Computationally Designed RNA Aptamers Enable Selective Detection of FUS Pathology in ALS

Elsa Zacco [1], Martina Gilodi [1], Alexandros Armaos [1], Fergal M. Waldron[2], Jakob Rupert [1], Neil Schneider [3], Jenna M. Gregory [3] & Gian Gaetano Tartaglia[1]

1. RNA Systems Biology, Istituto Italiano di Tecnologia, Genoa, Via Enrico Melen, 83, 16152, Genoa, Italy
2. Institute of Medical Sciences, University of Aberdeen, Foresterhill, Aberdeen AB25 2ZD, United Kingdom
3. Department of Neurology, Columbia University Irving Medical Center, 710 West 168th Street, New York, NY 10032, USA

ABSTRACT

FUS is a multifunctional RNA/DNA-binding protein whose mislocalization and aggregation into insoluble inclusions are defining features of neurodegenerative diseases such as amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD). Conventional antibody-based detection methods are often hindered by epitope masking and the structural heterogeneity of FUS aggregates. In this study, we present a computational pipeline for the *in silico* design of RNA aptamers capable of targeting FUS, including its aggregated forms. By integrating CLIP data, de novo motif enrichment, secondary structure prediction, and catRAPID-based interaction scoring, we screened over 28,000 RNA sequences and identified high-affinity aptamer candidates predicted to bind solvent-exposed, aggregation-resistant regions of FUS. The top candidates adopted stable hairpin structures favorable for interaction with the RNA recognition motif (RRM) and contained a GGU motif in the 3′ region to enhance binding to the zinc finger domain. Biolayer interferometry confirmed strong binding affinities in the low nanomolar range, with a high correlation to computational predictions. Circular dichroism analysis further validated the presence of stable hairpin conformations. When tested in post-mortem brain tissue from ALS patients harboring the p.P525L FUS mutation, the aptamers selectively labeled pathological FUS aggregates—including nuclear inclusions and nucleolar localization patterns—while demonstrating greater specificity than conventional antibodies. Together, these findings establish a rapid, cost-effective, and structure-guided approach for designing RNA aptamers that reliably detect FUS pathology with high sensitivity and specificity.

INTRODUCTION

FUsed in Sarcoma (FUS), also known as Translocated in LipoSarcoma (TLS), is a multifunctional DNA/RNA-binding protein with significant roles in maintaining cellular homeostasis through its involvement in RNA metabolism, stress response, gene expression, and DNA repair [1]. FUS shuttles between the nucleus and the cytoplasm, and this dynamic localization is crucial for its function in both RNA metabolism and response to cellular stress [2]. Through its interaction with RNA, FUS regulates splicing, transport, stability, and translation and it is involved in the processing of pre-mRNA into mature mRNA [3].

Mutations or dysregulation of FUS protein have been linked to several neurological disorders, particularly neurodegenerative diseases. For example, mutations in the FUS gene have been identified in a subset of familial amyotrophic lateral sclerosis (ALS) cases [4] and alterations in the FUS gene have also been associated with certain forms of frontotemporal lobar degeneration (FTLD) [5,6]. In these and other pathologies, the physiological functions of FUS are compromised by the accumulation of the protein in toxic aggregates, insoluble structures in which FUS transit from a liquid state to a solid-like form [7]. FUS aggregation –or liquid-to-solid phase transition (LSPT)– is thought to contribute to disease pathogenesis particularly with regards to ALS and FTLD, for which the relocation of FUS from the nucleus to the cytosol and the consequent cytoplasmic protein accumulation are considered pathological markers of disease [8,9]. The exact molecular mechanism underlying FUS aggregation and the way FUS LSPT correlates to disease development is still unknown. Various factors –including genetic mutations, changes in cellular stress response, impaired protein quality control mechanisms, and disruption of RNA metabolism– are believed to contribute to FUS aggregation and its pathological consequences [10,11] .

Efforts to understand and track FUS aggregation are critical for prompt diagnosis and intervention. Currently, ALS-FTD diagnosis is generally made based on clinical symptoms, electromyography findings, and genetic testing for FUS mutations or for other disease related proteins (such as SOD1, C9orf72, and TDP-43). The presence of FUS aggregates in tissues strengthens the diagnosis, particularly in familial ALS cases or those with FTD associated with ALS. However, the use of antibodies to visualise aggregated FUS can be challenging and limiting: on the one hand, protein aggregates can vary greatly in structure and conformation, and antibodies targeting one specific aggregate structure are generally not effective for others; on the other hand, aggregated proteins can share similar epitopes with others or may have cryptic epitopes —hidden regions exposed only in the aggregated state— that can lead to non-specific binding, causing unwanted cross-reactivity. The structural diversity and flexibility, essential for FUS diverse functions [12], including its interaction with RNA [12], are also what makes this protein so challenging to target with conventional approaches.

There is therefore a need for an alternative means by which to target and track FUS aggregation with higher accuracy. In this context, employing RNA-based aptamers to interact and engage with FUS aggregates might represent a promising solution. Aptamers are short single-stranded oligonucleotides that fold into a particular three-dimensional structure and that bind a target molecule with high specificity [13,14] . Using aptamers, a combined effect of structural and sequence recognition could guarantee an affinity towards the target within the picomolar-nanomolar range. The small size —usually between 5 and 15 kDa— and their chemical features give several advantages to the aptamers over antibodies in terms of tissue penetration ability, chemical stability and solubility

properties. Their nature (DNA or RNA) renders them non-immunogenic, chemically versatile and synthetically accessible.

In the case of FUS, attempts have been made to narrow down the RNA sequences and structures with which this protein might show preferential interaction. The literature presents a few cases in which, taking advantage of its natural ability to bind to nucleic acids, some oligonucleotide sequences have been identified to interact with FUS with certain specificity and have been employed mostly to clear mechanistic details of FUS binding to its RNA partners. Sequential Evolution of Ligands by Exponential Enrichment approach (SELEX) [15] revealed that FUS displays stronger binding avidity towards RNA sequences containing the GGU motif [16], while studies on the structural preferences of FUS showed that the protein interacts favorably with hairpin RNAs or with G-rich sequences forming the G-quadruplex quaternary structure [17,18]. Moreover, nuclear magnetic resonance (NMR) experiments demonstrated that both RRM and Zn finger domains of FUS take concomitant part to the binding of RNA. While the first motif is most likely to be involved in the binding with the structural components (hailpins/G-quadruplexes), the latter confer the protein its preference towards GGU-rich RNA sequences [12].

With this knowledge at hand, the possibility of tailor-design RNA sequences to selectively bind FUS appears more accessible, if yet difficult to accomplish with the currently available experimental approaches. With this work, we aim at offering a more approachable *in silico* aptamer selection method for the protein FUS and prove the use of selected aptamers in tracking the aggregation of FUS *in vitro*, paving the way for development of new diagnostic tools for FUS-driven neurodegenerative diseases. The generation of specific FUS-targeting RNA aptamers would potentially open the door to multiple applications, spreading from targeted therapy to biomarker detection, passing through the development of novel research tools.

