Nucleic Acids Research, 2023, 1–13
https://doi.org/10.1093/nar/gkad857
Computational Biology

RNA sequestration driven by amyloid formation: the alpha synuclein case

Jakob Rupert[1,2,†], Michele Monti[1,†], Elsa Zacco[1,*] and Gian Gaetano Tartaglia[1,2,3,*]

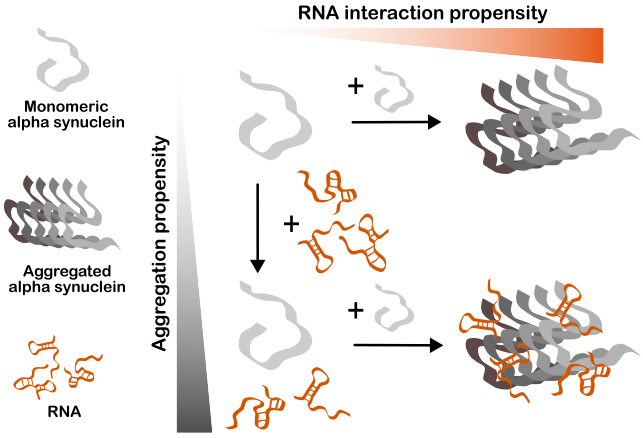
[1]Centre for Human Technologies (CHT), Istituto Italiano di Tecnologia (IIT), Via Enrico Melen, 83, 16152, Genova, Italy
[2]Department of Biology and Biotechnologies 'Charles Darwin', Sapienza University of Rome, P.le A. Moro 5, Rome 00185, Italy
[3]Catalan Institution for Research and Advanced Studies, ICREA, Passeig Lluís Companys 23, 08010, Barcelona, Spain

[*]To whom correspondence should be addressed. Email: elsa.zacco@iit.it
Correspondence should be addressed to Gian Gaetano Tartaglia. Tel: +39 010 28976204; Fax: +39 010 2897621; Email: gian.tartaglia@iit.it
[†]The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

Abstract

Nucleic acids can act as potent modulators of protein aggregation, and RNA has the ability to either hinder or facilitate protein assembly, depending on the molecular context. In this study, we utilized a computational approach to characterize the physico-chemical properties of regions involved in amyloid aggregation. In various experimental datasets, we observed that while the core is hydrophobic and highly ordered, external regions, which are more disordered, display a distinct tendency to interact with nucleic acids. To validate our predictions, we performed aggregation assays with alpha-synuclein (aS140), a non-nucleic acid-binding amyloidogenic protein, and a mutant truncated at the acidic C-terminus (aS103), which is predicted to have a higher tendency to interact with RNA. For both aS140 and aS103, we observed an acceleration of aggregation upon RNA addition, with a significantly stronger effect for aS103. Due to favorable electrostatics, we noted an enhanced nucleic acid sequestration ability for the aggregated aS103, allowing it to entrap a larger amount of RNA compared to the aggregated wild-type counterpart. Overall, our research suggests that RNA sequestration might be a common phenomenon linked to protein aggregation, constituting a gain-of-function mechanism that warrants further investigation.

Graphical abstract



Introduction

The order and structural robustness of amyloids render them a powerful platform for attracting other cellular factors and macromolecules (1,2). Nucleic acids are embroiled in the modulation of protein phase transition (3,4), as recent experiments show that they can be isolated from both physiological condensates and pathological inclusions (5,6). RNA in particular might affect aggregate accumulation rate and nature (7). In support to this observation, numerous proteins found aggregated in neurological disorders are RNA-binding and form ribonucleoprotein assemblies (1), and their ability to aggregate appears to be at least in part determined by the type and amount of RNA found in their proximity (4). Yet, only a few studies have focused on understanding the ability of proteins to interact with nucleic acids after the transition towards amyloid species has taken place, or whether such transition might confer the ability to attract nucleic acids to proteins not classified as RNA/DNA-binding (8).

Many canonical RNA-binding proteins (RBPs) exhibit a modular structure, comprising one or more RNA-binding domains (RBDs) and typically at least one prion-like domain that can facilitate phase transitions (9). This phenomenon

Received: April 25, 2023. **Revised:** August 15, 2023. **Editorial Decision:** September 14, 2023. **Accepted:** September 26, 2023

describes a process whereby soluble complexes of macromolecules, upon reaching their solubility threshold, transition from a reversible, liquid-like state into an irreversible, and frequently toxic, solid-like state. This transformation in proteins is often referred to as aggregation, and it underscores the intricate behavior of macromolecular systems, revealing insights into both their functionality and potential pathology. Prion regions represent the core of the amyloid fibril structure upon aggregation (10,11) and the RBDs, usually found at least partially folded in the aggregated species, are external to the core and might retain RNA-binding ability (12). This is the case of hnRNPDL-2, for which a cryo-EM structure of the aggregate shows two RBDs in a disordered, solvent-exposed portion of fibrils outside of the amyloid core (12)—the 'fuzzy coat'. hnRNPDL-2 fibrils have a comparable affinity towards oligonucleotides both in the monomeric and the aggregated form (12), confirming that canonical RBP aggregates can retain their functionality even upon partial refolding into different conformations.

For non-canonical RBPs and proteins that do not display RNA binding activity, the functional consequences of amyloid formation are less clear. In the case of microtubule-associated protein tau, it has been shown that fibrils can sequester RNA both in cell (6) and *in vitro* (13), implying the protein conformational changes upon aggregation can promote RNA binding. This observation is further strengthened by the fact that tau fibrils, grown *in vitro* and without any additional cofactors, can sequester RNA (14). We note that tau is not known as a canonical RBP, however it has been shown that RNA colocalises with this protein (15) and promotes its aggregation (16) into morphologically distinct fibril strains (17). A recent cryo-EM structure has shown that RNA can also trigger the formation of a previously undetected tau aggregate morphology with a central core spanning residues E391 to A426 (13). In previously solved tau fibril structures, this specific segment has been reported to be part of the fuzzy coat, implying that RNA can actively change the fibril morphology. Moreover, electron density showed an unknown density that could correspond to the RNA molecules being sequestered in the fibrils.

In this work, we use a computational approach to investigate nucleic acid sequestration in protein aggregates by considering the physico-chemical properties of both their amyloid cores, hidden within the fibers and occupied in cross-β bonds, and their exterior, available to establish novel interactions and exert new functions. We validated our model on alpha synuclein (aS), which represents a particularly suitable case, since its aggregation is abundantly studied and the protein does not bind RNA in the monomeric state under physiological conditions. aS is an intrinsically disordered protein involved in synaptic vesicle trafficking and machinery assembly, mitochondrial homeostasis and DNA repair (18–20). It has been identified as the primary component of Lewy bodies (21), pathological aggregates found in certain neurodegenerative conditions such as Parkinson's Disease (PD) and Multiple System Atrophy (MSA) (21–23). aS has a modular organization, with a basic N-terminus (residues 1–95), containing the amyloid central region (NAC, residues 61–95) (24), and an acidic, intrinsically disordered C-terminus (residues 96–140), shown to negatively affect fibrillation and oligomerization (25–27). The C-terminal domain can undergo several post-translational modifications such as truncations at various residues (28,29), which occur naturally in the cell, presumably as incomplete protein degradation products. However, certain disease-related C-terminal truncations also boost the protein fibrillation rates and result in diverse aggregate morphologies, leading to an increase in cellular toxicity (26,30–32). In the full length protein, long-range electrostatic interactions between the N-terminus and C-terminus influence the partially folded intermediary states of the protein and are highly affected by changes in pH and ionic strength (27,33–35). These factors, as well as the presence of various charged polymers, strongly affect the aggregation rates, highlighting the role of electrostatic charge in aS misfolding (36–39). In the absence of the acidic C-terminus, such in the case of disease-related truncations, it is possible that polyanions such as RNA could have an even more significant impact on aS aggregation.

Based on our calculations, we posit that proteins capable of forming amyloids, such as aS, may acquire the ability to sequester RNA as they aggregate. We put forth that RNA binding is not an inherent ability of aS but rather an acquired trait of its aggregates, and we believe that this distinction plays a pivotal role in explaining the observed trend in the modulation of aggregation rates. More broadly, our findings suggest that the process of misfolding during aggregation can alter the RNA-binding capability of a given protein. This phenomenon could potentially enable a protein that ordinarily does not bind to RNA to acquire the ability to do so in its aggregated form, expanding our understanding of protein-RNA interactions and their role in cellular biology.

Materials and methods

Aggregation propensity and analysis of the inner and outer regions of amyloids

Previous work indicates that residues with high aggregation propensities are buried within the amyloid fibrils, while those with low aggregation propensity are exposed to the solvent (40). In Figure 1, we present the comparison of predicted and experimentally validated inner and outer regions of the amyloid structures of calcitonin (41,42), amylin (43), glucagon (44), Aβ42 (45,46), alpha synuclein (47–49), hnRNPDL-2 (12) and tau (13). The optimal threshold for distinguishing the inner and outer regions of amyloids was identified using the Zyggregator score of 0.8. This value corresponds to the Youden index within the Receiver Operating Characteristic (ROC) analysis and corresponds to a True Positive Rate (TPR) of 0.75 and a False Positive Rate (FPR) of 0.25. Notably, the TPR and FPR values demonstrate stability within the Zyggregator range of 0.76–0.84. The balance between sensitivity and specificity is further reflected in the overall Area Under the ROC Curve (AUC), which is quantified at 0.84. For every individual residue in the sequences analyzed (Figure 2A-C and Supplementary Table S1), we computed the Zyggregator scores (40) using the 0.8 score threshold to distinguish between inner and outer regions of amyloids. Zyggregator enables the prediction of aggregation propensity by calculating hydrophobicity and secondary structure properties (40).

RNA/DNA binding domains and aggregation

Annotations of RNA/DNA binding domains available from UniProt were exploited to characterize their physicochemical properties in relation to the aggregation propensity (Supplementary Tables S2 and S3, Figure 2G). Specifically, the fraction of RNA/DNA domains belonging to the predicted outer region of the amyloids was computed. The same

analysis was carried out for prion domains predicted using PLAAC (50), as in a previous work (9) (Figure 2).

Mutagenesis

An *in silico* random mutation analysis was performed to define the role of each residue of the aS sequence in the overall aggregation propensity of the protein (Figure 3B). Random positions in the sequence were selected and replaced, one by one, with amino acids from an uniform distribution. The Zyggregator algorithm (40) was used to compute the aggregation propensity of the protein before and after the mutation (40). 10 000 random mutations were calculated and, for each one, the differences with the wild type in terms of Zyggregator score, expressed as a percentage, were calculated.

