins associated with neurodegenerative diseases. a-b Western blot analysis of soluble and aggregated proteins after RNase treatment of lysate from human neurons, **a**, or HEK293 cells expressing GFP-Aβ or GFP, **b**. **c** Effect of an RNase A inhibitor (RNasin) on RNase A mediated protein aggregation in neuronal cell lysate.

Figure 3. RNA is required for renaturation and function of proteins aggregated by enzymatic degradation of RNA in human neuronal lysates. a Diagram showing the renaturing assay. **b** Effect of RNA/protein ratio on protein aggregation after renaturing. **c** Coomassie-stained gel of soluble (Sup 1/2) and aggregated (Pel 1/2) proteins after removal of GuHCl in the presence (+) or absence (-) of total RNA. Asterix (*) in **c** denotes added RNase A. **d** As in **c** but analysed by Western blot. **e** ATP-hydrolysing activity of proteins renatured with total RNA or buffer (Ve-) after removal of GuHCl. All experiments were performed with total RNA. Data in **b** and **e** are expressed in arbitrary units (AU) and represent the mean ± s.d of two (**b**) or three (**e**) independent experiments. **=p<0.01, *=p<0.05 by ANOVA post-hoc analysis.

Figure 4. Features of solubilising nucleic acids. a Genomic attributes of RNA associated with renatured NF-H, Tau, or Aβ. Inset shows which fractions of the peaks that derive from exons or introns have their origin in coding or non-coding transcripts, respectively. **b** Characterization of DNA associated with soluble, renatured Aβ or all proteins (Total), captured either by immunoprecipitation (Aβ) or absorption to membranes (Total). **c** Sequence logo of the computationally identified motif (M1) in the solubilising genomic DNA of the Total samples.

Figure 5. Renaturing characteristics of selected motif. a Proportion of aggregated proteins after renaturation with either single-stranded (Forward or Reverse) or complementary strands (For/Rev) of the selected M1x4 or the control oligonucleotide, M2x4. **b** Western blot analysis of proteins renatured with either single-stranded or double-stranded M1x4 or M2x4 DNA oligonucleotides. **c** The effect on protein re-aggregation from varying the number of DNA motif repeats when complementary oligonucleotides are given together. **d** Protein aggregation after renaturation with the M1x4 DNA oligonucleotides pre-annealed with 100 mM NaCl or vehicle (None). Data are expressed as a fraction of vehicle (Ve-) and represent the mean ± s.d of three to four independent experiments. **=p<0.01, *=p<0.05 by ANOVA post-hoc analysis.

Figure 6. Structural characteristics of protein-solubilising oligonucleotides and renaturation of ALS brain-derived protein aggregates.

a Cartoons of ds/ss/ds oligonucleotides having either pyrimidines (T or C) or purines (A) in the ss-region. **b** Proportion of protein aggregation following renaturing with the ds/ss/ds oligonucleotides shown in (a). **c** Renaturing capacity of structurally different oligonucleotides. The diagrams on the right show a theoretical structure of each oligonucleotide. All oligonucleotides, except the 3x-loop and 3x-bulges, contain a stretch of

30 Ts in the single-stranded region and the same sequences (15 nucleotides each) in the double stranded regions. The 3x-loops and bulges oligonucleotides have 3 stretches of 9 Ts and the same sequence in the ds-regions. **d** Proportion of protein aggregation in Jurkat cell lysate supplemented with various amounts and configurations of the M1x4 or M2x4 DNA oligonucleotides. **e** Insoluble proteins from two ALS brain tissues were chemically denatured in guanidine hydrochloride and treated with either buffer (Vehicle), total RNA from Jurkat cells or the complementary strands of the M1x4 DNA oligonucleotides (M1 F/R) (Pellet 1, top panel). After removal of GuHCl, the soluble fraction from these samples was treated with RNase and any aggregated proteins (Pellet 2), analysed by western blot (middle panel). Remaining supernatants were then treated with Benzonase to degrade any remaining nucleic acids and aggregated proteins collected by centrifugation and analysed by western blot (bottom panel). Proteins aggregated by enzymatic RNA degradation in Jurkat cell lysates, which do not contain NF-H, were used as a control. Data in **b** and **c** are expressed as a fraction of vehicle (Ve-), while data in **d** are expressed as the amount of aggregation observed without any oligonucleotides present (A/T1). All bars represent the mean ± s.d of two to four independent experiments. **=$p<0.01$, *=$p<0.05$ by ANOVA post-hoc analysis.

EXPANDED VIEW FIGURE LEGENDS

Expanded View Figure 1. Protein aggregation caused by enzymatic degradation of RNA, and effect of pre-hydrolysed RNA or RNase inhibition. a SDS-PAGE analysis of soluble (Supernatant) and aggregated (Pellet) proteins after treatment of mouse brain tissue lysate with increasing amounts of a mixture of RNase A and T1 (A/T1) or vehicle (Ve-). **b** Agarose gel electrophoresis analysis of RNA isolated from RNase-treated human neuronal cell lysate. **c-d** Insoluble proteins collected by centrifugation after co-treatment of human neuronal cell lysate with RNase A and an RNase A inhibitor (RNasin, **c**), or following the addition of RNA pre-hydrolysed by RNase A or alkaline hydrolysis (NaOH, **d**).

Expanded View Figure 2. Computational analysis of proteins aggregated by enzymatic degradation of RNA. (**a-c**) Top ten gene ontology classes by Cellular component **a**, Molecular function **b**, or Biological process **c**. **d-e** Cumulative distribution of the proportion of predicted low-complexity regions **d** or unstructured regions **e** in the RNase-aggregated proteins (Red) or random sets of proteins (Blue). *p*-values in **d-e** were obtained by the two-sample Kalmogrov-Smirnov test and corrected for multiple testing using Bonferroni correction.

Expanded View Figure 3. Protein aggregation caused by enzymatic degradation of RNA in mouse brain tissue lysate. Western blot analysis of aggregated proteins, collected by centrifugation, after treatment of mouse brain lysate with a mixture of RNase A and RNase T1 (RNase A/T1), or vechicle (Ve-). Input represents the starting material, Sup the supernatant (soluble fraction), and Pellet the aggregated fraction after centrifugation.

Expanded View Figure 4. Protein renaturation: effects on enzymatic activity and effects of various polyanions. a Coomassie-stained SDS-PAGE gels of RNase-aggregated proteins from human neurons renatured with various polyanions. Input represents 1/10th of each

sample, taken before the first centrifugation, and Pel 1 and Sup 1 are the pellet and the supernatant, respectively, recovered after the first centrifugation. Pel 2 represents the pellet obtained after treating the Sup 1 fraction with RNase A/T1 followed by centrifugation, please see experimental outline in Fig. 3a for a full description. **b** Renaturing of recombinant TDP-43 using increasing amounts of total RNA. Pel 1 represents aggregated protein after renaturing. Pel 2 is the aggregated protein after the soluble fraction (Sup 1) has been treated with RNase A/T. **c** Gel-electrophoresis analysis of proteins from human neurons renatured with vehicle or total RNA and then captured with ATP-agarose beads (ATP). Input represents 1/10th of each sample taken directly after renaturation, before the first centrifugation, and Blank represents the sample without any ATP-agarose beads present. **d** ATP-hydrolysing activity of RNase-aggregated proteins from Jurkat T-cells renatured in the presence (RNA) or absence (Ve-) of RNA. Data in **c** are expressed in arbitrary units (AU) and represent the mean ± s.d of three independent experiments. **=p<0.01, *=p<0.05 by ANOVA post-hoc analysis using Bonferroni correction for multiple testing.

Expanded View Figure 5. Comparison of identified motifs. a Sequence of the selected M1 motif and a random control oligo (Motif 2) investigated for their renaturing capacity. Both motifs were assembled into four consecutive repeats of the same motif (M1x4). The underlined region of M1 represents the computationally identified motif while the two extra 5' nucleotides (GC) where added for cloning purposes. **b** Alignment of motifs identified in the nucleic acids associated with soluble proteins after renaturation. All motifs are compared to the M1 motif identified in the Total DNA samples. The quoted original E-value represents the value obtained in the MEME-Chip analysis and the Similarity p-value is the p-value given by the TomTom motif comparison software package, compared to the Total DNA motif.