RESULTS

A computational pipeline for the *in silico* design of RNA aptamers targeting FUS

One of the main goals of this work was to define a pipeline that would exploit computational tools to design RNA sequences capable of selectively binding proteins by leveraging its natural nucleic acid-binding properties [14]. By incorporating an *in silico* strategy for aptamer design, we significantly expedited and enhanced the aptamer selection process, using computational algorithms and experimental data to identify and optimize high-affinity candidates. This approach enabled a rapid screening and refinement of sequences, ensuring that aptamers with the greatest binding potential to FUS were prioritized.

Aptamers were specifically designed to interact with the RNA-binding portion of FUS, encompassing amino acids 269–454 (hereafter referred to as $FUS_{269–454}$). $FUS_{269–454}$ includes the RRM, Zn finger, and one of the three RGG domains, all elements implicated in RNA interactions [12,19]. The aptamer candidates were selected based on three principal criteria: (i) the presence of enriched sequence motifs derived from FUS CLIP datasets; (ii) the ability to adopt a stable hairpin fold, favoring binding to the RRM domain [17]; and (iii) the presence of a GGU triplet downstream of the stem-loop, which facilitates interaction with the Zn finger domain of FUS [12]. Most importantly, previous studies [20] indicate that in aggregation-prone proteins, the structured amyloid core is typically buried, while RNA-binding domains tend to remain solvent-exposed and accessible on the aggregate surface. Consistent with this model, aggregation propensity analysis of FUS shows that the RNA-binding regions included in $FUS_{269–454}$ are predicted to be external to the aggregation core and thus available for aptamer binding even in aggregated states (**Suppl. Fig. 1**). This key property underlies the success of targeting FUS aggregates via RNA aptamers.

From ENCODE eCLIP data (HepG2 and K562), 4,203 high-confidence FUS-binding sites were identified (fold enrichment > 3). Of these, 401 overlapped with PAR-CLIP data and were used as the positive set for motif discovery, while 1,075 control sequences were sampled from low-confidence, non-overlapping eCLIP peaks (**Fig. 1A**). *cat*RAPID effectively distinguished binding from non-binding regions identified by eCLIP (AUC = 0.868, **Fig. 1B; Materials and Methods**) [21]. Motif analysis identified 11 FUS-recognized sequence motifs, which were used to generate 28,102 candidate aptamers with 0–20 nt 3' extensions.

We assessed domain-specific binding propensities by computing *cat*RAPID scores across our full pool of 28,102 candidate sequences. Median interaction scores for the RGG1, RGG2, and RGG3 domains were –0.66, –0.36, and –0.60, respectively, whereas the RRM and the combined RRM + RGG + ZnF region scored 7.04 and 5.89 (**Fig. 1C**), confirming strong predicted affinity of our motifs for the canonical RNA-binding surfaces.

A custom scoring function was developed to evaluate secondary structures [22] , prioritizing those with a perfect hairpin loop at the 5' end, followed by a linear, unstructured sequence at the 3' end [12]. A total of 1,463 aptamers achieved an optimal secondary structure score (SS_score = 1), and 2,613 aptamers scored above 0.9 (SS_score > 0.9). Among these, 605 aptamers contained at least one GGU subsequence in the 3' extension and none GGU subsequence in the hairpin loop. These

aptamers were finally ranked for their predicted interaction score with $FUS_{269-454}$ defining a chart of most promising candidates (**Fig. 1D**).

Experimentally determined binding strengths agree with computational predictions

To verify the accuracy of the computational pipeline to identify RNA aptamer candidates targeting FUS with good interaction strength, a series of RNA sequences have been selected to experimentally determine their binding strength towards $FUS_{269-454}$. The dissociation constant (K_d) was used as a measure of interaction strength between the RNA and the protein and was determined via biolayer interferometry (BLI), an optical technique to measure interactions between molecules. From the aptamer list generated computationally, both the most promising candidates and a few controls deriving from intermediate steps of the aptamers development were tested. The aptamers screened experimentally and the measured K_d are reported in **Table 1** according to their binding strength. The seven aptamers derived from the latest steps of the computational design pipeline are indeed the ones displaying the strongest affinity towards $FUS_{269-454}$, with a K_d ranging from 23 nM to 740 nM (**Fig. 2A-G**). The two aptamers deriving from intermediate steps of the pipeline, extracted at a stage in which they were not yet scored for their hairpin structure, display a slightly worse K_d of 1100 and 1900 nM (**Fig. 2H-I**).

Taken together, there is an excellent correlation between the predicted interaction scores and the experimentally measured K_d (Pearson correlation = -0.837, **Suppl. Fig. 2**).

To further verify the superiority of the *in silico*-designed best aptamer candidates compared to other RNA sequences developed by others using different means, the K_d between $FUS_{269-454}$ and other aptamers found in literature was determined. Two aptamers, here called ComboNMR17 and ComboNMR22, were derived from RNA sequences deposited in the protein data bank (PDB) in association with either the RRM [23] or the ZnF of FUS [24]. The RNA oligonucleotides were combined together to generate aptamers that could potentially bind both domains of FUS. In both these cases, a K_d of 1600 nM was determined (**Suppl. Fig. 3A-B and Suppl. Table 1**), making these aptamers comparable to the ones generated during the intermediate steps of the computational pipeline.

The top aptamer candidates were also compared to other three RNA oligonucleotides reported in literature to bind FUS (**Suppl. Fig. 3C-E**). The RNA named PrD was chosen because, as the aptamers designed *in silico*, it folds into a stable hairpin and it contains the GGU motif on its tail [25]. The other two aptamers, named PSD95GQ2 and Shank1A, instead share the ability to fold into G-quadruplexes [18]. This structural feature was not preferred during the development of the *in silico*-disigned sequences because of its known promiscuity of binding [26]. Nevertheless, the affinity of these sequences towards FUS was reported to be quite high. For these three RNA aptamers, the K_d reported in the literature was determined by methods different from BLI and the affinity was tested towards a different FUS-derived construct compared to $FUS_{269-454}$. For these reasons, the binding strength of these sequences was again determined using the same conditions as the ones employed to calculate it for the *in silico*-generated aptamers (*e.i.* using BLI to derive the K_d with $FUS_{269-454}$).

The K_d determined by BLI for PrD was 2100 nM (**Suppl. Fig. 3C**), placing this aptamer in the same range of the controls ComboNMR17 and ComboNMR22, with which PrD shares the hairpin structure. PSD_95_GQ2 and Shank1A displayed instead a K_d with $FUS_{269-454}$ of 220 nM and 713 nM, respectively (**Suppl. Fig 3D-E**). These binding strengths are comparable to some of the RNA aptamers generated

via the computational pipeline but are not as strong binders as the top four RNA sequences generated *in silico*.

To conclude the interrogation of possible RNA aptamer controls, two sequences of lengths comparable to the most promising candidates, lacking both the hairpin structure and the GGU motif, were selected (see **Material and Methods** section). For both these sequences, named according to their length Neg24nt and Neg42nt, it was not possible to determine a K_d for $FUS_{269-454}$, since no binding occurred in the selected conditions (**Suppl. Fig. 3F-G**).