Physico-chemical properties of the outer amyloid regions

Calculations of hydrophobicity (51), disorder (52) and RNA-binding ability (53) were carried out through the *clever*Machine approach that predicts physico-chemical properties of protein sequences (54) (Figure 1-3). Data reported in the Supporting Information (Supplementary Figure S1 and S2) show that the selected scale used for nucleic acids binding and hydrophobicity is not strictly important, as the behavior of the histograms of the chemical physical property does not change when different scales are applied.

*cat*RAPID predictions of protein-RNA interactions

In the *cat*RAPID algorithm, the interaction propensity between a protein-RNA pair is computed using secondary structure properties, van der Waals forces and hydrogen bonding potentials (55). The algorithm is able to discriminate between interacting and non-interacting pairs with an area under the curve (AUC) on a receiver operating characteristic (ROC) curve of 0.78, and with false discovery rate (FDR) significantly below 0.25 when the *Z*-score values are >2 (56). In this work, the N-terminus of aS (residues 1–25) is predicted to attract RNA, while the C-terminal region (residues 104–140) shows much poorer propensity. For the calculations presented in this work (57), the whole human transcriptome was used and the overall interaction propensity (>0.4, corresponding to low binding affinity) for different aS regions was computed. TDP43, TRA2A and SRSF3 are used as positive controls. Full length aS103 and aS140 are not predicted to interact with RNA as monomers (Figure 3E).

Protein production

The cDNA for aS140 was inserted in a pET21a vector under the control of a T7 promoter. aS103 cDNA was produced by PCR-led deletion of the appropriate sequence from the full length construct with specific primers (aS103 forward primer: 5′-TAAGCGGCCGCACTCGAGCAC-3′ and aS103 reverse primer: 5′-ATTCTTGCCCAACTGGTCCTTTTTGA CAAAGC-3′). Competent *Escherichia coli* BL21 [DE3] cells were transformed with both constructs. 2 l of fresh LB broth supplemented with 50 μg/ml ampicillin were inoculated with overnight cultures with a 100:1 ratio. Cultures were grown at 37°C up to an optical density of ca. 0.7 at 600 nm. Protein expression was induced with 0.5 mM IPTG and the cultures were shaken for 4 h. Cells were harvested by centrifugation at 4500 g at 4°C for 30 min, afterwards the pellet was resuspended on ice in 20 mM potassium phosphate buffer pH 7.2 and frozen immediately. For protein purification the cells were thawed in ice-cold water and lysed by boiling at 95°C for 30 min. The boiled suspension was centrifuged at 18 000 g at 4°C for 30 min, then streptomycin sulfate was added to the supernatant up to a concentration of 10 mg/ml. The solution was centrifuged at 18 000 g at 4°C for 30 min and the pellet containing precipitated impurities was discarded. The protein was slowly precipitated by addition of crystalline ammonium sulfate up to 360 mg/ml and the suspension was centrifuged again at 18 000 g at 4°C for 30 min. The pellet was resuspended in 20 mM Tris–HCl pH 7.4 with 400 mM KCl and dialysed against 20 mM Tris–HCl pH 7.4 overnight at 4°C. aS140 was purified using ion exchange chromatography using a HiTrap™ Q FastFlow column, while aS103 was purified via heparin affinity chromatography on a HiTrap™ Heparin column. Proteins were eluted with heparin elution buffer (20 mM potassium phosphate pH 7.2, 800 M NaCl) and loaded onto a HiLoad™ 16/600 Superdex™ 75 gel filtration column equilibrated in 20 mM potassium phosphate pH 7.2, 100 mM KCl. The eluted proteins were pooled, their concentration was determined with BCA assay and their purity assessed via spectropolarimetry and SDS PAGE, before being stored aliquoted at –80°C.

Protein aggregation assays

A frozen protein aliquote of aS140 or aS103 was quickly thawed and filtered with a 0.22 μm syringe filter immediately before each assay. Total yeast RNA (totRNA, Roche) was resuspended in aggregation assay buffer (20 mM potassium phosphate pH 7.2, 100 mM KCl, 5 mM $MgCl_2$). For aggregation assays in the presence of DNA, total DNA from herring sperm (Sigma Merck) was resuspended in sterile, nuclease-free water for the stock solution. Protein, totRNA and DNA were diluted to the final concentration in the aggregation assay buffer with a 1000-time diluted Proteostat™ detection dye. For the BSA control, ultra-pure, DNA- and RNA-free BSA was used (Sigma Merck) and diluted to the final concentration in the sample buffer as described above for aS. The samples were distributed in 6 replicates on a 96-well black plate with transparent bottom and a single borosilicate glass bead was added to each well to ensure sample homogeneity and reproducibility. The excitation wavelength was set to 505 nm and the emission to 590 nm. Plates were incubated at 37°C with constant double-orbital shaking at 200 rpm. Reads were taken every 10 min for aS140 and every 5 min for aS103 for 24 h, with each read represented as an average of 10 scans. All protein samples were aggregated in parallel to total RNA alone at the same starting concentrations and conditions for RNA degradation control and the values reported have already been accounted for degradation (Figure 4A–C; Supplementary Figures S3–S5).

Aggregation data analysis

Data from the aggregation assays, obtained with Proteostat™ fluorescence dye, were fitted into a sigmoid curve using the Hill function (1) (Figure 4A–C):

$$h\ (x)\ =\ A\ \frac{x^h}{x^h\ +\ K^h} \qquad (1)$$

where A, h and K are the fitting parameters.

To compute the aggregation rate, the slope of the fitted curves before plateau was determined. With the previously set parameters and a Taylor expansion of $h\ (x)$ around K, the slope of the aggregation α curve was calculated as (2):

$$\alpha \equiv h\ (x = K) = h\ (K) + \frac{\partial h\ (x)}{\partial x}; x = K$$

$$\alpha = A\frac{h}{2K} \tag{2}$$

The 'delay time' was defined as the intersection of the fitting line with slope α centered in $x = K$ and $y = \frac{A}{2}$. The delay time τ was defined resolving the linear system (3):

$$y\ (x) = \alpha x + q$$

$$y\left(\frac{K}{2}\right) = \alpha\frac{K}{2} + q = \frac{A}{2} \rightarrow q = \frac{A - \alpha K}{2}$$

$$\alpha\tau + \frac{A - \alpha K}{2} = 0 \rightarrow \tau = \frac{\alpha K - A}{2\alpha}$$

$$\tau = K\frac{(h - 1)}{h} \tag{3}$$

Nucleic acid quantification and extraction from protein aggregates

After aggregation, individual replicates were transferred from the 96-well plate to 1,5 mL Eppendorf tubes and centrifuged at 4°C, 18 000 g for 30 min. The soluble fraction was transferred to another tube, RNA found in solution was quantified using Qubit RNA BR, and DNA with Qubit DNA BR kit (Thermo Fisher Scientific, USA) (Figure 4D–F; Supplementary Figure S6). For RNA extraction, the insoluble pellet was resuspended in 25 μl of digestion buffer (2 M urea, 100 μg/ml proteinase K, 3 mM DTT) and incubated at room temperature for 15 min. 125 μl of 6 M guanidine hydrochloride was added to a final concentration of 5 M and the sample was incubated at room temperature for further 30 min. 1 ml of Trizol™ was added and RNA was extracted using standard protocol by the manufacturer using the RNA Clean & Concentrator-5 kit (Zymo Research, USA). Eluted RNA was quantified using Qubit RNA BR dyes (Thermo Fisher Scientific, USA). DNA was extracted from the insoluble protein pellet with the QIAamp DNA Mini Kit (Qiagen, Germany) according to the manufacturer protocol for DNA extraction from blood. The final elution was performed in 50 μl of buffer AE. Extracted nucleic acids were again quantified using the Qubit fluorescent dyes.

Results

Our work stems from the consideration that the structural core of amyloid fibrils can be effectively predicted by computational approaches prioritizing their hydrophobic core. This characteristic hints at the existence of a property gradient within the fibril structure, wherein the regions peripheral to the amyloid core predominantly exhibit heightened hydrophilicity. We theorize that the solvent-exposed hydrophilic amino acids can be endowed with favorable electrostatic attributes, enabling a higher propensity for nucleic acid binding. This biochemical interaction can have profound implications for intracellular aggregate formation. Indeed, nucleic acids may modulate the assembly kinetics of proteins, thereby affecting aggregate formation.

To validate our hypothesis, we used our computational analysis to design experiments on an amyloid forming protein, alpha-synuclein. We studied its aggregation in the presence of RNA and explored alterations for a pathologically relevant mutant.

Calculations recapitulate experimental results on the ability of tau fibrils to attract RNA

We started by assessing the robustness of our calculations of aggregation propensities on the known case of tau. A cryo-EM structure of tau fibrils (PDB 7SP1), grown in the presence of total RNA, shows an electron density in a pocket on the external region of the aggregate where an RNA molecule was later additionally docked into (13). The 36-amino acid region of the C-terminal disordered domain (residues 391–426) is identified in the cryo-EM study as the fibril core (13) and another work indicates that the third microtubule-binding repeat (residues 306–335) (58,59) contributes to tau aggregation. Using the Zyggregator algorithm (Materials and Methods), which predicts aggregation-prone elements in protein sequences (40), we are able to successfully identify these two regions as the ones with highest aggregation propensity (peaks at residues 320 and 409, Figure 1A). This analysis indicates that Zyggregator is able to identify the inner regions of the tau aggregate in detail.

At the experimental level, the binding of docked RNA is stabilized through hydrogen bonds between the phosphate backbone, the 2-OH group of the ribose, and additional contacts between the RNA bases and the aggregate core (13). More in detail, the cryo-EM structure of aggregated tau reveals that residues R406 and H407 on the tau fibril surface directly interact with sequestered RNA and that residues E391-S396, V398, S400, D402, S404, D421, P423, I425, A426 are exposed to the solvent (Figure 1B). An examination of residues 391–426 reveals that the calculated physico-chemical properties are consistent with cryo-EM results. This confirms that the experimentally identified surface regions exhibit a higher propensity for RNA-binding. Additionally, these regions are predicted to possess less hydrophobicity and more disorder, as detailed in Figure 1B and the Materials and Methods section.