Expanded View Figure 6. Renaturation effects of selected oligonucleotides. Coomassie-stained SDS-PAGE gel of aggregated proteins (Pellet) following renaturation with either single-stranded (For or Rev) or complementary (For/Rev) M1x4 or M2x4 DNA oligonucleotides. gDNA represents genomic DNA, and Ve- represents vehicle.

Expanded View Figure 7. Protein aggregation in cell lysate supplemented with structural oligonucleotides and treated with RNase. Amount of protein aggregation observed after treating cell lysate from Jurkat T-cells with various oligonucleotides and RNase A/T1 (A/T1) and, if indicated, together with RNase H (A/T1 + H). Data are expressed as a percentage of the aggregation observed in samples treated with RANse A/T1 only and represent the mean ± s.d of three independent experiments. **=$p<0.01$ by ANOVA post-hoc analysis using Bonferroni correction for multiple testing.

EXPANDED VIEW MOVIE LEGEND

Expanded View Video 1. Effect of RNase-treatment on a clear cell lysate. Time-lapse movie of Jurkat T-cell lysate treated with RNase A/T1 or vehicle for 30 minutes at room temperature.

Expanded View Table 1.

Oligonucleotides used in this study.

Name	Sequence	Used in Fig:
Abeta_XhoI_F	CTA GCT CGA GGC CAC CGA TGC AGA ATT CCG ACA TGA	2b
Abeta_BamHI_R	CTA GGG ATC CGG ACA ACA CCG CCC ACC ATG A	2b
TARDBP BspHI_F	CTA GTC ATG ATG TCT GAA TAT ATT CGG GTA AC	4b
TARDBP Not I_R	CTA GGC GGC CGC CAT TCC CCA GCC AGA AG	4b
M1x4 For	**GAT CC**G CTA ATT TTT GTA TTT TTA GTA GCT AAT TTT TGT ATT TTT AGT AGC TAA TTT TTG TAT TTT TAG TAG CTA ATT TTT GTA TTT TTA GTA **C**	5a-d, 6d-e & S6
M1x3 For	**GAT CC**G CTA ATT TTT GTA TTT TTA GTA GCT AAT TTT TGT ATT TTT AGT AGC TAA TTT TTG TAT TTT TAG TA**C**	5c
M1x2 For	**GAT CC**G CTA ATT TTT GTA TTT TTA GTA GCT AAT TTT TGT ATT TTT AGT A**C**	5c
M1x1 For	**GAT CC**G CTA ATT TTT GTA TTT TTA GTA **C**	5c
M1x4 Rev	**TCG AG**T ACT AAA AAT ACA AAA ATT AGC TAC TAA AAA TAC AAA AAT TAG CTA CTA AAA ATA CAA AAA TTA GCT ACT AAA AAT ACA AAA ATT AGC **G**	5a-d, 6d-e & S6
M1x3 Rev	**TCG AG**T ACT AAA AAT ACA AAA ATT AGC TAC TAA AAA TAC AAA AAT TAG CTA CTA AAA ATA CAA AAA TTA GC**G**	5c
M1x2 Rev	**TCG AG**T ACT AAA AAT ACA AAA ATT AGC TAC TAA AAA TAC AAA AAT TAG C**G**	5c
M1x1 Rev	**TCG AG**T ACT AAA AAT ACA AAA ATT AGC **G**	5c
M2x4 For	**GAT CC**G AGT AAG AAT CTA TTA TAT ATG GAG TAA GAA TCT ATT ATA TAT GGA GTA AGA ATC TAT TAT ATA TGG AGT AAG AAT CTA TTA TAT ATG **C**	
M2x4 Rev	**TCG AG**C ATA TAT AAT AGA TTC TTA CTC CAT ATA TAA TAG ATT CTT ACT CCA TAT ATA ATA GAT TCT TAC TCC ATA TAT AAT AGA TTC TTA CTC **G**	5a-b, 6d, S6
ds/T28/ds For	CTT GCG GCC GCG CGC TTT TTT TTT TTT TTT TTT TTT TTT TGG CCA AGC TTA CCC GCC	6a, b

ds/T28/ds Rev	GGC GGG TAA GCT TGG CCT TTT TTT TTT TTT TTT TTT TTT TTT GCG CGC GGC CGC AAG	6a, b
ds/C28/ds Rev	CTT GCG GCC GCG CGC CCC CCC CCC CCC CCC CCC CCC CCC CGG CCA AGC TTA CCC GCC	6a, b
ds/C28/ds For	GGC GGG TAA GCT TGG CCC CCC CCC CCC CCC CCC CCC CCC CCC GCG CGC GGC CGC AAG	6a, b
ds/A28/ds Rev	CTT GCG GCC GCG CGC AAA AAA AAA AAA AAA AAA AAA AAA AAA AGG CCA AGC TTA CCC GCC	6a, b
ds/A28/ds For	GGC GGG TAA GCT TGG CCA AAA AAA AAA AAA AAA AAA AAA AAA AAA GCG CGC GGC CGC AAG	6a, b
ss T50	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	6c
ss A50	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	6c
15ds/T30/15ds For	GGG CCC GGG CCC GGG TTT TTT TTT TTT TTT TTT TTT TTT TTT GCG CGC GCG CGC GCG	6c, S6b
15ds/T30/15ds Rev	CGC GCG CGC GCG CGC TTT TTT TTT TTT TTT TTT TTT TTT TTT CCC GGG CCC GGG CCC	6c, S6b
15ds/A30/15ds For	GGG CCC GGG CCC GGG AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA GCG CGC GCG CGC GCG	6c, S6b
15ds/A30/15ds Rev	CGC GCG CGC GCG CGC AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA CCC GGG CCC GGG CCC	6c, S6b
15-Hairpin/T30	GGG CCC GGG CCC GGG TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT CCC GGG CCC GGG CCC	6c, S6b
15-Hairpin/A30	GGG CCC GGG CCC GGG AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA CCC GGG CCC GGG CCC	SbB
15ds/T30 For	GGG CCC GGG CCC GGG TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT	6c, S6b
15ds/T30 Rev	TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT CCC GGG CCC GGG CCC	6c, S3b
3x9T-loop For	GGG CCC GGG C TTT TTT TTT T GCA ACT CTT G TTT TTT TTT T C GCG CGC GCG TTT TTT TTT T GGT ACT AGA T	6c, S6b
3x9T-loop Rev	A TCT AGT ACC A TTT TTT TTT T CGC GCG CGC G TTT TTT TTT T C AAG AGT TGC TTT TTT TTT T G CCC GGG CCC	6c, S6b

3x9T-bulge Rev	A TCT AGT ACC A TT CGC GCG CGC G TT C AAG AGT TGC TT G CCC GGG CCC	6c
3'-15ds Rev*	CGC GCG CGC GCG CGC	6c
5'-15ds Rev*	CCC GGG CCC GGG CCC	6c
15T/15ds/15T/15ds For	TTT TTT TTT TTT TTT GGG CCC GGG CCC GGG TTT TTT TTT TTT TTT GCG CGC GCG CGC GCG	6c
15T/15ds/15T/15ds Rev	TTT TTT TTT TTT TTT CGC GCG CGC GCG CGC TTT TTT TTT TTT TTT CCC GGG CCC GGG CCC	6c

*Used together with "*15ds/T30/15ds For"* to form 15ds/1xT30/15ds used in Fig 6c. **Bold** sequences represent restriction site overhangs (Bam HI in For and Xho I for Rev), not used in this study. Underlined sequences represent highlight motif repeats.