In conclusion, the RNA aptamers designed through the computational pipeline described above are superior or equal to other RNA sequences reported in the literature in binding the RNA binding region of FUS comprising the RRM, the ZnF and one RGG repeat. Having been designed entirely *in silico*, these RNA aptamers present the additional advantage of a faster and cost effective generation compared to others. The use of computational tools also enables the screening of sequences that not only are predicted to strongly bind FUS but also to target it with high specificity, by selecting RNA aptamers predicted to prefer FUS to other proteins of similar composition and length.

Circular dichroism reveals the presence of stable hairpins in designed aptamers

Since one of the main features of the aptamers designed with the computational pipeline was a high hairpin score, the most promising candidates were investigated for their ability to fold into this structure within the conditions selected for this study. Alongside the three most performing aptamers showing the strongest binding strength towards $FUS_{269-454}$, also a negative control was analyzed.

The top three aptamers —TopScorerLong, Enriched and TopScorerShort— share a similarly-shaped CD profile, with a main maximum at 265 nm and a minimum at 240 nm (**Fig. 3**). These features are generally indicative of an RNA sequence, and indeed we find these characteristics also in the CD profile of Neg24nt, albeit in this case both maximum and minimum are shifted towards higher UV regions (270 nm and 245 nm, respectively) (**Fig. 3**). It is also observable an important difference in the intensities of the peaks among the three top scorer aptamers: position and amplitude of the peaks may markedly differ with sequence, not only because the chromophore differ but also because of the different conformational properties of each base. What it is important to observe for this study is the presence, in all three cases, of a weak shoulder at 285-290 nm, which is typical feature indicative of an hairpin [27]. This shoulder is absent in the CD profile of Neg24nt.

To confirm the stability of the hairpins in the selected conditions, the CD spectra were acquired at increasing temperature and the CD profiles at 30 °C were compared with the ones recorded at 70 °C (**Fig. 3**). For the same sequence at the same concentration, a loss of intensity is indicative of an increased instability of the RNA. The top three aptamer candidates display a loss of intensity of the maximum at 265 nm of around 1.2 folds, while the shoulder indicative of the hairpin remains unvaried in all cases. Instead, the maximum at 270 nm of Neg24nt drops more than twice from its initial intensity (**Fig. 3**). This suggests strong secondary structure features for the three top aptamers and the lack of any stable configuration for Nag24nt.

These results indicate that the aptamers designed computationally to fold into a hairpin structure clearly display this feature in the conditions selected for this study, while the negative control, programmed to remain linear, does not show any particular secondary structure characteristics.

FUS aptamers recognise aggregates in tissue more selectively than antibodies

To test the ability of selected aptamers to bind to physiologically relevant FUS aggregates *in vivo*, we obtained post-mortem tissue from two people who had an underlying p.P525L FUS mutation, one who had received FUS ASO therapy and one who had not [28] and a non-neurological control. We performed immunohistochemical staining using the TopScorerLong and TopScorerShort sequences tagged with a 3' biotin molecule to facilitate detection using an anti-biotin antibody immunohistochemistry method as published previously [29]. We compared staining profiles of the two aptamers against a commercially available FUS antibody. FUS antibody staining in the non-neurological control case demonstrated a normal, predominantly nuclear staining pattern of FUS with no evidence of nuclear or cytoplasmic aggregation. In contrast to this, FUS antibody staining in the untreated p.P525L *FUS* mutation carrier demonstrated both normal and abnormal FUS staining including large cytoplasmic aggregates (**Fig. 4A**). The antibody staining in the ASO treated p.P525L *FUS* mutation carrier demonstrated no evidence of either normal, nor pathological FUS staining, commensurate with near complete knockdown and clearance of FUS protein. Contrastingly, the FUS aptamer staining identified only abnormal FUS staining patterns, with no staining visible in the non-neurological control and extensive pathological FUS staining in the untreated individual. Aptamer staining additionally revealed large and extensive nuclear aggregation events that are not readily identified using antibody staining alone. Similar to the antibody staining, no immunoreactivity is seen in the ASO treated individual using FUS aptamer staining. Digital image analysis was performed in QuPath using superpixel masks (**Fig. 4B**) [30]. Quantification, using superpixel H-score demonstrates equivalent sensitivity of aptamer staining compared to antibody and improved specificity (no detection in control and ASO treated individuals) of aptamer staining compared to antibody. No difference was identified between the long and short FUS aptamer sequences, both demonstrated equivalent sensitivity and specificity (Figure 4C). Amongst features readily identified by aptamer staining, nuclear pathology is most notable (Figure 5A). Nuclear pathology is less readily visualised in antibody-stained cases as normal nuclear localisation may be obscuring the nuclear pathology. Nuclear pathological features include (i) nucleolar decoration, (ii) large nuclear inclusions, and (iii) nuclear membrane pathology (**Fig. 5A**). These are in addition to the large cytoplasmic aggregates, and cytoplasmic puncta that can be readily visualised by both the antibody and the aptamer sequences (**Fig. 5B**).

CONCLUSIONS

In this work, we developed and validated a novel computational approach for the generation of RNA aptamers selectively targeting FUS aggregates. FUS is a multifunctional RNA/DNA-binding protein essential for numerous aspects of RNA metabolism, including splicing, transport, translation, and stability. In pathological conditions such as ALS and FTLD, FUS undergoes aberrant aggregation, resulting in the disruption of its normal functions and contributing to neuronal dysfunction and death.

The need for developing RNA aptamers for pathological FUS aggregates lies in the inherent limitations of conventional antibodies when it comes to detecting proteins that undergo state transition. Aggregated proteins often exhibit significant conformational variability—a hallmark of neurodegenerative conditions like ALS and FTLD— and antibodies typically recognize specific epitopes that may become obscured or structurally altered in aggregated states, reducing detection reliability [31,32]. RNA aptamers possess greater structural adaptability and we can computationally tailor them for high-affinity interactions, making them well-suited to bind the diverse and often concealed surfaces of misfolded protein aggregates [14,33]. Notably, our RNA aptamers are entirely designed *in silico*, prioritizing candidates with high predicted affinity for the FUS RNA-binding region, using *cat*RAPID interaction propensity scores alongside secondary structure predictions favoring hairpin conformations.

Experimental validation confirmed that the top-ranked aptamers bind FUS with high affinity in the nanomolar range, as determined by biolayer interferometry. Circular dichroism analysis further confirmed the stability of their secondary structure under physiological conditions. Importantly, the results of our post-mortem immunohistochemical analyses highlight the diagnostic potential of RNA aptamers in detecting FUS pathology, particularly in ALS cases linked to the P525L mutation where antibody-based methods are limiting [34]. FUS-targeting aptamers demonstrated high specificity, staining only pathological aggregates in the untreated mutation carrier, with no signal in the non-neurological control or ASO-treated individual [28]. In contrast, antibody staining also detected normal nuclear FUS in controls, indicating broader reactivity that may obscure subtle pathology. The absence of aptamer signal in ASO-treated tissue—consistent with antibody results—further supports their target specificity in complex samples.