We broadened our analysis to encompass the entirety of the tau sequence, thereby incorporating regions that lie outside of the core aggregate at residues 391–426 (Figure 1C). Upon examining the regions associated with low Zyggregator scores, it is notable that peaks within residues 175–220 and 230–260, which correspond to the proline-rich region and first microtubule-binding repeat, have been highlighted in previous research for their direct involvement in interactions with nucleic acids (60). This finding suggests a high likelihood of their localization within the external regions of the amyloid structure (Figure 1C). We observe that these regions, characterized by inherent disorder, may evade detection through traditional experimental techniques such as cryo-EM. Consequently, their potential contribution to the sequestration of RNA during aggregation cannot be dismissed. This possibility has also been acknowledged by the authors of the cryo-EM study (13).

This initial result indicates that regions outside the core of tau amyloid fibrils have the potential to interact with RNA

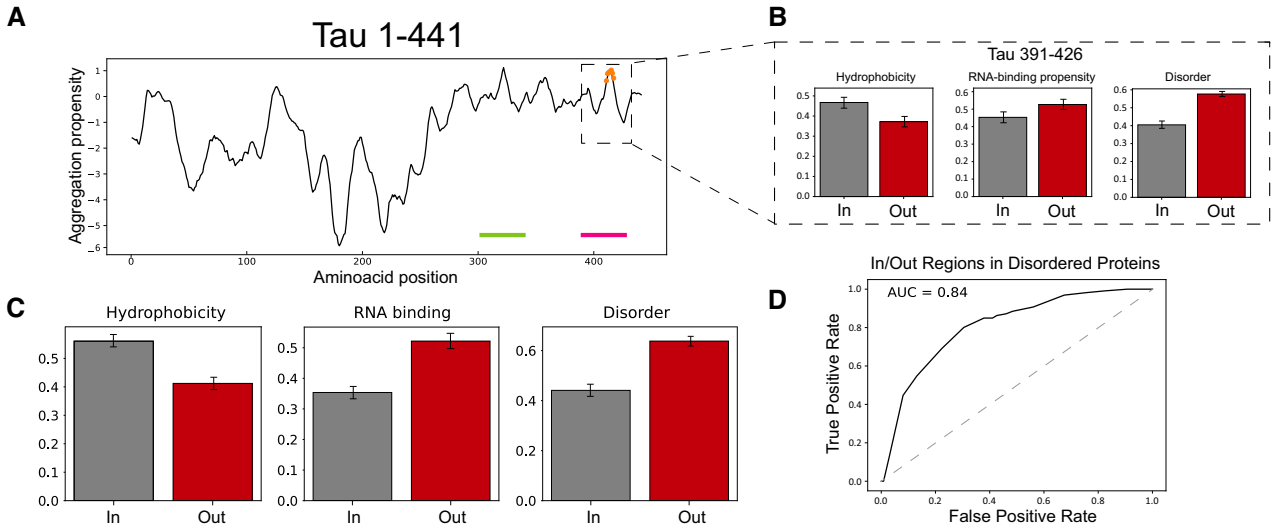


Figure 1. Analysis of tau aggregation and prediction of physico-chemical properties. (A) Aggregation propensity profile for tau, with orange dots representing the predicted internal part of the fibril, and green and pink bars indicating regions experimentally identified as part of the fibril core. B) Comparison of hydrophobicity, RNA-binding, and disorder propensities for residues 391–426, with features of experimentally identified internal (In) and external (Out) regions marked in black and red, respectively (mean and standard deviation shown for these regions). (C) Analysis of hydrophobicity, RNA-binding propensity, and disorder across the entire tau sequence, using the Zyggregator method to distinguish between internal and external regions (with a threshold of 0.8; error bars are calculated by varying the Zyggregator threshold within the 0.76–0.84 range). (D) Performance evaluation of the Zyggregator method in differentiating between the internal and external regions of experimentally identified protein aggregates, with an optimal threshold of 0.8, as indicated by the Youden index of the Receiver Operating Characteristic (ROC) curve. With regard to panel C, we note that in the range 0.76–0.84 the Zyggregator scores are associated with a True Positive Rate (TPR) of 0.75 and a False Positive Rate (FPR) of 0.25.

and we are able to identify them in great detail using our computational approaches. This finding also suggests that other aggregating proteins might behave as tau.

Internal and external regions of amyloids have different biophysical properties

In an effort to explore the prevailing trend that regions capable of binding RNA are located on the exterior of amyloids, we sourced a dataset of amyloids from a previous publication (61). We then analyzed their physico-chemical properties employing the Zyggregator algorithm (Supplementary Table S1). This allowed us to pinpoint the regions constituting the amyloid core, as outlined in the Materials and Methods section. We focused on proteins that partially exhibit an unfolded state in their monomeric form. We especially prioritized proteins that not only can be found in their amyloid form, but that have been also empirically determined and stringently validated. Our focus has been on the following proteins: calcitonin (41,42), amylin (43), glucagon (44), Aβ42 (45,46), alpha synuclein (47–49), hnRNPDL-2 (12) and tau (13). The model achieves an Area Under the Curve (AUC) of 0.84 on the Receiver Operating Curve (ROC), thereby effectively identifying the amyloid core of the aggregates. The optimal Zyggregator threshold is noted to be 0.8, as indicated by the Youden Index of the ROC curve, which corresponds to a True Positive Rate (TPR) of 0.75 and a False Positive Rate (FPR) of 0.25 (Figure 1D).

As expected, our calculations show that the inner region of an amyloid is more hydrophobic than the outer one (Figure 2A). There is a significant shift of the RNA-binding propensity and structural disorder in favor of the external region (Figure 2B and 2C) while the electrostatic charge distribution does not differ substantially (Supplementary Figure S1A and B). Notably, the trend is observed with different physico-chemical propensity predictors (Supplementary Figure S2). Given the overall negative charge of the phosphate backbone of RNA and no significant difference in the aggregate charge distribution, we find strong RNA-binding propensity exclusively for positively charged protein regions.

To better visualize this trend we chose the structure of the amyloid formed by *Podospora anserina* protein HET-s (PDB 2RNM) as an example (Figure 2D-F). The structure was determined using solid-state NMR and included the flexible outer region of the amyloid. Using different color scales, the distribution trends are clearly visible, especially in the hydrophobic amyloid core (Figure D). The RNA-binding propensity is considerably higher for the unstructured external portion of the fibers, similarly to the structural disorder (Figure 2D-F, Supplementary Figure S2B). Overall, the analysis of the HET-s amyloid structure shows excellent agreement with the predicted results from the computational analysis.

Within the set of amyloids (Figure 2A–C), we analyzed proteins that have RNA and DNA binding domains annotated in Uniprot (Supplementary Tables S2 and S3). Consistent with our calculations of RNA-binding abilities, the nucleic acid-binding domains are predicted to be displayed on the surface of the aggregate (Figure 2G). This is well exemplified by hnRNPDL-2, a protein in which residues 224–280 were found by cryoEM to constitute the core of the fibril (PDB 7ZIR), while two RNA-binding domains are external (12). The aggregation propensity predicted by Zyggregator captures the inner part of the aggregate very well, and the RNA-binding domains are indeed predicted to be more exposed (Figure 2H). Further investigating this phenomenon at the human proteome

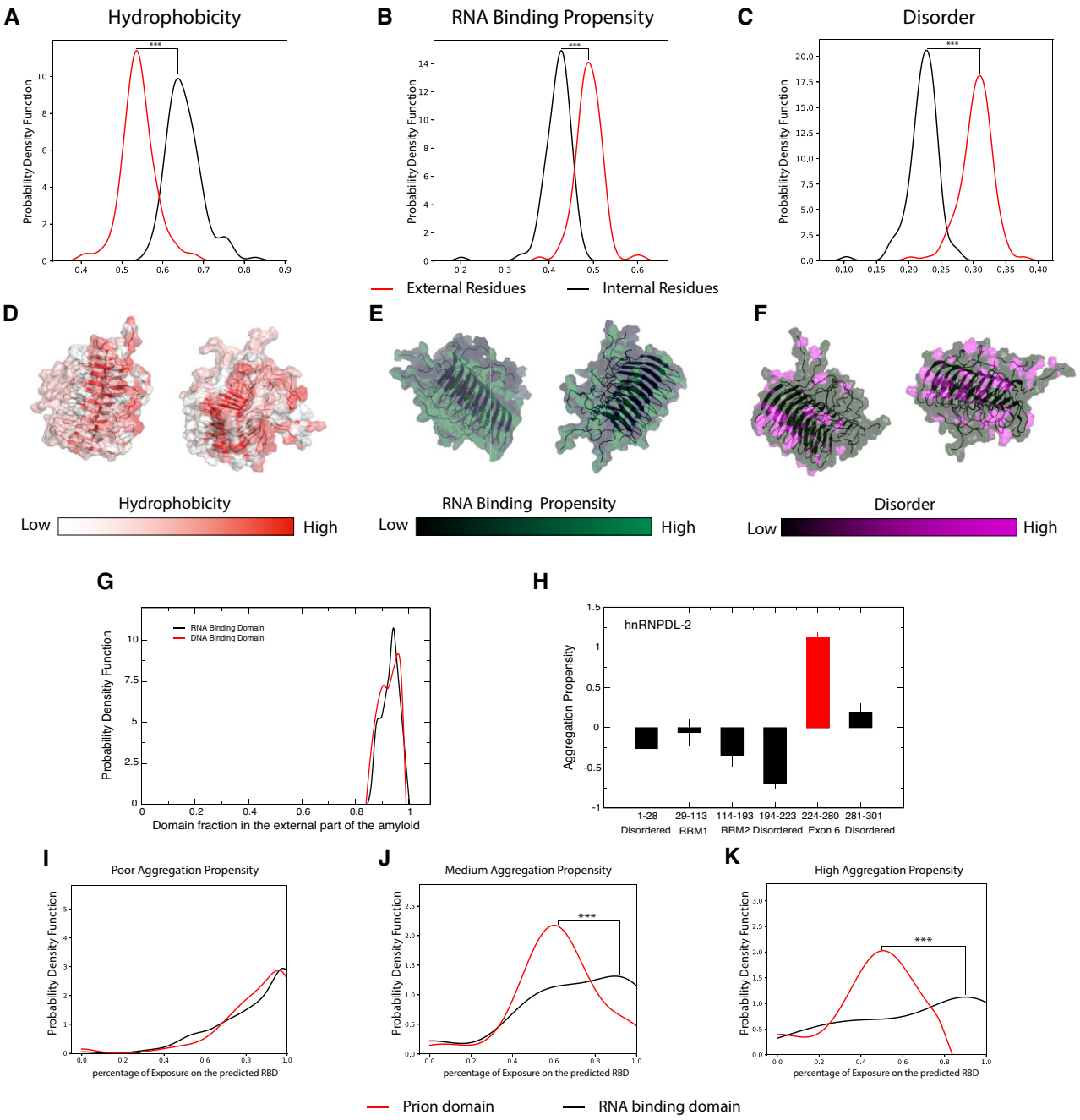


Figure 2. Physico-chemical properties of internal and external regions of amyloids. (A–C) Differences in hydrophobicity, RNA-binding, and disorder propensities were calculated for the internal and external portions of amyloids, which were identified using the Zyggregator method (Supplementary Table S1; Supplementary Figure S2). (D–F) Molecular structure of the HET-s prion protein (PDB 2RNM) solved by solid-state NMR, depicted by different color scales for hydrophobicity, RNA-binding propensity, and disorder, respectively. (G) Fraction of the RNA and DNA binding domains predicted to belong to the outer region of amyloid fibrils. (H) Zyggregator scores for hnRNPDL-2 domains, showing the average values along with the corresponding standard error of the mean. The red color highlights exon 6, which has been experimentally demonstrated to form the core of the aggregate. (I–K) Fraction of exposed residues in the RNA-binding and prion domains of the human proteome at different Zyggregator scores. *** P-values < 0.0001 (Kolmogorov–Smirnov test).

level, we carried out predictions of prion and RNA-binding domains, respectively using PLAAC (50) and *cat*RAPID signature (62). Proportionally to the overall aggregation propensity computed with Zyggregator, we estimated that prion domains are in the core of the fibril while RNA-binding domains tend to be on the outer regions (Figure 2I-K, Supplementary Table S4). The literature is in agreement with our analysis of hnRNPDL-2 (12).