Aarum: "**Enzymatic degradation of RNA causes widespread protein aggregation in cell and tissue lysates**"

Figures 1-6

Figure 1

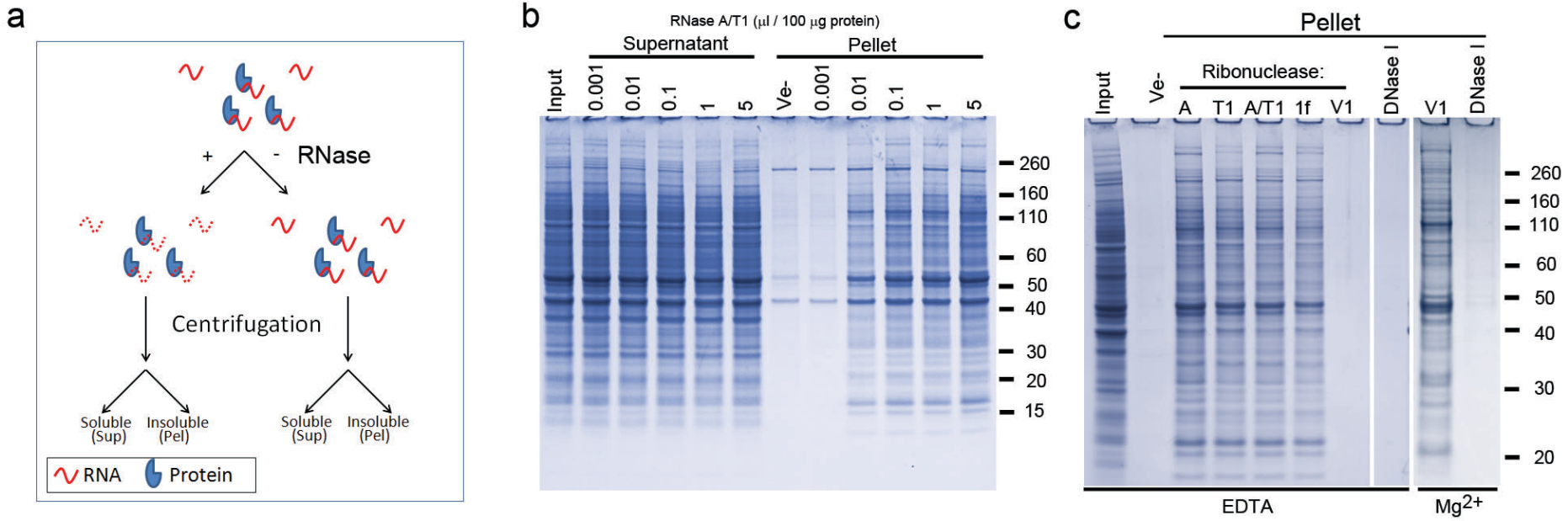


Figure 2

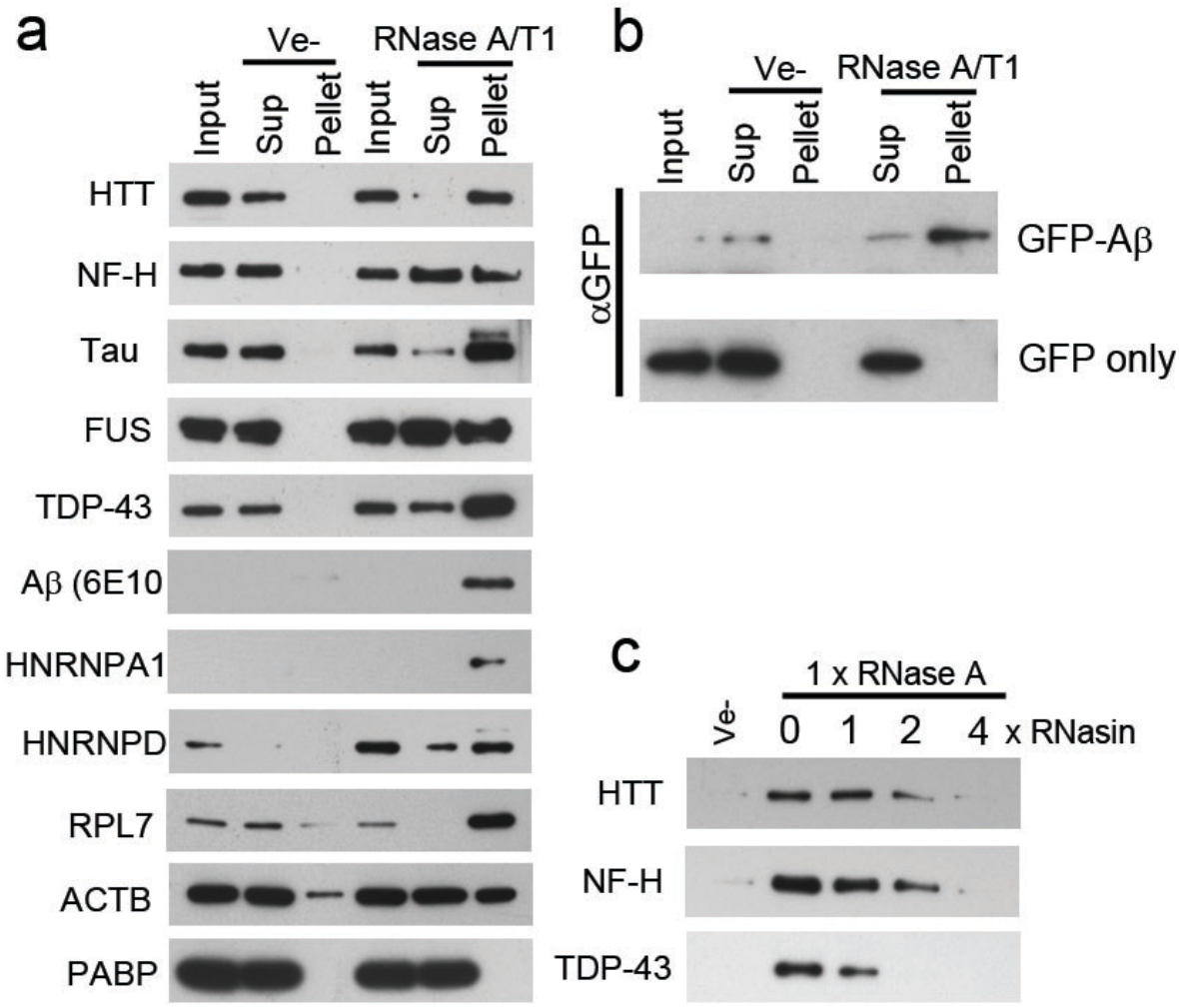


Figure 3

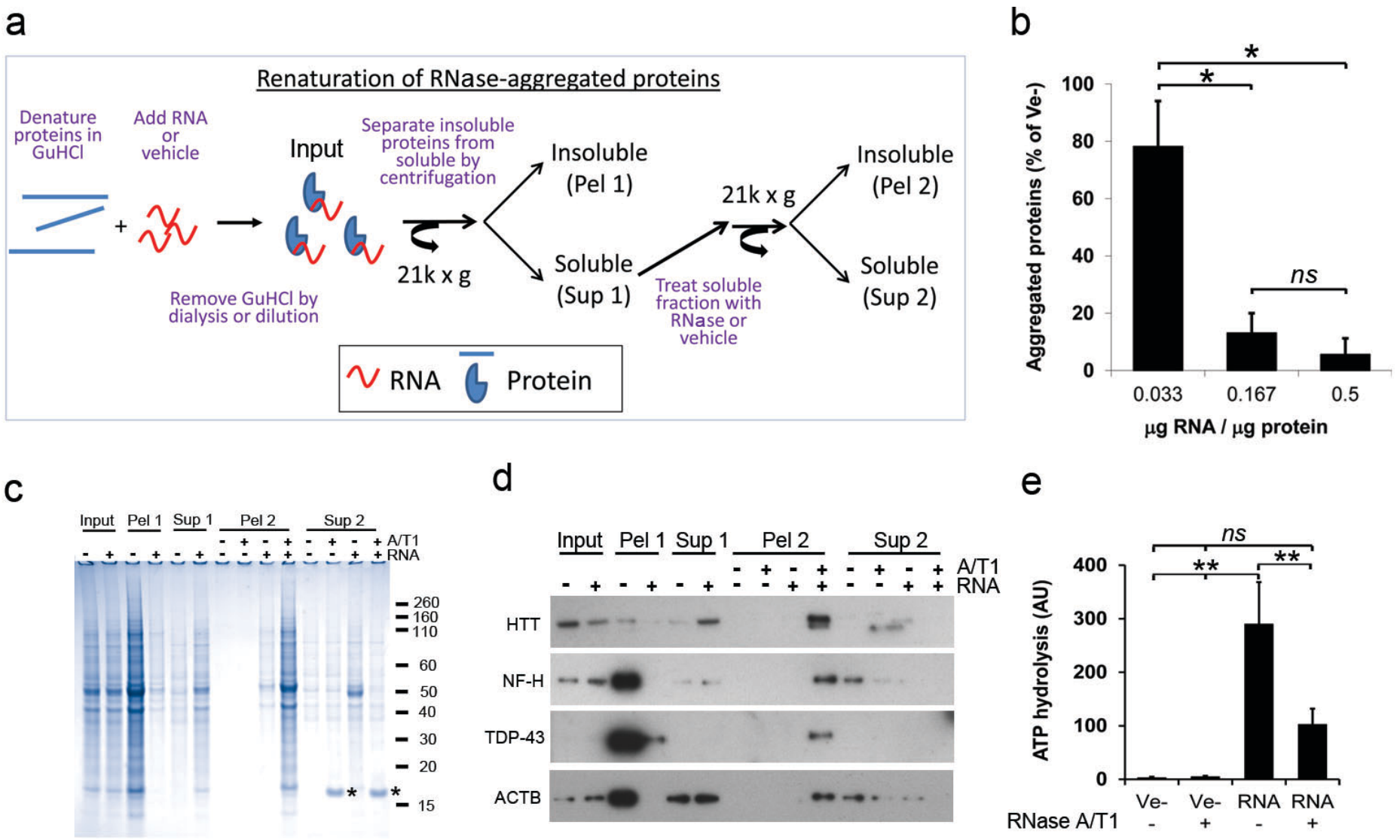

a
Renaturation of RNase-aggregated proteins
Denature proteins in GuHCl
Add RNA or vehicle
Input
Separate insoluble proteins from soluble by centrifugation
Insoluble (Pel 1)
Soluble (Sup 1)
21k x g
Remove GuHCl by dialysis or dilution
21k x g
Insoluble (Pel 2)
Soluble (Sup 2)
Treat soluble fraction with RNase or vehicle
RNA
Protein
b
Aggregated proteins (% of Ve-)
100
80
60
40
20
0
*
*