In addition, aptamer staining revealed nuclear pathological features not readily detected by antibodies, marking a major advance in understanding FUS-associated ALS. Because antibodies label both normal and pathological FUS, nuclear inclusions may be masked. Aptamers, however, can highlight abnormal nuclear patterns including nucleolar decoration, large intranuclear inclusions, and nuclear membrane abnormalities, suggesting that nuclear aggregation may precede or occur alongside cytoplasmic pathology.

Quantitative analysis confirmed that aptamers match antibodies in sensitivity while offering improved specificity. These findings support the clinical potential of aptamer-based diagnostics, which require high signal-to-noise ratios and reproducibility. Aptamers also mirrored antibodies in identifying cytoplasmic aggregates and puncta, confirming their broad detection capability. However, their unique ability to detect nuclear pathology makes them a valuable complement rather than a

replacement. Dual-staining approaches using both tools could offer a more complete pathological picture.

The absence of staining in ASO-treated tissue across both modalities further validates that aptamer signals are tightly linked to FUS presence and not structural artifacts, highlighting their potential in tracking therapeutic response [28]..

In conclusion, we established a rapid, cost-effective, and scalable computational strategy for the generation of high-affinity RNA aptamers against aggregation-prone proteins. These same aptamers can be employed as a highly specific and sensitive tool for identifying pathological FUS, with added value in revealing nuclear features invisible to antibodies. The ability of the aptamers to discriminate between physiological and pathological forms of FUS could lead to better understanding of disease mechanisms, particularly regarding nuclear and cytoplasmic aggregation dynamics.

These findings lay the groundwork for broader application in neurodegenerative diagnostics and for expanding aptamer use to other protein aggregation disorders. Nevertheless, several limitations of this study must be acknowledged. The aptamers were designed against the soluble RNA-binding region of FUS, and although aggregation is expected to partially preserve these domains, it remains possible that certain aggregate conformations or post-translational modifications may reduce binding efficiency. Additionally, while we demonstrated specificity in post-mortem brain tissue, the behavior of these aptamers in more complex biological environments, such as living tissue or biofluids, remains to be fully assessed. Moreover, although specificity towards FUS was prioritized computationally, cross-reactivity with other RNA-binding proteins sharing similar domains cannot be entirely excluded and warrants further investigation. Finally, the histological application of aptamers, while powerful, will require optimization and standardization before it can be integrated into routine diagnostic workflows.

Acknowledgements

We thank the “RNA Flagship” at IIT, as well as all members of the J.M.G. and G.G.T. research groups, for their support and contributions. This research was supported by the European Research Council [RIBOMYLOME no. 309545 to G.G.T. and ASTRA no. 855923 to G.G.T.], and the H2020 programme [IASIS no. 727658 to G.G.T. and INFORE no. 825080 to G.G.T.].

Competing interests

The authors declare the following competing interests: E.Z., M.G., A.A., and G.G.T. have filed the Italian priority patent application no. 102024000004522, which covers the technology described in the present publication.

METHODS

Aptamer design. We targeted the RNA-binding region of FUS (residues 269–454: RRM, RGG2, ZnF) to exploit its natural nucleic-acid affinity. High-confidence FUS-binding sites were called from ENCODE eCLIP data (ENCODE project [35]) (HepG2 & K562; two replicates each) after filtering for $\log_2$(fold-enrichment) > 3 (**Fig. 1A**). Merging all four replicates yielded 4,203 reproducible binding sites; To measure the ability of *cat*RAPID to reproduce the experimental data, we used these eCLIP binding sites as the positive set and an equally sized random selection of sites from the negative set as controls. A receiver operating characteristic (ROC) curve was computed to assess the ability of catRAPID to distinguish true FUS-binding sites from non-binding regions. The resulting area under the curve (AUC) was 0.868, indicating strong predictive performance (**Fig. 1B**).

Of the eCLIP sites, 401 of these overlapped with PAR-CLIP [36] and were designated as our positive set for the motif detection (**Fig. 1A**). As controls, we pooled all eCLIP peaks seen in only one replicate, failing the same enrichment cutoff and non-overlapping with PAR-CLIP, giving 200,664 sites. From these we randomly sampled 1,075 sequences to create a negative input set that was matching the length distribution of the positive set. Both sets were introduced to STREME (MEME Suite) for de novo motif discovery (motif lengths 4–16 nt), identifying eight enriched motifs (5–13 nt) in the positive set. These were supplemented by 3 FUS-recognized motifs from the *cat*RAPID *omics* v2 library [21], generating a final set of 11 motifs. 28,102 candidate aptamers were generated by extracting RNA sequences centered around these motif occurrences, with an additional 0–20 nucleotide 3′ extension. Secondary structure predictions were performed using RNAfold [22], and a custom scoring function was developed to prioritize aptamers displaying a perfect hairpin structure at the 5′ end followed by an unstructured 3′ extension. In parallel, the sequences were evaluated for the presence of at least one GGU motif in the single-stranded tail, a known enhancer of binding to FUS' Zn finger domain [36–38] (**Fig. 1C-D**)

Sequences were further prioritized based on their predicted binding propensity to $FUS_{269–454}$ using *cat*RAPID [21]. Aptamers with a secondary structure score above 0.9 and favorable *cat*RAPID interaction scores were retained for experimental validation. This yielded 83 high-confidence aptamers fulfilling all criteria, while low-scoring sequences were retained as negative controls. To complement the de novo approach, several aptamers were directly designed based on high catRAPID binding predictions, favoring structured hairpins with enriched sequence motifs, even when they were not directly extracted from the CLIP datasets.

Protein purification. Purification of RRM12-RGG-ZnfC ($FUS_{269-454}$) was carried out from kanamycin-resistant pET-Sumo expression vectors with a six-histidine tag (his-tag) and expressed in Rosetta (DE3) E. coli cells. Cells transformed with the plasmid were grown overnight at 37°C in LB medium containing 50 µg/ml kanamycin. Cell cultures were diluted 1:100 in fresh LB with 50 µg/ml kanamycin and grown until an optical density of ca. 0.6, before adding 0.5 mM IPTG and 0,1 mM $ZnCl_2$ to induce protein expression overnight at 18°C. Cells were collected by centrifugation at 4800 g for 20 min at 4°C, then resuspended in lysis buffer (20 mM potassium phosphate buffer pH 7.2, 150 mM KCl, 5 mM imidazol, 0,1 mM $ZnCl_2$, 1 mM DTT, 5 % v/v glycerol, 1 mg/ml lysozyme, cOmplete™ EDTA-free

Protease Inhibitor, 1 µg/ml DNase I). Cells were broken by probe-sonication and soluble proteins were recovered by centrifugation at 18.000 g for 1 h at 4°C. The total soluble protein mixture was loaded onto a column packed with nickel-coated resin and the 6xHis-SUMO-construct was eluted with high-salt phosphate buffer (20 mM potassium phosphate buffer pH 7.2, 150 mM KCl, 5 uM $ZnCl_2$, 1 mM DTT, 5 % v/v glycerol) with the addition of 300 mM imidazole. The eluate was dialyzed overnight at 4°C against low-salt phosphate buffer (20 mM potassium phosphate pH 7.2, 15 mM KCl, 5 uM $ZnCl_2$, 1 mM DTT, 5 % v/v glycerol) in the presence of the Tobacco Etch Virus (TEV) protease at a 1:20 protein construct:TEV molar ratio, to remove the 6×His-SUMO tag. Following a further affinity purification step with a nickel-coated resin, a HiTrap Heparin column was used to remove nucleic acids. The protein was then eluted with 1.5 M KCl and submitted to size-exclusion chromatography with a HiLoad 16/60 Superdex 75 column. Protein identity and purity were assessed by PAGE. The proteins were flash frozen and stored in low salt buffer (20 mM potassium phosphate pH 7.2, 15 mM KCl, 5 uM $ZnCl_2$, 1 mM DTT) at −80 °C.