Physico-chemical properties of alpha synuclein aggregates

We next analyzed the aggregation propensity of alpha synuclein (aS or aS140), a protein with a well-characterized amyloid core (40). We computed the aggregation propensity of the wild type sequence (Figure 3A) and performed a mutational analysis with 10 000 random single point variations (63). As supported by data reported in the literature (64), the

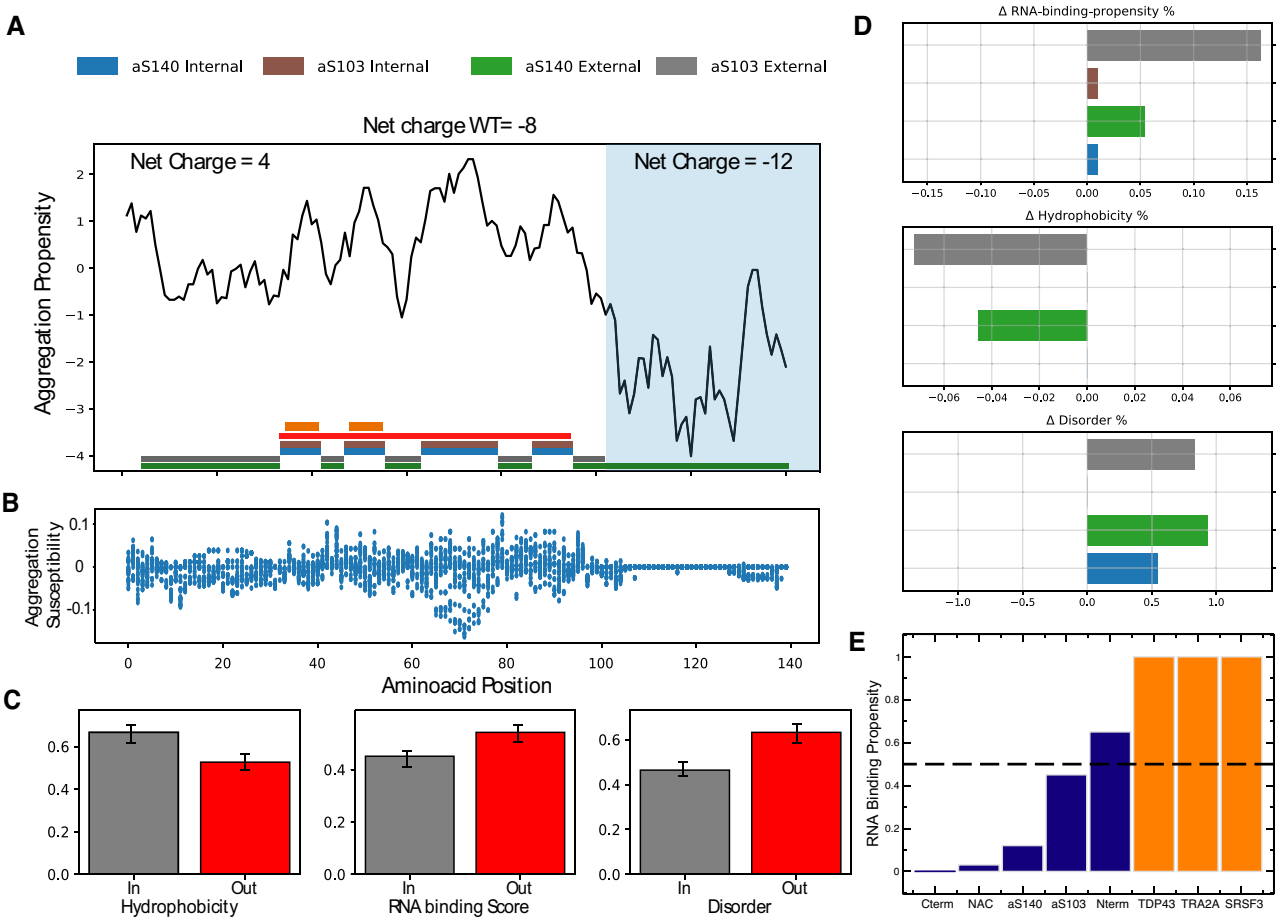


Figure 3. Aggregation propensity of aS140 and aS103. (A) Zyggregator profile of aS140 and aS103, with colored bars indicating predicted inner and outer amyloid regions. The experimentally determined amyloid core of aS is marked with a red bar. Orange indicates minor Zyggregator peaks playing a role in modulating aS aggregation. (B) Random mutagenesis analysis showing that the NAC region and N-terminus of aS are susceptible to significant changes in aggregation propensity. (C) Physico-chemical properties of inner (in, gray) and outer (out, red) regions of aS140 and aS103. (D) RNA-binding propensity, hydrophobicity and disorder of aS relative to its internal and external parts. Differences with respect to the mean of the amyloid histograms (Figure 2A–C) are shown. (E) RNA-binding predictions of different regions of aS140 and aS103, compared with canonical RBP (scores > 0.5 indicate non-negligible interaction propensities).

central region of the protein emerges as a key determinant in protein aggregation, exhibiting a prominent Zyggregator peak that aligns with the NAC region (residues 61–95; Figure 3A). Furthermore, the work by Doherty and colleagues (24) sheds light on additional smaller regions, specifically residues 36–42 and 45–57, which coincide with minor Zyggregator peaks and play a crucial role in modulating aS aggregation (Figure 3A).

Mutations in the NAC region display a strong impact on aS140 aggregation (Figure 3B) and the outer regions of aS140 aggregates are predicted to be disordered, to display low overall hydrophobicity, and to be able to interact with RNA molecules (Figure 3C). Amino acids 67–82 appear to be particularly sensitive to change and dramatically affect the propensity of aS to aggregate, in agreement with previous experimental work (65). Random single mutations in the C-terminus have instead negligible effects. Yet, 14 out of the 24 negatively charged residues in aS140 (18 Glu and 6 Asp) are found in this region, indicating that the truncated form of aS140 -namely aS103-, lacking all amino acids in positions 104–140, could behave differently compared to the wild type protein in terms of aggregation propensity as well as RNA-binding ability.

Computing the scores of the physico-chemical properties for both aS140 and aS103 reveals a strong difference between the two protein variants (Figure 3D). The removal of the 12 negatively charged residues in the C-terminal portion results in a net charge increment. Accordingly, the external regions of aS103 and aS140 show a lower degree of hydrophobicity compared to the average of the dataset and an increase in RNA-binding propensity (Figure 3D) due to the abundance of charged and polar residues outside the NAC region. We note that removing the C-terminus slightly affects the predicted structural disorder, which is consistent with the overall unfolded nature of the whole aS (Figure 3D).

We next used the *cat*RAPID approach (55,57) to estimate the ability of aS140 and aS103 to interact with RNA (Materials and Methods). While the C-terminus and NAC show low interaction propensity, the N-terminal domain is predicted to be exposed to the solvent in the aggregate and available to interact with RNA (Figure 3E). Based on this observation, we hypothesize that aS might be able to sequester