ns
0.033
0.167
0.5
µg RNA / µg protein
c
Input
Pel 1
Sup 1
Pel 2
Sup 2
A/T1
RNA
260
160
110
60
50
40
30
20
15
d
Input
Pel 1
Sup 1
Pel 2
Sup 2
A/T1
RNA
HTT
NF-H
TDP-43
ACTB
e
ns
**
**
ATP hydrolysis (AU)
400
300
200
100
0
Ve-
Ve-
RNA
RNA
RNase A/T1

Figure 4

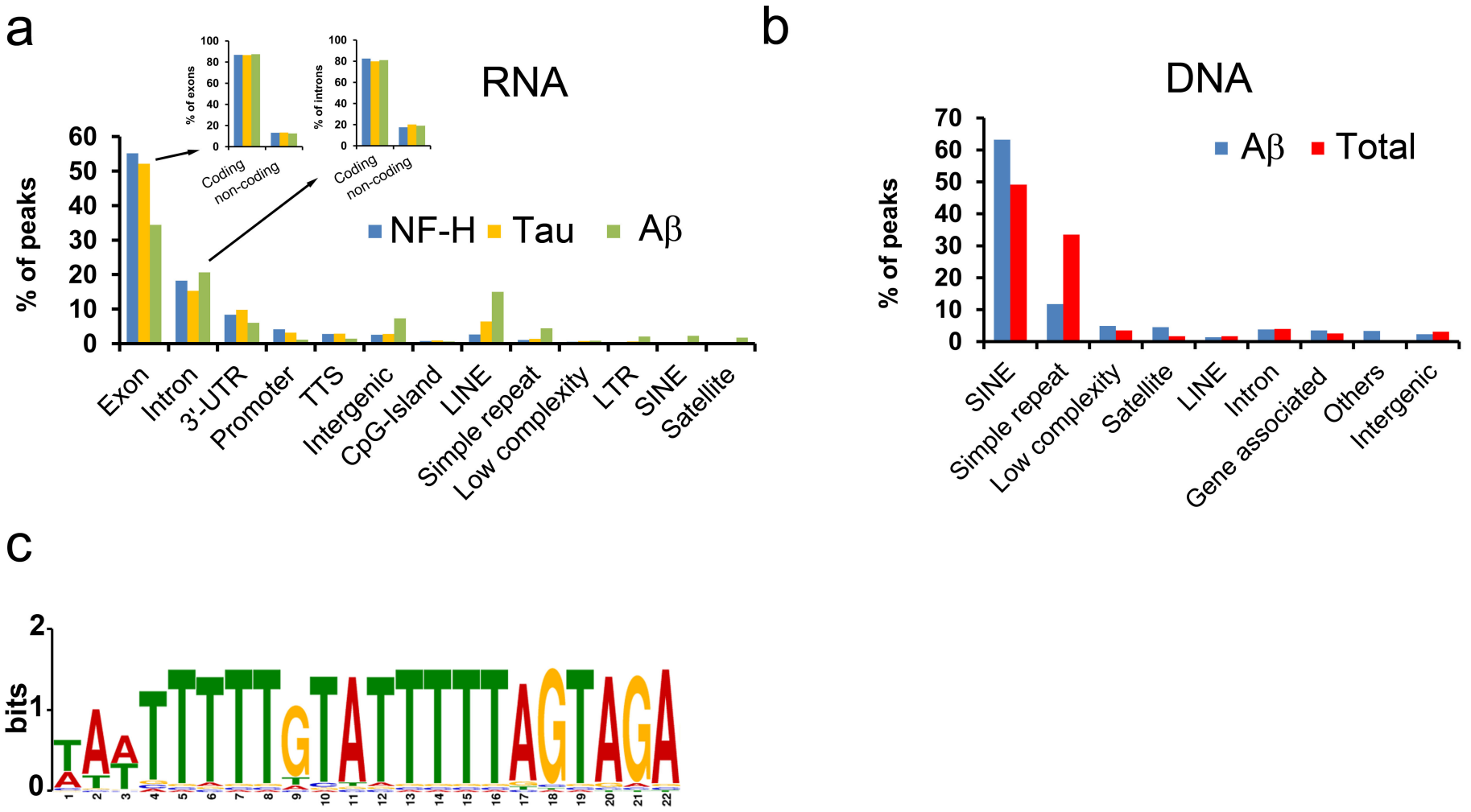

a
RNA
% of peaks
60
50
40
30
20
10
0
NF-H
Tau
Aβ
% of exons
% of introns
100
80
60
40
20
0
Coding
non-coding
Exon
Intron
3'-UTR
Promoter
TTS
Intergenic
CpG-Island
LINE
Simple repeat
Low complexity
LTR
SINE
Satellite
b
DNA
% of peaks
70
60
50
40
30
20
10
0
Aβ
Total
SINE
Simple repeat
Low complexity
Satellite
LINE
Intron
Gene associated
Others
Intergenic
c
bits
2
1
0