Quantification of the aptamer binding constants. Biolayer interferometry experiments were acquired on an Octet Red instrument (ForteBio, Inc., Menlo Park, CA) operating at 25 °C. The binding assays were performed in 20 mM potassium phosphate buffer (pH 7.2) with 150 mM KCl and 0.05% Tween20. Streptavidin-coated biosensors were loaded with 1-5 µg/ml RNA aptamer modified with biotin on one of the ends and exposed to increasing protein concentrations varying from 10 nM to 10 µM, according to the strength of the binding. K_d values were estimated by fitting the response intensity (shift in the wavelength upon binding) as a function of the protein concentration, at the steady state. The assay was repeated at least 3 times, each time in triplicate.

Circular dichroism measurements. Circular dichroism spectra were recorded on a JASCO-1100 spectropolarimeter. The spectra were acquired in 0.1 mm path length quartz cuvettes under a constant N_2 flush at 4.0 L/min. All CD datasets were an average of ten scans. For the determination of the global secondary structure, the far-UV (200–320 nm) spectrum was recorded at 30°C in a low-salt phosphate buffer. Samples were gradually heated to 70°C (1°C/min) for the determination of the melting temperature (Tm). The ellipticity variation as expressed by the intensity of the band at 222 nm was followed as a function of temperature. The spectra were corrected for the buffer signal and expressed as mean residue molar ellipticity θ ($deg*cm^2*dmol^{-1}$).

Human post-mortem tissue staining. All postmortem data were generated from tissue samples collected and banked at the New York Brain Bank of Columbia University with consent obtained from the patient's next of kin, according to New York State law and the guidelines of the Department of Pathology of Columbia University and New York Presbyterian Hospital. In the State of New York, research involving autopsy material does not meet the regulatory definition of 'human subject research' and is not subject to institutional review board oversight. Formalin-fixed paraffin embedded (FFPE) tissue from BA4 was sectioned on a Leica microtome in 4 um sections onto a superfrost microscope slide. Sections were dried overnight at 40 °C and immunostaining was performed using the Novolink Polymer detection system with an anti-FUS antibody (BioTechne NB100-595) at a 1 in 500 dilution. DAB chromogen was used and counterstaining was performed with haematoxylin, according to standard operating procedures. Pretreatment was 30 minutes in a pressure cooker in pH6 buffered citric acid. FUS aptamer staining was performed as previously

published using identical method and concentration as TDP-43 aptamer staining [39]. Digital burden scoring was performed using the freely available QuPath software implementing superpixel analysis using the following code: setImageType('BRIGHTFIELD_H_DAB'); setColorDeconvolutionStains('{"Name" : "H-DAB default", "Stain 1" : "Hematoxylin", "Values 1" : "0.65111 0.70119 0.29049 ", "Stain 2" : "DAB", "Values 2" : "0.26917 0.56824 0.77759 ", "Background" : " 255 255 255 "}'); setPixelSizeMicrons(0.625,0.625); createSelectAllObject(true); selectAnnotations (); runPlugin('qupath.imagej.superpixels.DoGSuperpixelsPlugin', '{"downsampleFactor": 1.0, "sigmaMicrons": 3, "minThreshold": 10.0, "maxThreshold": 230.0, "noiseThreshold": 0.0}'); selectDetections(); runPlugin('qupath.lib.algorithms.IntensityFeaturesPlugin', '{"downsample": 1.0, "region": "ROI", "tileSizePixels": 200.0, "colorOD": false, "colorStain1": false, "colorStain2": true, "colorStain3": false, "colorRed": false, "colorGreen": false, "colorBlue": false, "colorHue": false, "colorSaturation": false, "colorBrightness": false, "doMean": true, "doStdDev": false, "doMinMax": false, "doMedian": false, "doHaralick": false, "haralickDistance": 1, "haralickBins": 32}'); setDetectionIntensityClassifications("ROI: 2.00 µm per pixel: DAB: Mean", 0.2689, 0.3007, 1). H-Score was calculated using the following weighted score (1x number of 1+ detections + 2x number of 2+ detections + 3x number of 3+ detections) normalised by the total number of pixels in the region of interest. Graphs were plotted using GraphPad Prism v10.0. Mann Whitney U test (with correction for multiple testing) was used to compare between cases and stains.

FIGURES and TABLES captions

Figure 1: Representation of the pipeline for FUS aptamers development. A: Boxplot showing the distribution of maximum fold enrichment for binding sites detected across one to four biological replicates. Sites overlapping with PAR-CLIP data (red) show higher fold enrichment compared to those not overlapping (grey); **B:** Receiver operating characteristic (ROC) curve evaluating the ability of *cat*RAPID to discriminate true FUS-binding sites from controls. The area under the curve (AUC) is 0.868, indicating strong predictive performance; **C:** Boxplot illustrating the *cat*RAPID interaction scores between candidate aptamers and different FUS domains. Sequences targeting the RRM and RRM.RGG.ZnF domains (red) show significantly higher predicted interaction propensity compared to controls (grey), whereas RGG-only domains show low interaction scores; **D:** Summary of final aptamer selection. Among the sequences with high secondary structure scores (>0.9), a subset containing the GGU motif in the 3' extension (red) was prioritized for experimental validation. An illustrative example of the predicted secondary structure of a selected aptamer is shown, highlighting the hairpin conformation and the location of GGU motifs.

Figure 2. Biolayer interferometry-derived binding curves of selected aptamers with $FUS_{269-454}$. A: TopScorerLong; **B**: Enriched; **C**: TopScorerShort; **D**: Enriched_catRAPID; **E**: Tuned 1278; **F**: Top_YNY14; **G**: Tuned_533; **H**: Apt1_catRAPID; **I**: U1SL34_catRAPID. The dots in the graphs indicate the shift of the light expressed in nm recorded by the instrument upon binding of the RNA aptamers with the protein at a given concentration. The response has been normalised to 1 for direct comparison between the different curves. The K_ds have been determined at steady state.

Figure 3. Circular dichroism analysis of selected aptamer candidates. A: TopScorerLong; **B:** Enriched; **C:** TopScorerShort; **D:** Neg24nt. Continue line:spectra recorded at 30 °C; dashed line: spectra recorded at 70 °C.