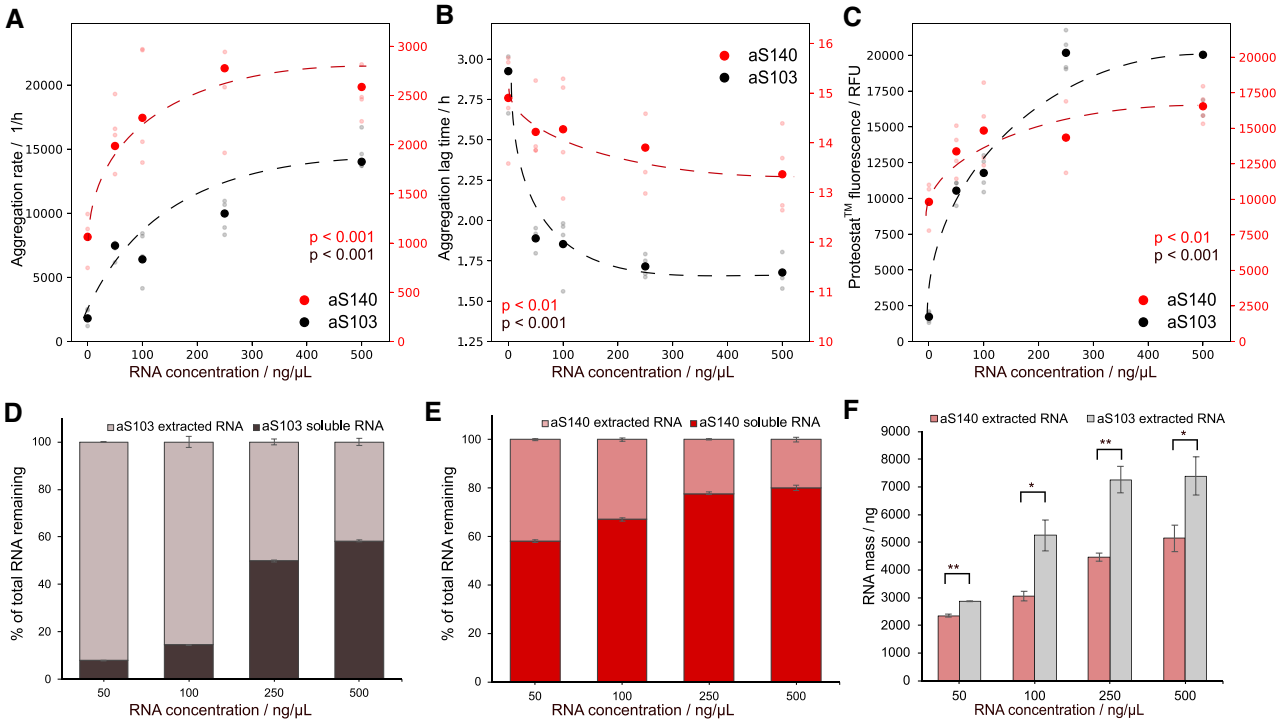


Figure 4. In vitro aggregation assays of aS140 and aS103 in the presence of RNA. A-C) Data points corresponding to each protein are plotted against their own respective y-axis (aS103 to the left and aS140 to the right axis). A) Calculated aggregation rates from aggregation data (Supplementary Fig. 3, Supplementary Fig. 4B) of time resolved Proteostat™ fluorescence. Data shown are calculated from parameters derived from the fitted curve to 6 technical replicates of 3 separate experiments and statistical significance was calculated for rates in the presence of RNA as compared to the protein-only samples (see Materials and Methods). The presence of RNA significantly increases the rate of aggregation for both aS103 (Fisher t-test, $p < 0.001$) and aS140 (Fisher t-test, $p < 0.001$), with the aS140 rate reaching a plateau with concentrations higher than 250 ng/µL. B) aS103 has a significantly shorter aggregation lag time than aS140. The addition of RNA further decreases the lag phase of both proteins (aS103, $p < 0.001$ and aS140, $p < 0.01$, Fisher t-test), however increasing concentrations do not seem to drastically affect aS103 further, while the opposite is true for aS140. C) Increasing concentration of RNA also elevates the plateau of fluorescence for both proteins, especially significantly for aS103 ($p < 0.001$ compared to p 0.01 for aS140, Fisher t-test). D and E) Quantification of nucleic acids from aS aggregation assays indicates different behavior for aS140 and aS103. Comparing RNA in solution with the fraction extracted from aggregates of aS103 and aS140, we found that RNA is preferentially sequestered by aS103 aggregates. F) The average amount of RNA extracted from aS103 and aS140 aggregates at different initial starting concentrations show again that aS103 aggregates have a higher propensity of sequestering soluble RNA upon aggregation (Student t-test, * $p < 0.05$, ** $< p 0.01$).

RNA upon aggregation. As we will show in the following sections, this effect could shape the dynamic of the aggregation propensity, tuning the concentration of the RNA in solution.

RNA affects the aggregation of alpha synuclein isoforms in different ways

To study experimentally the effect that RNA might have on the aggregation kinetics of aS140 and aS103, we performed aggregation assays in the presence of a fluorescent aggregate intercalator. Both aS103 and aS140 were incubated with increasing concentrations of total RNA, from 0 to 500 ng/µl, for 24 h (Figure 4A, Supplementary Figure S3A-B, Supplementary Figure S4). Since the experiments were performed in the presence of a reporter dye, visualization of the aggregates at the end of the experiments with a fluorescence microscope confirmed the presence of protein aggregates in all samples and in every tested condition (Supplementary Figure S5A).

As previously reported in literature, aS103 aggregates significantly faster than aS140, irrespective of the concentration of RNA present in solution (Figure 4A and 4B, Supplementary Figures S3 and S4) (66). In the presence of RNA, the elongation phase of aS103 is significantly faster than that of aS140 (Figure 4A) and is accompanied by a consistently shorter lag phase (Figure 4B). Our data analysis shows that the estimated aggregation rate of aS103 is approximately two-fold higher in the absence of RNA and up to 10-fold higher in the presence of 500 ng/µl RNA. Importantly, increasing concentrations of RNA lead to varied effects on the aggregation rate of both proteins. For aS140, the aggregation rate in the presence of 250 ng/µl RNA increases, up to ca. 5-fold. With higher RNA concentrations, the rate seems to reach a plateau and does not increase further (Figure 4A). By contrast, incremental additions of RNA increase the aggregation rate of aS103 over ten-fold, reaching a maximum at 500 ng/µl RNA, with no evidence of reaching a plateau (Figure 4A, Supplementary Figure S4). RNA affects the lag time in the opposite way, with increasing concentrations effectively decreasing the lag time of aS103 until reaching a stable plateau at 1.6 h. The lag time of aS140 however only starts drastically decreasing at RNA concentrations higher than 250 ng/µl (Figure 4B). The actual maximum value of fluorescence at the end of the exponential phase does not differ much between aS103 and aS140 (Figure 4C). RNA incrementally increases the level of maximum fluorescence up to 250 ng/µl, when the value stabilizes. These results suggest not only that RNA has a profound effect on aS aggregation, but that it also affects aS140 and aS103 in different ways as hinted by our computational analysis.

Having observed the effect RNA has on the aggregation kinetics of aS, we then set to test the original hypothesis that

aS can sequester the RNA upon aggregation. First we verified that the monomeric forms of both proteins do not interact with RNA (Supplementary Figure S5B). We then quantified the RNA found in both soluble and insoluble fractions at the of the aggregation assays, as described in Materials and Methods. The concentration was measured individually for each well and normalized according to the initial RNA concentration in the same well, in order to account for individual rates of RNA degradation. Also, in order to avoid any interference by remaining non-aggregated protein and smaller aggregates, we measured the RNA concentration of using an RNA-specific fluorescent dye. Importantly, to bypass any potential differences between the volumes of the soluble and insoluble fractions, concentrations were recalculated as RNA mass and the data are here presented as the percentage of total RNA mass after aggregation in different fractions (Figure 4D and 4E). These percentages were individually adjusted to 100% to again take into account the different rates of RNA degradation during experiment. The results clearly indicate that both aS140 and aS103 are able to sequester RNA through aggregation. Furthermore, there is a clear difference between the amount of RNA sequestered by individual protein constructs. As predicted by the computational analysis, aS103 aggregates can sequester RNA to a much higher degree than aS140 (Figure 4F). The difference is less noticeable at lower initial RNA concentrations, but becomes more pronounced at higher concentrations. As the starting concentration of the protein is always 50 μM, it is safe to assume the difference is due to the higher sequestration capacity of aS103 aggregates. This is further supported by the fact that less than 10% of the starting amount of RNA is found soluble at the concentration of 50 ng/μl (Figure 4D). This sequestration capacity seems to be saturated at 250 ng/μl for aS103 while it keeps increasing for aS140 (Figure 4F).

To verify whether the different distribution of RNA among the soluble and insoluble fractions is in fact due to its sequestration by protein aggregates, we repeated the experiments in the same conditions with bovine serum albumin, a non-aggregating, non-RNA-binding protein. The quantification of RNA at the end of the assay showed comparable values to the control values of total RNA without any protein added (Supplementary Figure S6A). The wells in both cases were inspected by fluorescence microscopy and no aggregates were detected (Supplementary Figure S5A).

We repeated the aggregation of aS103 and aS140 in the presence of total cell DNA extract, to determine whether what we observed is an RNA-specific effect. In order to take into account the different rates of degradation of DNA and RNA (Supplementary Figure S6B), we again measured the individual well concentrations both before and after aggregation assay with a DNA-specific fluorescent dye. aS103 still shows the ability to sequester DNA at lower initial DNA concentrations, however higher concentrations apparently limit the binding ability (Supplementary Figure S6C). aS140 barely sequesters any DNA in the aggregates, with less than 10% of the initial DNA extracted from the insoluble part (Supplementary Figure S6D).

These results show that both aS140 and aS103 have the ability to sequester nucleic acids upon aggregation. The grade of incorporation is protein sequence-dependent and confirms the computational analysis predicting aS103 aggregates as the higher-propensity binders. The results further emphasize that protein aggregation is a dynamic process, and that aggregated species not only display different biophysical features but could also acquire new functional properties distinct from the ones attributed to the monomers.

Discussion

Aggregates are able to sequester molecules, including other proteins (1) and several metabolites (67). In this study, we explored the possibility that aggregated protein structures, such as amyloids, might also serve as traps for nucleic acids. Our hypothesis was formulated considering that aggregation-prone regions of proteins are mostly hydrophobic and buried in the core of fibrils (40), while external regions, prevalently hydrophilic and often structurally disordered, can interact with nucleic acids, if their electrostatics is favorable (68). Several lines of evidence, both computational and experimental, confirmed our original idea. First, regions outside the core of amyloids are predicted to interact with RNA and annotated RNA-binding domains have a strong propensity to be solvent-exposed in amyloids. Most importantly, the recently reported cryoEM structures of tau and hnRNPDL-2 show that RNA is indeed found on the surface of the aggregates (13) and the RNA-binding domains are outside the amyloid core (12).

Intriguingly, a number of artificial nucleic acid molecules, RNA and DNA aptamers, have been designed to impact protein aggregation (4,69). For instance, an aptamer called T-SO517, developed by Ikebukuro *et al.* against oligomeric αS140, bound not only to the designated target but also to Aβ40 oligomers (70). This finding confirms our hypothesis that aggregates can bind to nucleic acids, although it suggests that the binding is not specific. Thus, it appears plausible that the non-specific sequestration of RNA within a cellular environment might add to the toxicity of amyloid aggregates. This idea aligns with a previously posited hypothesis, suggesting that toxic effects could occur with proteins that co-aggregate together (1).

Here, we hypothesized that the formation of the amyloid core in aS fibrils, specifically involving the most hydrophobic residues (namely, the NAC region), leaves charged amino acids (such as the N-terminal and C-terminal) accessible for interactions with other molecules. This concept is supported by various experimental methodologies, including site-directed spin labeling coupled with EPR, hydrogen/deuterium exchange, and limited proteolysis (34,47–49). These methods indicate that the central portion of the protein is engaged in forming the core of the aggregates, leaving the remaining protein regions exposed to the solvent. We posited that the elimination of the negative charges corresponding to amino acids 104–140, which are not part of the core aggregate, would significantly augment the RNA-binding propensity of the regions outside the NAC. Consequently, we examined the hypothesis concerning the RNA-binding gain-of-function upon aggregation. The decision to remove the aS C-terminus was made in light of the unfavorable electrostatics of this disordered region concerning RNA molecules. In previous studies it has been shown that favorable electrostatics can promote RNA-binding by enabling the protein to adopt multiple conformations and engage with various partners (71,72). This structural disorder in RNA-binding proteins allows them to interact with RNA in a more flexible and promiscuous manner (73). In certain cases, disordered proteins may even undergo a disorder-to-order conformational transition upon interacting with RNA. Such complex behavior can be observed in mesh-like condensates such as stress granules, where RNAs are capable of entrapping proteins (73,74). These interactions

within stress granules can provoke conformational alterations in RNA-binding proteins, culminating in the formation of large assemblies (75,76).