Figure 5

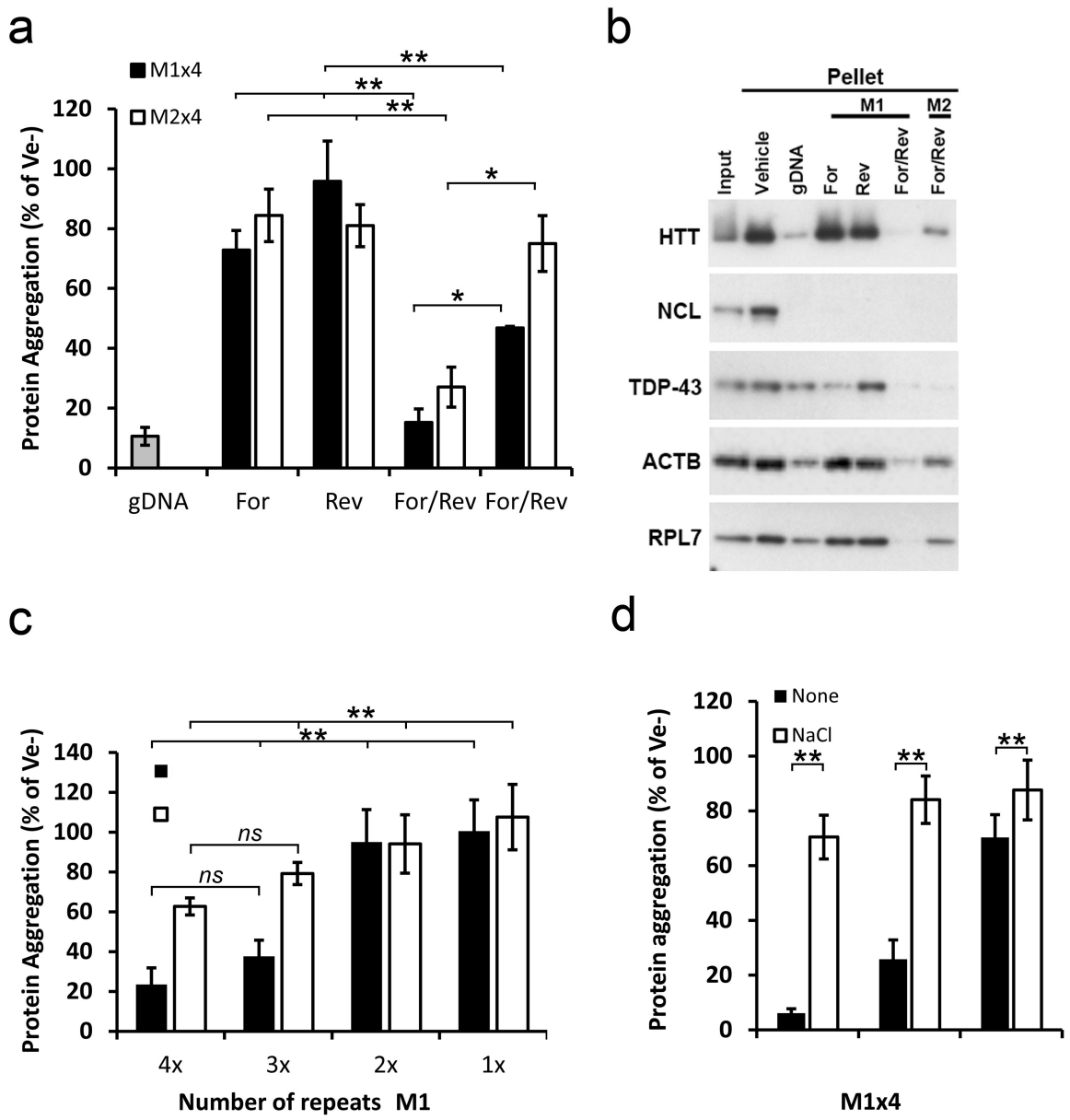
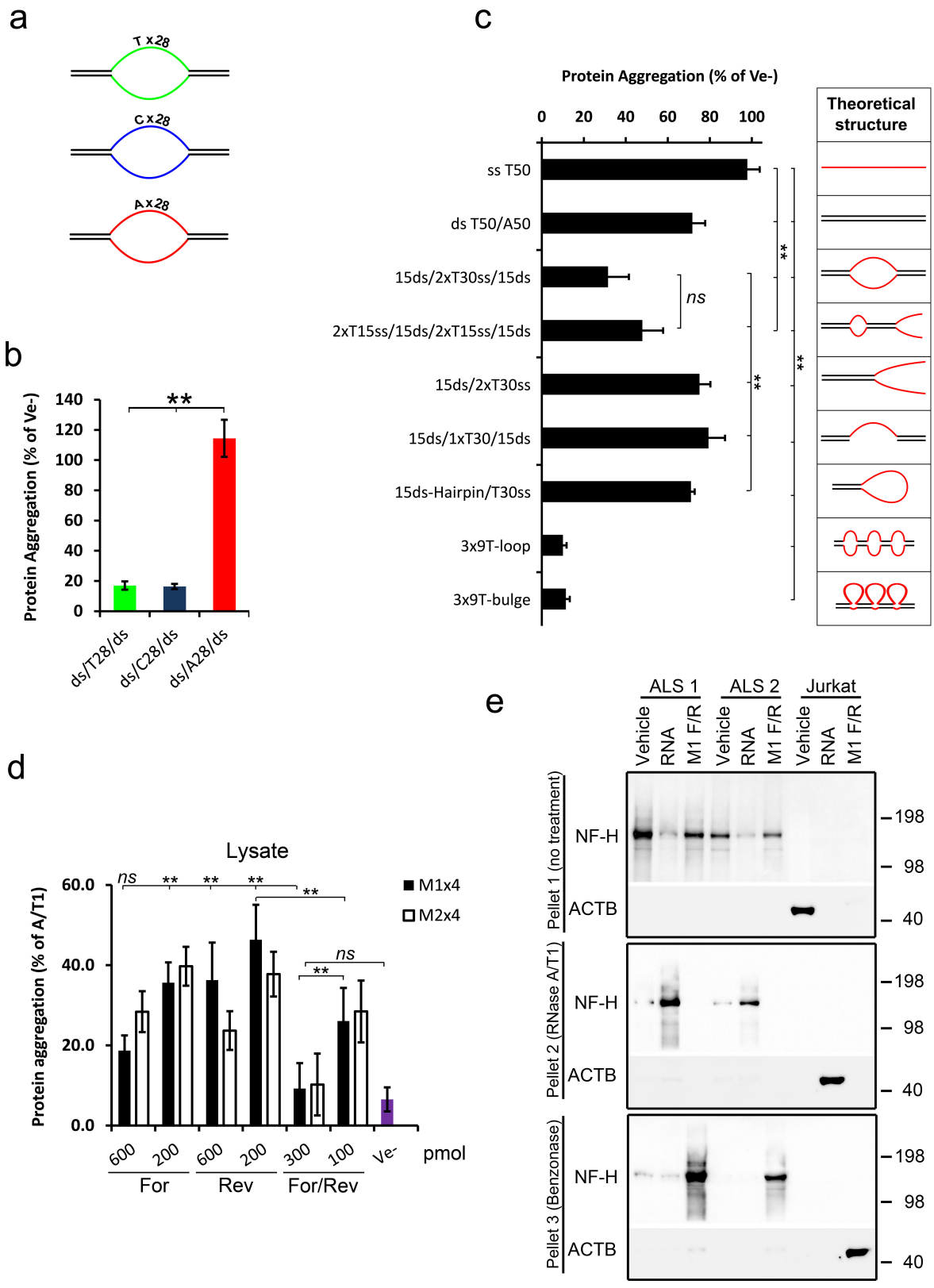
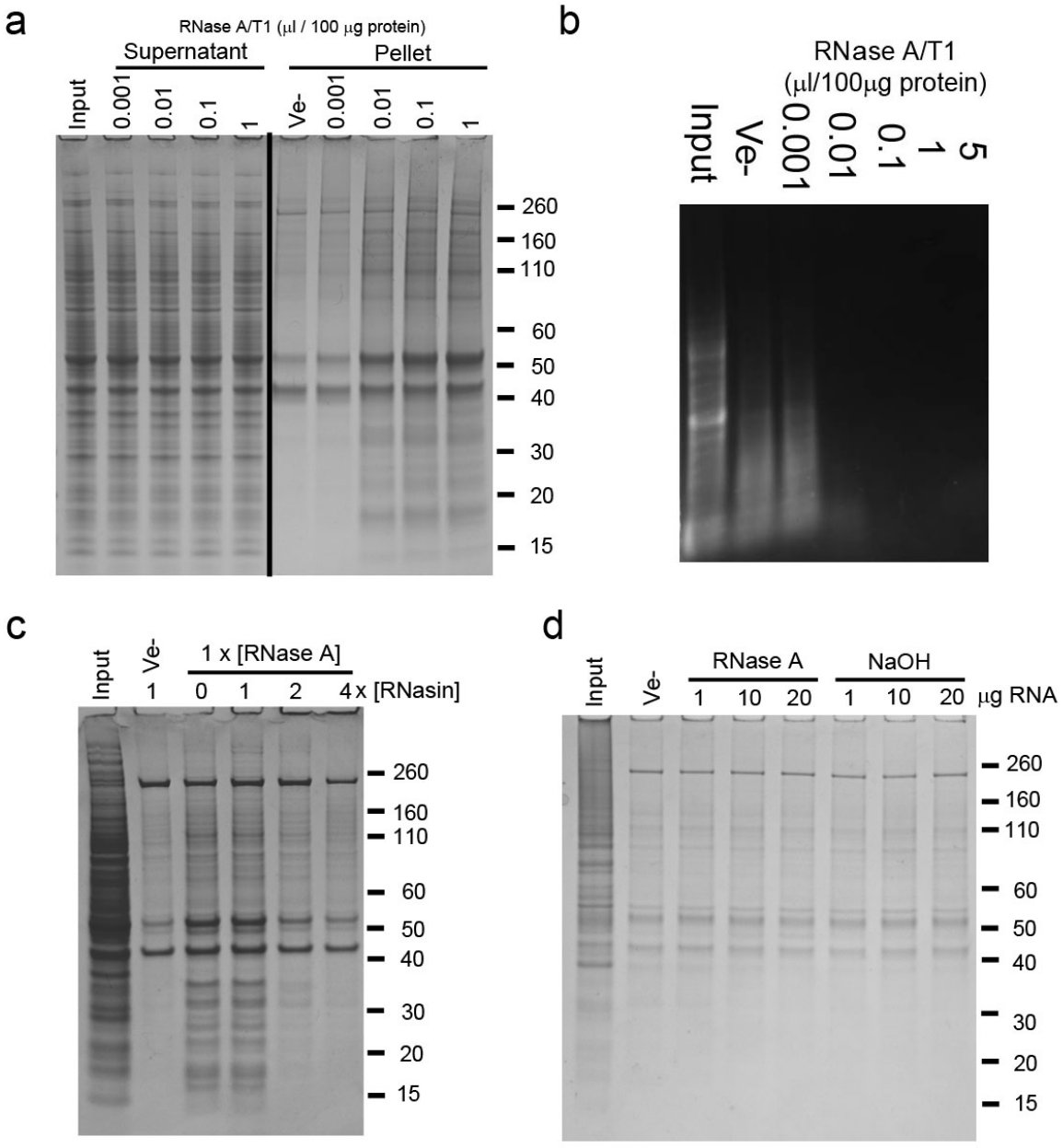