Figure 4. FUS aptamers recognise aggregates in tissue more selectively than antibodies. A. Representative photomicrographs implementing FUS antibody and aptamer staining using immunohistochemistry. Staining is performed on non-neurological control tissue from a person with an underlying p.P525L mutation who did not undergo ASO therapy, and a person with an underlying p.P525L mutation who did undergo ASO therapy. Antibody staining demonstrates normal nuclear localisation in controls and admixed normal and abnormal FUS staining including large cytoplasmic aggregates. Aptamer staining identifies only abnormal FUS staining patterns, with no staining visible in the non-neurological control and extensive pathological FUS staining in the untreated individual. No staining is seen in the ASO treated individual with antibody or aptamer. Aptamer staining additionally reveals large and extensive nuclear aggregation events that are not seen as frequently with the antibody staining. **B**. Superpixel digital masks used to quantify staining intensity. Blue pixels are baseline DAB (background) and yellow, orange, and red pixels represent 1, 2 and 3 standard deviations (respectively) above baseline average intensity. **C**. Superpixel H-score used to represent staining intensity demonstrates equivalent sensitivity (compared to antibody) and improved specificity (no detection in control and ASO treated individuals). Each point is the mean of three replicates from randomly generated regions of interest within layer 5 of the motor cortex of a single patient, error bars are SD. Image analysis was performed using QuPath.

Figure 5. Pathological features identified using FUS aptamer histological staining. A. Nuclear pathology is readily visualised using aptamer immunohistochemistry. Photomicrographs demonstrate nucleolar decoration, large nuclear aggregates and nuclear membrane pathology. B. Cytoplasmic features show similar morphologies to antibody staining, including large cytoplasmic aggregates, neuropil staining and glial cytoplasmic puncta within processes. Features demonstrated with white arrowheads.

Table 1. Predicted and experimentally validated affinities.

1. Kodavati, M. *et al.* FUS unveiled in mitochondrial DNA repair and targeted ligase-1 expression rescues repair-defects in FUS-linked motor neuron disease. *Nat. Commun.* **15**, 2156 (2024).
2. Errichelli, L. *et al.* FUS affects circular RNA expression in murine embryonic stem cell-derived motor neurons. *Nat. Commun.* **8**, 14741 (2017).
3. Reber, S. *et al.* Minor intron splicing is regulated by FUS and affected by ALS-associated FUS mutants. *EMBO J.* **35**, 1504–1521 (2016).
4. Naumann, M. *et al.* Phenotypes and malignancy risk of different FUS mutations in genetic amyotrophic lateral sclerosis. *Ann. Clin. Transl. Neurol.* **6**, 2384–2394 (2019).
5. López-Erauskin, J. *et al.* ALS/FTD-linked mutation in FUS suppresses intra-axonal protein synthesis and drives disease without nuclear loss-of-function of FUS. *Neuron* **100**, 816–830.e7 (2018).
6. Nolan, M., Talbot, K. & Ansorge, O. Pathogenesis of FUS-associated ALS and FTD: insights from rodent models. *Acta Neuropathol. Commun.* **4**, 99 (2016).
7. Naumann, M. *et al.* Impaired DNA damage response signaling by FUS-NLS mutations leads to neurodegeneration and FUS aggregate formation. *Nat. Commun.* **9**, 335 (2018).
8. Scekic-Zahirovic, J. *et al.* Cytoplasmic FUS triggers early behavioral alterations linked to cortical neuronal hyperactivity and inhibitory synaptic defects. *Nat. Commun.* **12**, 3028 (2021).
9. Mackenzie, I. R. A. & Neumann, M. Fused in Sarcoma Neuropathology in Neurodegenerative Disease. *Cold Spring Harb. Perspect. Med.* **7**, (2017).
10. Rezvykh, A. P. *et al.* Cytoplasmic aggregation of mutant FUS causes multistep RNA splicing perturbations in the course of motor neuron pathology. *Nucleic Acids Res.* **51**, 5810–5830 (2023).
11. Nomura, T. *et al.* Intranuclear aggregation of mutant FUS/TLS as a molecular pathomechanism of amyotrophic lateral sclerosis. *J. Biol. Chem.* **289**, 1192–1202 (2014).
12. Loughlin, F. E. *et al.* The Solution Structure of FUS Bound to RNA Reveals a Bipartite Mode of RNA Recognition with Both Sequence and Shape Specificity. *Mol. Cell* **73**, 490–504.e6 (2019).
13. Ni, S. *et al.* Recent progress in aptamer discoveries and modifications for therapeutic applications. *ACS Appl. Mater. Interfaces* **13**, 9500–9519 (2021).
14. Zacco, E. *et al.* Probing TDP-43 condensation using an in silico designed aptamer. *Nat. Commun.* **13**, 3306 (2022).
15. Ellington, A. D. & Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346**,

818–822 (1990).

16. Lerga, A. *et al.* Identification of an RNA binding specificity for the potential splicing factor TLS. *J. Biol. Chem.* **276**, 6807–6816 (2001).
17. Wang, X., Schwartz, J. C. & Cech, T. R. Nucleic acid-binding specificity of human FUS protein. *Nucleic Acids Res.* **43**, 7535–7543 (2015).
18. Imperatore, J. A., McAninch, D. S., Valdez-Sinon, A. N., Bassell, G. J. & Mihailescu, M. R. FUS Recognizes G Quadruplex Structures Within Neuronal mRNAs. *Front Mol Biosci* **7**, 6 (2020).
19. Jutzi, D. *et al.* Aberrant interaction of FUS with the U1 snRNA provides a molecular mechanism of FUS induced amyotrophic lateral sclerosis. *Nat. Commun.* **11**, 6341 (2020).
20. Rupert, J., Monti, M., Zacco, E. & Tartaglia, G. G. RNA sequestration driven by amyloid formation: the alpha synuclein case. *Nucleic Acids Res.* **51**, 11466–11478 (2023).
21. Armaos, A., Colantoni, A., Proietti, G., Rupert, J. & Tartaglia, G. G. catRAPID omics v2.0: going deeper and wider in the prediction of protein–RNA interactions. *Nucleic Acids Res.* **49**, W72–W79 (2021).
22. Lorenz, R. *et al.* ViennaRNA Package 2.0. *Algorithms Mol. Biol.* **6**, 26 (2011).
23. Loughlin, F. E. & Allain, F. H.-T. Solution structure of FUS-RRM bound to stem-loop RNA. Preprint at https://doi.org/10.2210/pdb6gbm/pdb (2019).
24. Loughlin, F. E. & Allain, F. H.-T. Solution structure of FUS-ZnF bound to UGGUG. Preprint at https://doi.org/10.2210/pdb6g99/pdb (2019).
25. Schwartz, J. C. *et al.* FUS binds the CTD of RNA polymerase II and regulates its phosphorylation at Ser2. *Genes Dev.* **26**, 2690–2695 (2012).
26. Luige, J., Armaos, A., Tartaglia, G. G. & Ørom, U. A. V. Predicting nuclear G-quadruplex RNA-binding proteins with roles in transcription and phase separation. *Nat. Commun.* **15**, 2585 (2024).
27. Kypr, J., Kejnovská, I., Renciuk, D. & Vorlícková, M. Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res.* **37**, 1713–1725 (2009).
28. Korobeynikov, V. A., Lyashchenko, A. K., Blanco-Redondo, B., Jafar-Nejad, P. & Shneider, N. A. Antisense oligonucleotide silencing of FUS expression as a therapeutic approach in amyotrophic lateral sclerosis. *Nat. Med.* **28**, 104–116 (2022).
29. Spence, H. *et al.* RNA aptamer reveals nuclear TDP-43 pathology is an early aggregation event that coincides with STMN-2 cryptic splicing and precedes clinical manifestation in ALS. *Acta Neuropathol.* **147**,

50 (2024).