Our experimental results confirm the observation of acquired RNA sequestration ability and identify it as a way of modulating aS aggregation process by directly affecting the aggregation rates. RNA sequestration is protein-sequence dependent: aS103 shows a higher sequestration propensity compared to wild type aS140, likely due to the increased electrostatic attraction. RNA also modifies the aggregation rate of both aS variants.

As previously reported, the key to modulating aS aggregation propensity lies in the long-range intra- and intermolecular interactions between the N-terminal and C-terminal domains of the protein, which shield the NAC region and prevent its self-assembly (34,35). RNA likely enters in contact with the N-terminal, thus interfering with these interactions. The truncated aS103 lacks the C-terminal domain and *a priori* cannot engage in the shielding effect of the NAC region. It is plausible that, through transient interactions via the N-terminal domain, RNA acts as a bridge between protein molecules, bringing the NAC regions into closer contacts. As C-terminal truncations of aS are commonly found in brain samples from patients and aS103 is prevalent in the medial temporal lobe in Lewy body dementia, we speculate that RNA might play a role in the onset of the disease (28).

Very recently, the combined solid-state NMR-cryoEM approach to study the fibrils of aS by Zhang and colleagues has confirmed our observations regarding the outer disorder of aS fibrils (77). The structure (PDB 7YK2) also highlighted the previously unreported additional interaction between the core of the fibril and the G14-G25 region in the N-terminal part. This suggests that additional ordering of the structure might be present in the N-terminus upon aggregation. It is important to note, however, that the fibrils in the study were grown in the absence of RNA, which could potentially prevent these additional intramolecular interactions to occur upon misfolding by shielding the N-terminus. An additional case hinting at the potential of RNA to modify the structure of aS aggregates comes from the work of Schweighauser and colleagues (78). Their cryo-EM study of three different strains of fibrils found in the post-mortem brain of multiple system atrophy patients show a unique fibril fold with an unassigned electron density in the central cavity of the amyloid core. This density is enclosed in a pocket between the interface of two fibrils and formed by positively charged residues K43, K45 and H50 from both filaments. While there is no direct evidence that the electron density corresponds to trapped RNA, and the fold is unique amongst the aS amyloid architectures found in Amyloid Atlas (79) to contain such a central cavity, the possibility remains that RNA could be the culprit in inducing the formation of such an architecture.

We note that, for our experiments, we used near-to-physiological RNA concentrations. Indeed, we considered a mammalian cell has a cell volume between 1000 and 10 000 μm^3 [i.e. 10^{-6}–10^{-5} µl (80)] and an RNA amount of 10 - 30 pg [i.e. 0.01–0.03 ng (81)], which indicates that 500 ng/µl of RNA are in the expected normal range (i.e. 0.01 ng in 10^{-5} µl, which is 1000 ng/µl). Monomeric, physiological aS was found associated with nucleic acids (20,82), nucleosomes (83) and processing bodies (84). However, no proof of direct and functional RNA binding by aggregated aS has ever been discovered. While Siegert *et al.* have shown RNA has little effect and can actually abrogate aS liquid-liquid phase separation (85), the molecular aspects of the effect of RNA on aS liquid-to-solid phase transition have not yet been characterized. It should be mentioned that Cohlberg *et al.* have previously reported that negatively charged molecules such as heparin display a stronger effect on aS aggregation compared to 1 µg/µl RNA (38). Nevertheless, it must be noted that the authors used thioflavin T to quantify aS aggregation. This dye can interact also with RNA (86), interfering with the correct characterization of the kinetics of protein aggregation. For this reason, we selected a different aggregation reporter, Proteostat™, that does not bind to RNA (87), obtaining significantly different results.

An intriguing and significant aspect of our study is the potential pathological implications of RNA sequestration within the context of synucleinopathies. Recent research has highlighted the formation of nuclear alpha-synuclein inclusions via interneuronal transmission, with intranuclear aggregates potentially playing a role in driving the pathogenesis of synucleinopathies (88). An independent investigation has proposed that alpha-synuclein could facilitate the release of nuclear factors that further enhance its own aggregation (89). Yet, it has been observed that constitutive nuclear accumulation of endogenous alpha-synuclein in mice leads to motor impairment and cortical dysfunction, even in the absence of significant protein aggregation (90). Collectively, these findings highlight the potential of alpha-synuclein to aggregate within the nucleus and suggest that this phenomenon could substantially contribute to the progression of synucleinopathies. This understanding sheds light on the intricate interplay between RNA sequestration, nuclear aggregation of alpha-synuclein, and the development of these debilitating disorders.

At a fundamental biological level, our investigation presents an important revelation: the aggregation process, coupled with the concomitant misfolding, can induce distinct biophysical and functional attributes that markedly differ from those exhibited by the monomeric, soluble protein counterparts. This phenomenon holds significant implications beyond the scope of our study, extending its relevance to a broader array of proteins. In particular, our findings spotlight the case of aS, a protein traditionally deemed non-RNA-binding. Intriguingly, aS can, under specific protein sequence conditions, acquire the ability to sequester RNA. This discovery adds a new layer of significance to the intricate interplay between structural dynamics and functional shifts across diverse protein states. This observation resonates beyond aS, hinting at a broader paradigm: the dynamic transformations that aggregation can impose on a protein's intrinsic behavior and interactions. Such insights might illuminate new avenues of exploration across various proteins, where aggregation-associated changes could manifest in unanticipated and consequential ways. By revealing these dynamic shifts, our study contributes to a more comprehensive comprehension of protein behavior, emphasizing the intricate interplay between aggregation, misfolding, and functional diversity across the biological landscape.

Data availability

The authors confirm that the data supporting the findings of this study are available within the article and its supplementary materials.

Supplementary data

Supplementary Data are available at NAR Online.

Acknowledgements

The authors would like to thank the 'RNA initiative' at IIT and all the members of Tartaglia's lab at CRG, Sapienza and IIT.

Funding

H2020 European Research Council [855923]. Funding for open access charge: ERC ASTRA_855923 (to G.G.T.); PNRR grant from National Center for Gene Therapy and Drugs based on RNA Technology (CN00000041, EPNRRCN3 to GGT) and IVBM4PAP_101098989 (to G.G.T.); E.Z. received funding from the MINDED fellowship of the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie [754490].

Conflict of interest statement

None declared.