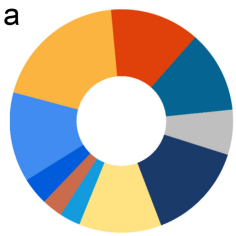
Figure 6



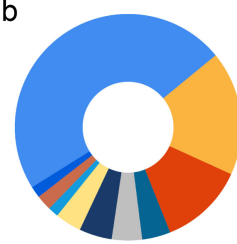
Aarum: "**Enzymatic degradation of RNA causes widespread protein aggregation in cell and tissue lysates". Expanded View Figures 1-7**



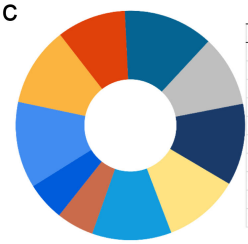
Expanded View Figure 1. Protein aggregation caused by enzymatic degradation of RNA, and effect of pre-hydrolysed RNA or RNase inhibition. a SDS-PAGE analysis of soluble (Supernatant) and aggregated (Pellet) proteins after treatment of mouse brain tissue lysate with increasing amounts of a mixture of RNase A and T1 (A/T1) or vehicle (Ve-). **b** Agarose gel electrophoresis analysis of RNA isolated from RNAse-treated human neuronal cell lysate. **c-d** Insoluble proteins collected by centrifugation after co-treatment of human neuronal cell lysate with RNase A and an RNase A inhibitor (RNasin, **c**), or following the addition of RNA pre-hydrolysed by RNase A or alkaline hydrolysis (NaOH, **d**).



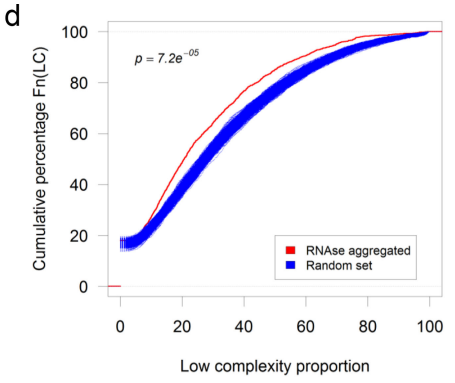
Cellular component	genes	%	FC	p value	q value
membrane	447	34.3	4.83	2E-196	1E-192
cytosol	665	51	2.83	2E-171	6E-168
extracellular vesicular exosome	446	34.2	3.31	8E-128	1E-124
nucleoplasm	407	31.2	2.8	2E-90	3E-87
mitochondrion	224	17.2	3.55	3E-65	4E-62
nucleus	491	37.7	1.99	1.4E-59	1E-56
cytoplasm	410	31.5	2.13	2.5E-55	2E-52
focal adhesion	103	7.9	5.31	5E-46	4E-43
mitochondrial inner membrane	103	7.9	4.91	1.1E-42	7E-40
nucleolus	136	10.4	3.39	7E-37	4E-34

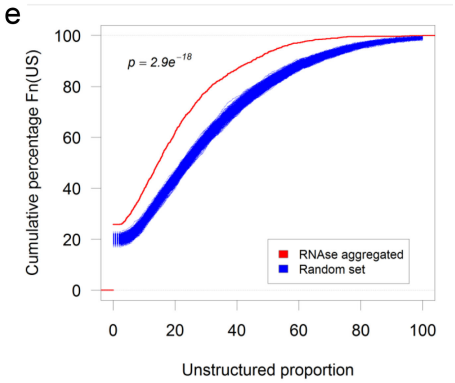


Molecular function	genes	%	FC	p value	q value
protein binding	932	71.5	2.14	5E-183	1E-178
RNA binding	345	26.5	5.18	2E-157	2E-153
ATP binding	236	18.1	3.24	3.6E-61	2E-57
structural constituent of ribosome	79	6.06	9.81	9.4E-59	5E-55
cadherin binding	84	6.45	6.16	4.6E-43	2E-39
GTP binding	91	6.98	4.94	5.3E-38	2E-34
GTPase activity	73	5.6	5.41	2E-33	6E-30
GDP binding	29	2.23	9.85	2.1E-22	5E-19
ATPase activity	43	3.3	5.62	6.7E-21	2E-17
unfolded protein binding	32	2.46	5.67	4.9E-16	1E-12

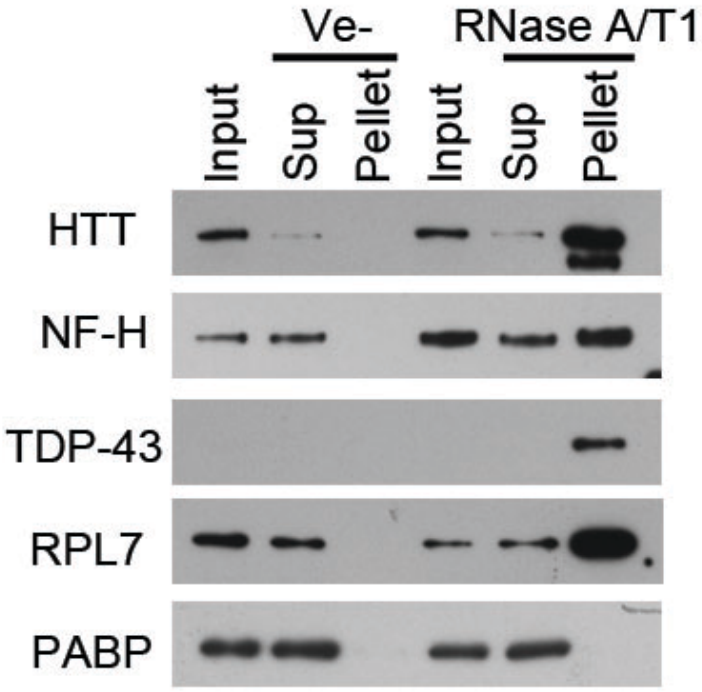


Biological process	genes	%	FC	p value	q value
translational initiation	74	5.68	11.4	1.7E-61	5E-57
viral transcription	68	5.22	12.8	1.8E-61	5E-57
SRP-dependent cotranslational protein targeting	59	4.53	13.4	4.9E-55	9E-51
translation	77	5.91	8.81	7.2E-53	1E-48
nuclear-transcribed mRNA catabolic process, NMD	61	4.68	10.7	1.4E-48	2E-44
rRNA processing	70	5.37	7.1	1.3E-40	1E-36
nuclear mRNA splicing, via spliceosome	65	4.99	5.59	8.7E-31	7E-27
viral reproduction	68	5.22	4.68	4E-27	3E-23
mitochondrial translational elongation	32	2.46	7.49	4.1E-20	2E-16
regulation of mRNA stability	33	2.53	6.4	2.8E-18	1E-14