30. Bankhead, P. *et al.* QuPath: Open source software for digital pathology image analysis. *Sci. Rep.* **7**, 16878 (2017).

31. Zacco, E. *et al.* RNA as a key factor in driving or preventing self-assembly of the TAR DNA-binding protein 43. *J. Mol. Biol.* **431**, 1671–1688 (2019).

32. Shelkovnikova, T. A., Robinson, H. K., Southcombe, J. A., Ninkina, N. & Buchman, V. L. Multistep process of FUS aggregation in the cell cytoplasm involves RNA-dependent and RNA-independent mechanisms. *Hum. Mol. Genet.* **23**, 5211–5226 (2014).

33. Germer, K., Leonard, M. & Zhang, X. RNA aptamers and their therapeutic and diagnostic applications. *Int. J. Biochem. Mol. Biol.* **4**, 27–40 (2013).

34. Mariani, D. *et al.* ALS-associated FUS mutation reshapes the RNA and protein composition of stress granules. *Nucleic Acids Res.* **52**, 13269–13289 (2024).

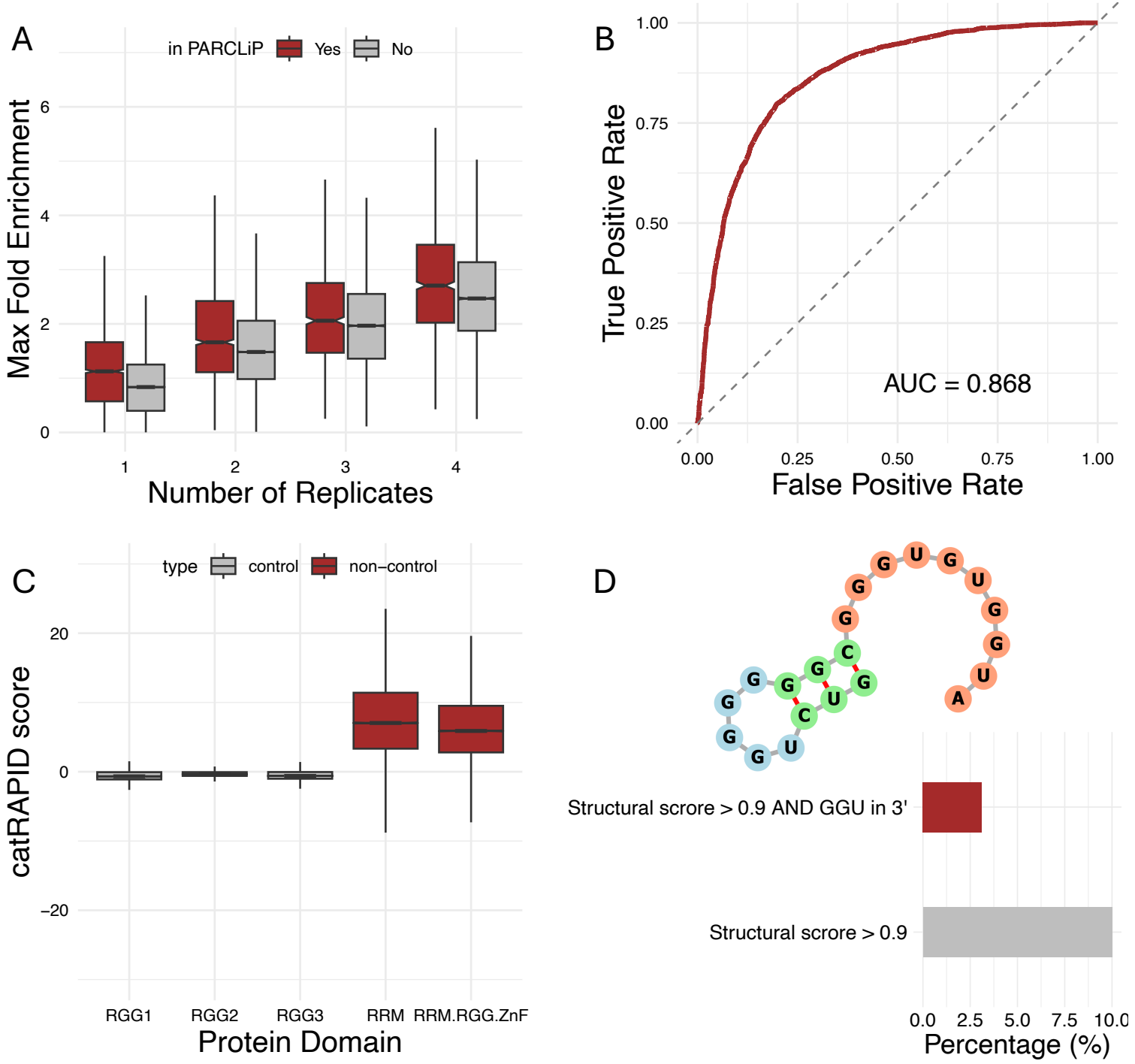
35. Van Nostrand, E. L. *et al.* A large-scale binding and functional map of human RNA-binding proteins. *Nature* **583**, 711–719 (2020).