References

1. Olzscha,H., Schermann,S.M., Woerner,A.C., Pinkert,S., Hecht,M.H., Tartaglia,G.G., Vendruscolo,M., Hayer-Hartl,M., Hartl,F.U. and Vabulas,R.M. (2011) Amyloid-like aggregates sequester numerous metastable proteins with essential cellular functions. *Cell*, **144**, 67–78.
2. Gallardo,R., Ranson,N.A. and Radford,S.E. (2020) Amyloid structures: much more than just a cross-β fold. *Fold. Bind. Proteins*, **60**, 7–16.
3. Maharana,S., Wang,J., Papadopoulos,D.K., Richter,D., Pozniakovsky,A., Poser,I., Bickle,M., Rizk,S., Guillén-Boixet,J., Franzmann,T.M., *et al.* (2018) RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science*, **360**, 918–921.
4. Zacco,E., Kantelberg,O., Milanetti,E., Armaos,A., Panei,F.P., Gregory,J., Jeacock,K., Clarke,D.J., Chandran,S., Ruocco,G., *et al.* (2022) Probing TDP-43 condensation using an in silico designed aptamer. *Nat. Commun.*, **13**, 3306.
5. Shmookler Reis,R.J., Atluri,R., Balasubramaniam,M., Johnson,J., Ganne,A. and Ayyadevara,S. (2021) "Protein aggregates" contain RNA and DNA, entrapped by misfolded proteins but largely rescued by slowing translational elongation. *Aging Cell*, **20**, e13326.
6. Lester,E., Ooi,F.K., Bakkar,N., Ayers,J., Woerman,A.L., Wheeler,J., Bowser,R., Carlson,G.A., Prusiner,S.B. and Parker,R. (2021) Tau aggregates are RNA-protein assemblies that mislocalize multiple nuclear speckle components. *Neuron*, **109**, 1675–1691.
7. Tyedmers,J., Mogk,A. and Bukau,B. (2010) Cellular strategies for controlling protein aggregation. *Nat. Rev. Mol. Cell Biol.*, **11**, 777–788.
8. Hallegger,M., Chakrabarti,A.M., Lee,F.C.Y., Lee,B.L., Amalietti,A.G., Odeh,H.M., Copley,K.E., Rubien,J.D., Portz,B., Kuret,K., *et al.* (2021) TDP-43 condensation properties specify its RNA-binding and regulatory repertoire. *Cell*, **184**, 4680–4696.
9. Gotor,N.L., Armaos,A., Calloni,G., Torrent Burgas,M., Vabulas,R.M., De Groot,N.S. and Tartaglia,G.G. (2020) RNA-binding and prion domains: the Yin and Yang of phase separation. *Nucleic Acids Res.*, **48**, 9491–9504.
10. Li,Q., Babinchak,W.M. and Surewicz,W.K. (2021) Cryo-EM structure of amyloid fibrils formed by the entire low complexity domain of TDP-43. *Nat. Commun.*, **12**, 1620.
11. Lee,M., Ghosh,U., Thurber,K.R., Kato,M. and Tycko,R. (2020) Molecular structure and interactions within amyloid-like fibrils formed by a low-complexity protein sequence from FUS. *Nat. Commun.*, **11**, 5735.
12. Garcia-Pardo,J., Bartolomé-Nafría,A., Chaves-Sanjuan,A., Gil-Garcia,M., Visentin,C., Bolognesi,M., Ricagno,S. and Ventura,S. (2023) Cryo-EM structure of hnRNPDL-2 fibrils, a functional amyloid associated with limb-girdle muscular dystrophy D3. *Nat. Commun.*, **14**, 239.
13. Abskharon,R., Sawaya,M.R., Boyer,D.R., Cao,Q., Nguyen,B.A., Cascio,D. and Eisenberg,D.S. (2022) Cryo-EM structure of RNA-induced tau fibrils reveals a small C-terminal core that may nucleate fibril formation. *Proc. Natl. Acad. Sci. U.S.A.*, **119**, e2119952119.
14. Chakraborty,P., Rivière,G., Liu,S., de Opakua,A.I., Dervişoğlu,R., Hebestreit,A., Andreas,L.B., Vorberg,I.M. and Zweckstetter,M. (2021) Co-factor-free aggregation of tau into seeding-competent RNA-sequestering amyloid fibrils. *Nat. Commun.*, **12**, 4231.
15. Zhang,X., Lin,Y., Eschmann,N.A., Zhou,H., Rauch,J.N., Hernandez,I., Guzman,E., Kosik,K.S. and Han,S. (2017) RNA stores tau reversibly in complex coacervates. *PLoS Biol.*, **15**, e2002183.
16. Dinkel,P.D., Holden,M.R., Matin,N. and Margittai,M. (2015) RNA binds to tau fibrils and sustains template-assisted growth. *Biochemistry*, **54**, 4731–4740.
17. Zwierzchowski-Zarate,A.N., Mendoza-Oliva,A., Kashmer,O.M., Collazo-Lopez,J.E., White,C.L. and Diamond,M.I. (2022) RNA induces unique tau strains and stabilizes Alzheimer's disease seeds. *J. Biol. Chem.*, **298**, 102132.
18. Lashuel,H.A., Overk,C.R., Oueslati,A. and Masliah,E. (2013) The many faces of α-synuclein: from structure and toxicity to therapeutic target. *Nat. Rev. Neurosci.*, **14**, 38–48.
19. Nguyen,M., Wong,Y.C., Ysselstein,D., Severino,A. and Krainc,D. (2019) Synaptic, mitochondrial, and lysosomal dysfunction in Parkinson's disease. *Trends Neurosci.*, **42**, 140–149.
20. Schaser,A.J., Osterberg,V.R., Dent,S.E., Stackhouse,T.L., Wakeham,C.M., Boutros,S.W., Weston,L.J., Owen,N., Weissman,T.A., Luna,E., *et al.* (2019) Alpha-synuclein is a DNA binding protein that modulates DNA repair with implications for Lewy body disorders. *Sci. Rep.*, **9**, 10919.
21. Spillantini,M.G., Crowther,R.A., Jakes,R., Hasegawa,M. and Goedert,M. (1998) α-Synuclein in filamentous inclusions of Lewy bodies from Parkinson's disease and dementia with Lewy bodies. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 6469–6473.
22. Chiti,F. and Dobson,C.M. (2017) Protein misfolding, amyloid formation, and human disease: a summary of progress over the last decade. *Annu. Rev. Biochem.*, **86**, 27–68.
23. ,Grazia Spillantini, Anthony Crowther,R., Jakes,R., Cairns,N.J., Lantos,P.L. and Goedert,M. (1998) Filamentous α-synuclein inclusions link multiple system atrophy with Parkinson's disease and dementia with Lewy bodies. *Neurosci. Lett.*, **251**, 205–208.
24. Doherty,C.P.A., Ulamec,S.M., Maya-Martinez,R., Good,S.C., Makepeace,J., Khan,G.N., van Oosten-Hawle,P., Radford,S.E. and Brockwell,D.J. (2020) A short motif in the N-terminal region of α-synuclein is critical for both aggregation and function. *Nat. Struct. Mol. Biol.*, **27**, 249–259.
25. Hoyer,W., Cherny,D., Subramaniam,V. and Jovin,T.M. (2004) Impact of the acidic C-terminal region comprising amino acids 109−140 on α-synuclein aggregation in vitro. *Biochemistry*, **43**, 16233–16242.
26. Levitan,K., Chereau,D., Cohen,S.I.A., Knowles,T.P.J., Dobson,C.M., Fink,A.L., Anderson,J.P., Goldstein,J.M. and Millhauser,G.L. (2011) Conserved C-terminal charge exerts a profound influence on the aggregation rate of α-synuclein. *J. Mol. Biol.*, **411**, 329–333.
27. Farzadfard,A., Pedersen,J.N., Meisl,G., Somavarapu,A.K., Alam,P., Goksøyr,L., Nielsen,M.A., Sander,A.F., Knowles,T.P.J., Pedersen,J.S., *et al.* (2022) The C-terminal tail of α-synuclein protects against aggregate replication but is critical for oligomerization. *Commun. Biol.*, **5**, 123.

28. Sorrentino,Z.A. and Giasson,B.I. (2020) The emerging role of α-synuclein truncation in aggregation and disease. *J. Biol. Chem.*, **295**, 10224–10244.
29. Pancoe,S.X., Wang,Y.J., Shimogawa,M., Perez,R.M., Giannakoulias,S. and Petersson,E.J. (2022) Effects of mutations and post-translational modifications on α-synuclein in vitro aggregation. *J. Mol. Biol.*, **434**, 167859.
30. Iyer,A., Roeters,S.J., Kogan,V., Woutersen,S., Claessens,M.M.A.E. and Subramaniam,V. (2017) C-Terminal truncated α-synuclein fibrils contain strongly twisted β-sheets. *J. Am. Chem. Soc.*, **139**, 15392–15400.
31. van der Wateren,I.M., Knowles,T.P.J., Buell,A.K., Dobson,C.M. and Galvagnion,C. (2018) C-terminal truncation of α-synuclein promotes amyloid fibril amplification at physiological pH. *Chem. Sci.*, **9**, 5506–5516.
32. Delenclos,M., Burgess,J.D., Lamprokostopoulou,A., Outeiro,T.F., Vekrellis,K. and McLean,P.J. (2019) Cellular models of alpha-synuclein toxicity and aggregation. *J. Neurochem.*, **150**, 566–576.
33. Uversky,V.N., Li,J. and Fink,A.L. (2001) Evidence for a partially folded intermediate in α-synuclein fibril formation. *J. Biol. Chem.*, **276**, 10737–10744.
34. Roeters,S.J., Iyer,A., Pletikapić,G., Kogan,V., Subramaniam,V. and Woutersen,S. (2017) Evidence for intramolecular antiparallel beta-sheet structure in alpha-synuclein fibrils from a combination of two-dimensional infrared spectroscopy and atomic force microscopy. *Sci. Rep.*, **7**, 41051.
35. Stephens,A.D., Zacharopoulou,M., Moons,R., Fusco,G., Seetaloo,N., Chiki,A., Woodhams,P.J., Mela,I., Lashuel,H.A., Phillips,J.J., *et al.* (2020) Extent of N-terminus exposure of monomeric alpha-synuclein determines its aggregation propensity. *Nat. Commun.*, **11**, 2820.
36. Munishkina,L.A., Henriques,J., Uversky,V.N. and Fink,A.L. (2004) Role of protein−water interactions and electrostatics in α-synuclein fibril formation. *Biochemistry*, **43**, 3289–3300.
37. Antony,T., Hoyer,W., Cherny,D., Heim,G., Jovin,T.M. and Subramaniam,V. (2003) Cellular polyamines promote the aggregation of α-synuclein. *J. Biol. Chem.*, **278**, 3235–3240.
38. Cohlberg,J.A., Li,J., Uversky,V.N. and Fink,A.L. (2002) Heparin and other glycosaminoglycans stimulate the formation of amyloid fibrils from α-synuclein in vitro. *Biochemistry*, **41**, 1502–1511.
39. Kam,T.-I., Mao,X., Park,H., Chou,S.-C., Karuppagounder,S.S., Umanah,G.E., Yun,S.P., Brahmachari,S., Panicker,N., Chen,R., *et al.* (2018) Poly(ADP-ribose) drives pathologic α-synuclein neurodegeneration in Parkinson's disease. *Science*, **362**, eaat8407.
40. Tartaglia,G.G., Pawar,A.P., Campioni,S., Dobson,C.M., Chiti,F. and Vendruscolo,M. (2008) Prediction of aggregation-prone regions in structured proteins. *J. Mol. Biol.*, **380**, 425–436.
41. Moriarty,D.F., Vagts,S. and Raleigh,D.P. (1998) A role for the C-terminus of calcitonin in aggregation and gel formation: a comparative study of C-terminal fragments of human and salmon calcitonin. *Biochem. Biophys. Res. Commun.*, **245**, 344–348.
42. Reches,M., Porat,Y. and Gazit,E. (2002) Amyloid fibril formation by pentapeptide and tetrapeptide fragments of human calcitonin. *J. Biol. Chem.*, **277**, 35475–35480.
43. Luca,S., Yau,W.-M., Leapman,R. and Tycko,R. (2007) Peptide conformation and supramolecular organization in amylin fibrils: constraints from solid state NMR. *Biochemistry*, **46**, 13505–13522.
44. Pedersen,J.S., Dikov,D. and Otzen,D.E. (2006) N- and C-terminal hydrophobic patches are involved in fibrillation of glucagon. *Biochemistry*, **45**, 14503–14512.
45. Török,M., Milton,S., Kayed,R., Wu,P., McIntire,T., Glabe,C.G. and Langen,R. (2002) Structural and dynamic features of Alzheimer's Abeta peptide in amyloid fibrils studied by site-directed spin labeling. *J. Biol. Chem.*, **277**, 40810–40815.
46. Lührs,T., Ritter,C., Adrian,M., Riek-Loher,D., Bohrmann,B., Döbeli,H., Schubert,D. and Riek,R. (2005) 3D structure of Alzheimer's amyloid-beta(1-42) fibrils. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 17342–17347.
47. Miake,H., Mizusawa,H., Iwatsubo,T. and Hasegawa,M. (2002) Biochemical characterization of the core structure of alpha-synuclein filaments. *J. Biol. Chem.*, **277**, 19213–19219.
48. Der-Sarkissian,A., Jao,C.C., Chen,J. and Langen,R. (2003) Structural organization of α-synuclein fibrils studied by site-directed spin labeling. *J. Biol. Chem.*, **278**, 37530–37535.
49. Del Mar,C., Greenbaum,E.A., Mayne,L., Englander,S.W. and Woods,V.L. (2005) Structure and properties of α-synuclein and other amyloids determined at the amino acid level. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 15477–15482.
50. Lancaster,A.K., Nutter-Upham,A., Lindquist,S. and King,O.D. (2014) PLAAC: a web and command-line application to identify proteins with prion-like amino acid composition. *Bioinforma. Oxf. Engl.*, **30**, 2501–2502.
51. Roseman,M.A. (1988) Hydrophilicity of polar amino acid side-chains is markedly reduced by flanking peptide bonds. *J. Mol. Biol.*, **200**, 513–522.
52. Campen,A., Williams,R.M., Brown,C.J., Meng,J., Uversky,V.N. and Dunker,A.K. (2008) TOP-IDP-scale: a new amino acid scale measuring propensity for intrinsic disorder. *Protein Pept. Lett.*, **15**, 956–963.
53. Castello,A., Fischer,B., Eichelbaum,K., Horos,R., Beckmann,B.M., Strein,C., Davey,N.E., Humphreys,D.T., Preiss,T., Steinmetz,L.M., *et al.* (2012) Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell*, **149**, 1393–1406.
54. Klus,P., Bolognesi,B., Agostini,F., Marchese,D., Zanzoni,A. and Tartaglia,G.G. (2014) The cleverSuite approach for protein characterization: predictions of structural properties, solubility, chaperone requirements and RNA-binding abilities. *Bioinforma.*, **30**, 1601–1608.
55. Bellucci,M., Agostini,F., Masin,M. and Tartaglia,G.G. (2011) Predicting protein associations with long noncoding RNAs. *Nat. Methods*, **8**, 444–445.
56. Lang,B., Armaos,A. and Tartaglia,G.G. (2019) RNAct: protein-RNA interaction predictions for model organisms with supporting experimental data. *Nucleic Acids Res.*, **47**, D601–D606.
57. Armaos,A., Colantoni,A., Proietti,G., Rupert,J. and Tartaglia,G.G. (2021) catRAPID omics v2.0: going deeper and wider in the prediction of protein-RNA interactions. *Nucleic Acids Res.*, **49**, W72–W79.
58. Hernández,F., Ferrer,I., Pérez,M., Zabala,J.C., Del Rio,J.A. and Avila,J. (2023) Tau aggregation. *Neuroscience*, **518**, 64–69.
59. Von Bergen,M., Friedhoff,P., Biernat,J., Heberle,J., Mandelkow,E.-M. and Mandelkow,E. (2000) Assembly of τ protein into Alzheimer paired helical filaments depends on a local sequence motif (306 VQIVYK 311) forming β structure. *Proc. Natl. Acad. Sci*, **97**, 5129–5134.
60. Wang,X., Wang,D., Zhao,J., Qu,M., Zhou,X., He,H. and He,R. (2006) The proline-rich domain and the microtubule binding domain of protein tau acting as RNA binding domains. *Protein Pept. Lett.*, **13**, 679–685.
61. Vendruscolo,M. and Fuxreiter,M. (2022) Sequence determinants of the aggregation of proteins within condensates generated by liquid-liquid phase separation. *J. Mol. Biol.*, **434**, 167201.
62. Livi,C.M., Klus,P., Delli Ponti,R. and Tartaglia,G.G. (2016) catRAPID signature: identification of ribonucleoproteins and RNA-binding regions. *Bioinformatics*, **32**, 773–775.
63. Monti,M., Armaos,A., Fantini,M., Pastore,A. and Tartaglia,G.G. (2021) Aggregation is a context-dependent constraint on protein evolution. *Front. Mol. Biosci.*, **8**, 678115.
64. Tartaglia,G.G., Cavalli,A., Pellarin,R. and Caflisch,A. (2004) The role of aromaticity, exposed surface, and dipole moment in determining protein aggregation rates. *Protein Sci.*, **13**, 1939–1941.