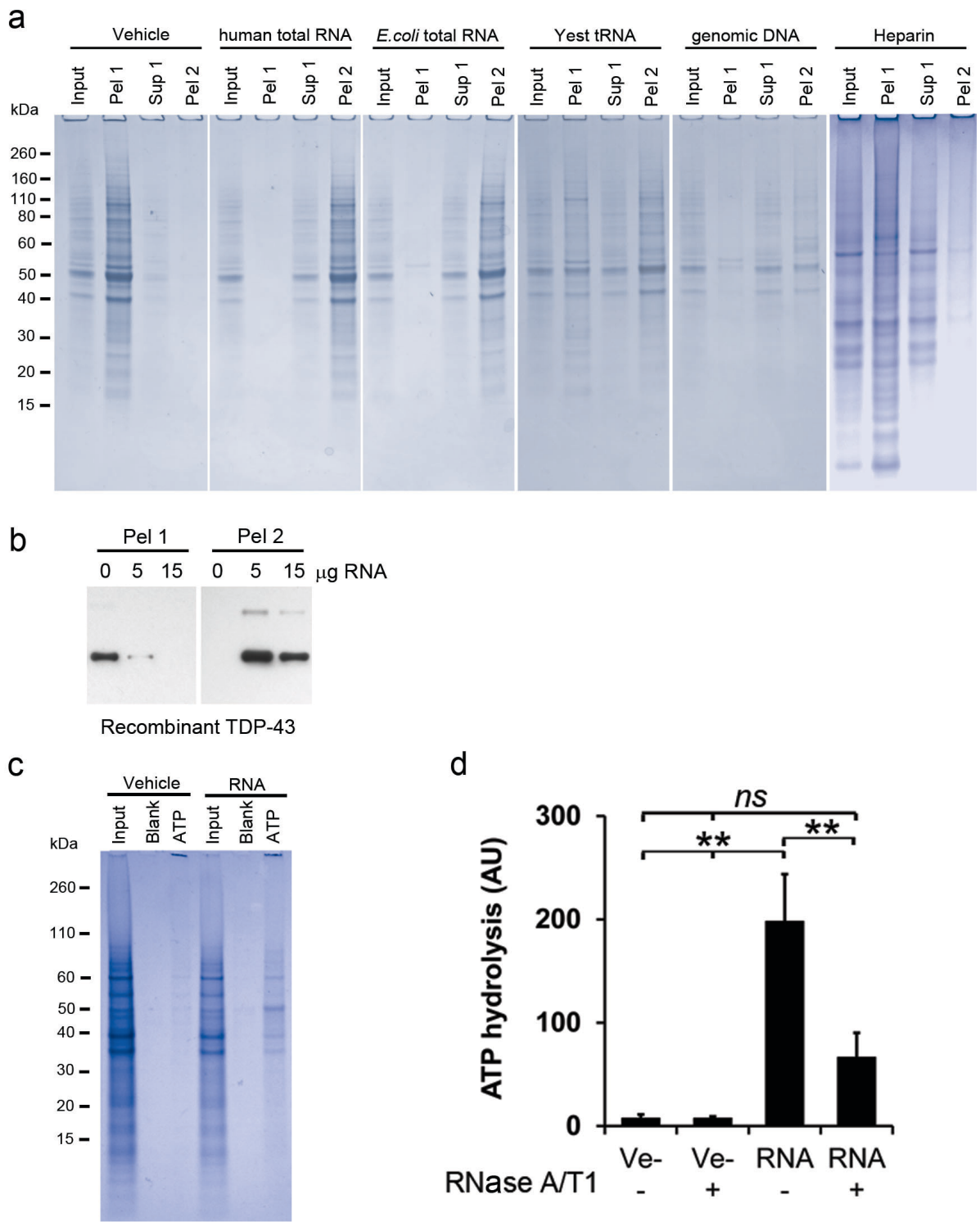




Expanded View Figure 2. Computational analysis of proteins aggregated by enzymatic degradation of RNA. (**a-c**) Top ten gene ontology classes by Cellular component **a**, Molecular function **b**, or Biological process **c**. **d-e** Cumulative distribution of the proportion of predicted low-complexity regions **d** or unstructured regions **e** in the RNase-aggregated proteins (Red) or random sets of proteins (Blue). *p*-values in **d-e** were obtained by the two-sample Kalmogrov-Smirnov test and corrected for multiple testing using Bonferroni correction.