65. Waxman,E.A., Mazzulli,J.R. and Giasson,B.I. (2009) Characterization of hydrophobic residue requirements for α-synuclein fibrillization. *Biochemistry*, **48**, 9427–9436.
66. Zhang,Z., Kang,S.S., Liu,X., Ahn,E.H., Zhang,Z., He,L., Iuvone,P.M., Duong,D.M., Seyfried,N.T., Benskey,M.J., *et al.* (2017) Asparagine endopeptidase cleaves α-synuclein and mediates pathologic activities in Parkinson's disease. *Nat. Struct. Mol. Biol.*, **24**, 632–642.
67. Anand,B.G., Prajapati,K.P. and Kar,K. (2018) Aβ1-40 mediated aggregation of proteins and metabolites unveils the relevance of amyloid cross-seeding in amyloidogenesis. *Biochem. Biophys. Res. Commun.*, **501**, 158–164.
68. Tartaglia,G.G. and Caflisch,A. (2007) Computational analysis of the S. cerevisiae proteome reveals the function and cellular localization of the least and most amyloidogenic proteins. *Proteins*, **68**, 273–278.
69. Murakami,K., Izuo,N. and Bitan,G. (2022) Aptamers targeting amyloidogenic proteins and their emerging role in neurodegenerative diseases. *J. Biol. Chem.*, **298**, 101478.
70. Tsukakoshi,K., Abe,K., Sode,K. and Ikebukuro,K. (2012) Selection of DNA aptamers that recognize α-synuclein oligomers using a competitive screening method. *Anal. Chem.*, **84**, 5542–5547.
71. Sanders,D.W., Kedersha,N., Lee,D.S.W., Strom,A.R., Drake,V., Riback,J.A., Bracha,D., Eeftens,J.M., Iwanicki,A., Wang,A., *et al.* (2020) Competing protein-RNA interaction networks control multiphase intracellular organization. *Cell*, **181**, 306–324.
72. Zagrovic,B., Bartonek,L. and Polyansky,A.A. (2018) RNA-protein interactions in an unstructured context. *FEBS Lett.*, **592**, 2901–2916.
73. Zaharias,S., Zhang,Z., Davis,K., Fargason,T., Cashman,D., Yu,T. and Zhang,J. (2021) Intrinsically disordered electronegative clusters improve stability and binding specificity of RNA-binding proteins. *J. Biol. Chem.*, **297**, 100945.
74. Yang,P., Mathieu,C., Kolaitis,R.-M., Zhang,P., Messing,J., Yurtsever,U., Yang,Z., Wu,J., Li,Y., Pan,Q., *et al.* (2020) G3BP1 is a tunable switch that triggers phase separation to assemble stress granules. *Cell*, **181**, 325–345.
75. Ma,W., Zhen,G., Xie,W. and Mayr,C. (2021) In vivo reconstitution finds multivalent RNA–RNA interactions as drivers of mesh-like condensates. *eLife*, **10**, e64252.
76. Vandelli,A., Cid Samper,F., Torrent Burgas,M., Sanchez de Groot,N. and Tartaglia,G.G. (2022) The interplay between disordered regions in RNAs and proteins modulates interactions within stress granules and processing bodies. *J. Mol. Biol.*, **434**, 167159.
77. Zhang,S., Li,J., Xu,Q., Xia,W., Tao,Y., Shi,C., Li,D., Xiang,S. and Liu,C. (2023) Conformational dynamics of an α-synuclein fibril upon receptor binding revealed by insensitive nuclei enhanced by polarization transfer-based solid-state nuclear magnetic resonance and cryo-electron microscopy. *J. Am. Chem. Soc.*, **145**, 4473–4484.
78. Schweighauser,M., Shi,Y., Tarutani,A., Kametani,F., Murzin,A.G., Ghetti,B., Matsubara,T., Tomita,T., Ando,T., Hasegawa,K., *et al.* (2020) Structures of α-synuclein filaments from multiple system atrophy. *Nature*, **585**, 464–469.
79. Sawaya,M.R., Hughes,M.P., Rodriguez,J.A., Riek,R. and Eisenberg,D.S. (2021) The expanding amyloid family: structure, stability, function, and pathogenesis. *Cell*, **184**, 4857–4873.
80. Sender,R., Fuchs,S. and Milo,R. (2016) Revised estimates for the number of human and bacteria cells in the body. *PLoS Biol.*, **14**, e1002533.
81. Russo,J. and Russo,I.H. (2014) In: *Techniques and Methodological Approaches in Breast Cancer Research*. Springer, NY.
82. Hegde,M.L. and Rao,K.S.J. (2007) DNA induces folding in α-synuclein: understanding the mechanism using chaperone property of osmolytes. *Arch. Biochem. Biophys.*, **464**, 57–69.
83. Pinho,R., Paiva,I., Jerčić,K.G., Fonseca-Ornelas,L., Gerhardt,E., Fahlbusch,C., Garcia-Esparcia,P., Kerimoglu,C., Pavlou,M.A.S., Villar-Piqué,A., *et al.* (2019) Nuclear localization and phosphorylation modulate pathological effects of alpha-synuclein. *Hum. Mol. Genet.*, **28**, 31–50.
84. Hallacli,E., Kayatekin,C., Nazeen,S., Wang,X.H., Sheinkopf,Z., Sathyakumar,S., Sarkar,S., Jiang,X., Dong,X., Di Maio,R., *et al.* (2022) The Parkinson's disease protein alpha-synuclein is a modulator of processing bodies and mRNA stability. *Cell*, **185**, 2035–2056.
85. Siegert,A., Rankovic,M., Favretto,F., Ukmar-Godec,T., Strohäker,T., Becker,S. and Zweckstetter,M. (2021) Interplay between tau and α-synuclein liquid–liquid phase separation. *Protein Sci.*, **30**, 1326–1336.
86. Sugimoto,S., Arita-Morioka,K., Mizunoe,Y., Yamanaka,K. and Ogura,T. (2015) Thioflavin T as a fluorescence probe for monitoring RNA metabolism at molecular and cellular levels. *Nucleic Acids Res.*, **43**, e92.
87. Zacco,E., Graña-Montes,R., Martin,S.R., de Groot,N.S., Alfano,C., Tartaglia,G.G. and Pastore,A. (2019) RNA as a key factor in driving or preventing self-assembly of the TAR DNA-binding protein 43. *J. Mol. Biol.*, **431**, 1671–1688.
88. Weston,L.J., Bowman,A.M., Osterberg,V.R., Meshul,C.K., Woltjer,R.L. and Unni,V.K. (2022) Aggregated alpha-synuclein inclusions within the nucleus predict impending neuronal cell death in a mouse model of Parkinsonism. *Int. J. Mol. Sci.*, **23**, 15294.
89. Srinivasan,E., Chandrasekhar,G., Chandrasekar,P., Anbarasu,K., Vickram,A.S., Karunakaran,R., Rajasekaran,R. and Srikumar,P.S. (2021) Alpha-synuclein aggregation in Parkinson's disease. *Front. Med.*, **8**, 736978.
90. Geertsma,H.M., Suk,T.R., Ricke,K.M., Horsthuis,K., Parmasad,J.-L.A., Fisk,Z.A., Callaghan,S.M. and Rousseaux,M.W.C. (2022) Constitutive nuclear accumulation of endogenous alpha-synuclein in mice causes motor impairment and cortical dysfunction, independent of protein aggregation. *Hum. Mol. Genet.*, **31**, 3613–3628.

Received: April 25, 2023. **Revised:** August 15, 2023. **Editorial Decision:** September 14, 2023. **Accepted:** September 26, 2023