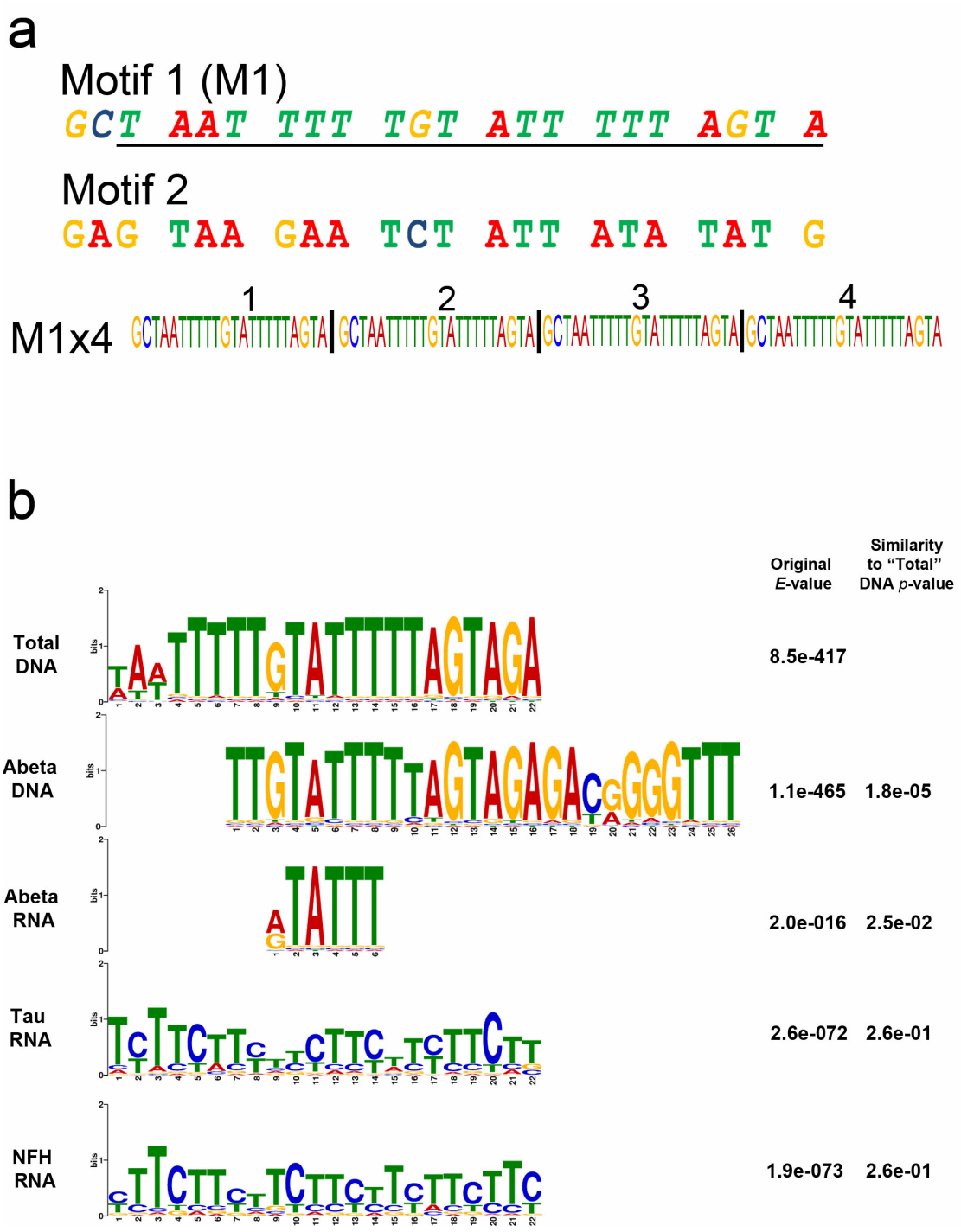


Expanded View Figure 3. Protein aggregation caused by enzymatic degradation of RNA in mouse brain tissue lysate. Western blot analysis of aggregated proteins, collected by centrifugation, after treatment of mouse brain lysate with a mixture of RNase A and RNase T1 (RNase A/T1), or vechicle (Ve-). Input represents the starting material, Sup the supernatant (soluble fraction), and Pellet the aggregated fraction after centrifugation.



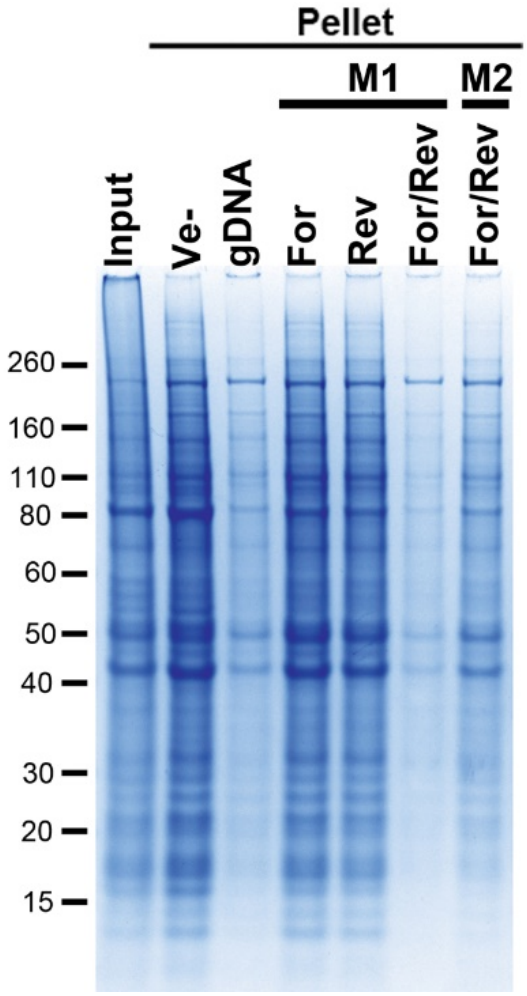
Expanded View Figure 4. (legend on next page)

Expanded View Figure 4. Protein renaturation: effects on enzymatic activity and effects of various polyanions. a Coomassie-stained SDS-PAGE gels of RNase-aggregated proteins from human neurons renatured with various polyanions. Input represents 1/10th of each sample, taken before the first centrifugation, and Pel 1 and Sup 1 are the pellet and the supernatant, respectively, recovered after the first centrifugation. Pel 2 represents the pellet obtained after treating the Sup 1 fraction with RNase A/T1 followed by centrifugation, please see experimental outline in Fig. 3a for a full description. **b** Renaturing of recombinant TDP-43 using increasing amounts of total RNA. Pel 1 represents aggregated protein after renaturing. Pel 2 is the aggregated protein after the soluble fraction (Sup 1) has been treated with RNase A/T. **c** Gel-electrophoresis analysis of proteins from human neurons renatured with vehicle or total RNA and then captured with ATP-agarose beads (ATP). Input represents 1/10th of each sample taken directly after renaturation, before the first centrifugation, and Blank represents the sample without any ATP-agarose beads present. **d** ATP-hydrolysing activity of RNase-aggregated proteins from Jurkat T-cells renatured in the presence (RNA), or absence (Ve-), of RNA. Data in **c** are expressed in arbitrary units (AU) and represent the mean ± s.d of three independent experiments. **=p<0.01, *=p<0.05 by ANOVA post-hoc analysis using Bonferroni correction for multiple testing.

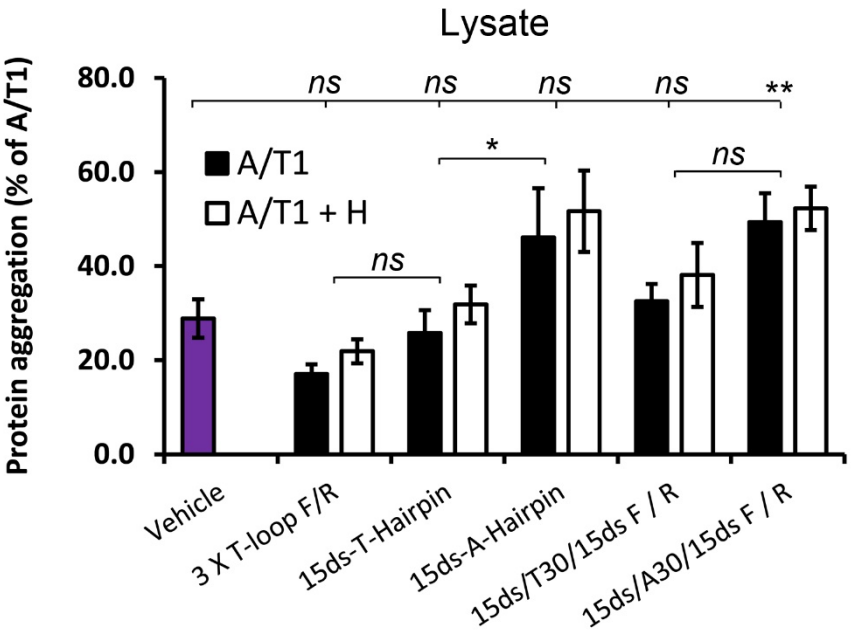


Expanded View Figure 5. (legend on next page)

Expanded View Figure 5. Comparison of identified motifs. a Sequence of the selected M1 motif and a random control oligo (Motif 2) investigated for their renaturing capacity. Both motifs were assembled into four consecutive repeats of the same motif (M1x4). The underlined region of M1 represents the computationally identified motif while the two extra 5' nucleotides (GC) where added for cloning purposes. **b** Alignment of motifs identified in the nucleic acids associated with soluble proteins after renaturation. All motifs are compared to the M1 motif identified in the Total DNA samples. The quoted original E-value represents the value obtained in the MEME-Chip analysis and the Similarity *p*-value is the *p*-value given by the TomTom motif comparison software package, compared to the Total DNA motif.



Expanded View Figure 6. Renaturation effects of selected oligonucleotides. Coomassie-stained SDS-PAGE gel of aggregated proteins (Pellet) following renaturation with either single-stranded (For or Rev) or complementary (For/Rev) M1x4 or M2x4 DNA oligonucleotides. gDNA represents genomic DNA, and Ve- represents vehicle.



Expanded View Figure 7. Protein aggregation in cell lysate supplemented with structural oligonucleotides and treated with RNase. Amount of protein aggregation observed after treating cell lysate from Jurkat T-cells with various oligonucleotides and RNase A/T1 (A/T1) and, if indicated, together with RNase H (A/T1 + H). Data are expressed as a percentage of the aggregation observed in samples treated with RNase A/T1 only and represent the mean ± s.d of three independent experiments. **=p<0.01 by ANOVA post-hoc analysis using Bonferroni correction for multiple testing.