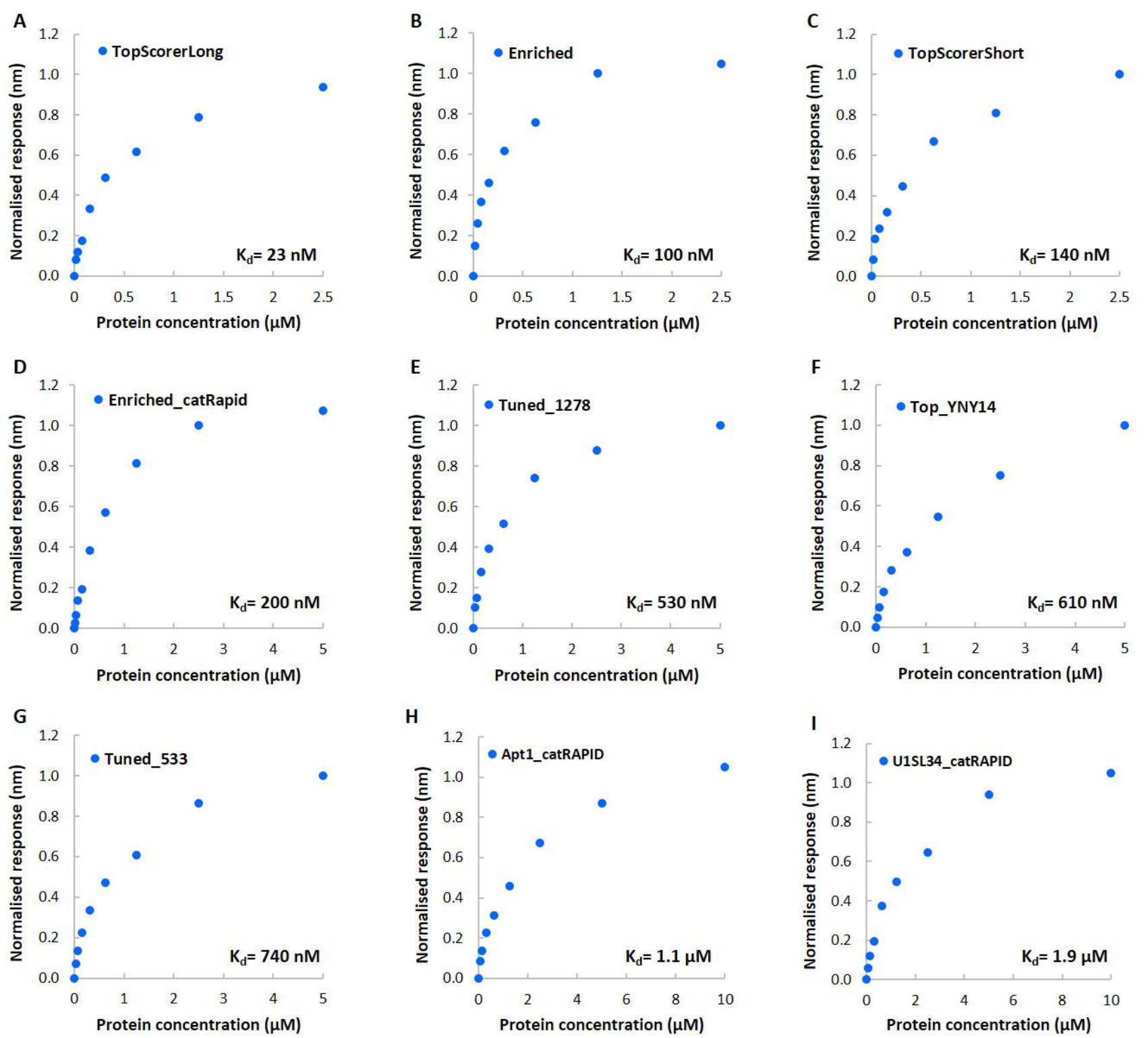
36. De Santis, R. *et al.* Mutant FUS and ELAVL4 (HuD) Aberrant Crosstalk in Amyotrophic Lateral Sclerosis. *Cell Rep.* **27**, 3818–3831.e5 (2019).

37. Bellucci, M., Agostini, F., Masin, M. & Tartaglia, G. G. Predicting protein associations with long noncoding RNAs. *Nat. Methods* **8**, 444–445 (2011).

38. Dominguez, D. *et al.* Sequence, structure, and context preferences of human RNA binding proteins. *Mol. Cell* **70**, 854–867.e9 (2018).

39. M Waldron, F., Spence, H. & Gregory, J. TDP-43 RNA aptamer staining to detect pathological TDP-43 in FFPE human tissue, as described in Spence and Waldron et al., 2024 (Acta Neuropathologica): A SOP and tick-sheet. v2. v2. (2024) doi:10.17504/protocols.io.eq2lyjo4mlx9/v2.


A
in PARCLiP
Yes
No
Max Fold Enrichment
0
2
4
6
1
2
3
4
Number of Replicates
B
True Positive Rate
False Positive Rate
0.00
0.25
0.50
0.75
1.00
AUC = 0.868
C
type
control
non−control
catRAPID score
20
0
−20
RGG1
RGG2
RGG3
RRM
RRM.RGG.ZnF
Protein Domain
D
Structural scrore > 0.9 AND GGU in 3'
Structural scrore > 0.9
0.0
2.5
5.0
7.5
10.0
Percentage (%)


A
TopScorerLong
K_d= 23 nM
B
Enriched
K_d= 100 nM
C
TopScorerShort
K_d= 140 nM
D
Enriched_catRapid
K_d= 200 nM
E
Tuned_1278
K_d= 530 nM
F
Top_YNY14
K_d= 610 nM
G
Tuned_533
K_d= 740 nM
H
Apt1_catRAPID
K_d= 1.1 µM
I
U1SL34_catRAPID
K_d= 1.9 µM
Normalised response (nm)
Protein concentration (µM)

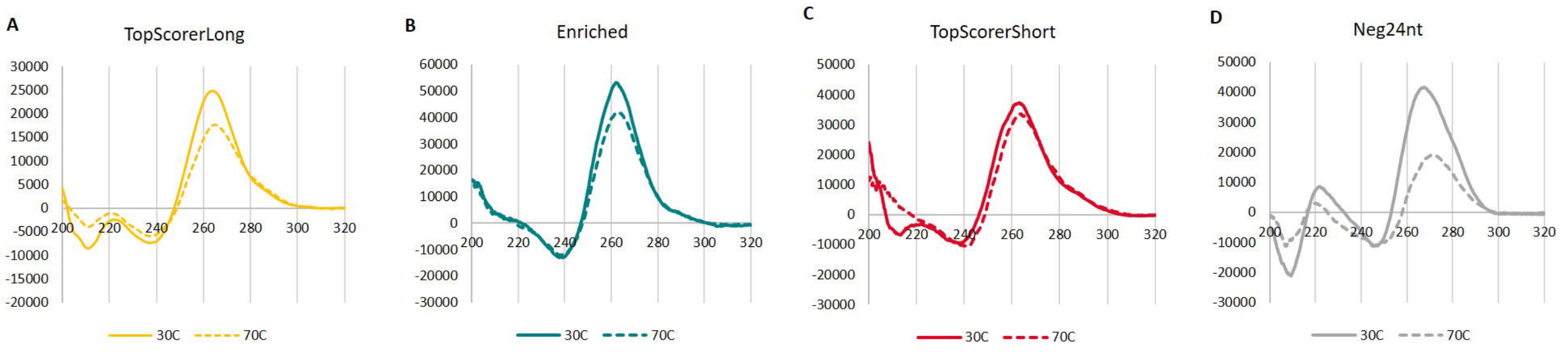

A
TopScorerLong
30000
25000
20000
15000
10000
5000
0
-5000
-10000
-15000
-20000
200
220
240
260
280
300
320
30C
70C
B
Enriched
60000
50000
40000
30000
20000
10000
0
-10000
-20000
-30000
200
220
240
260
280
300
320
30C
70C
C
TopScorerShort
50000
40000
30000
20000
10000
0
-10000
-20000
-30000
200
220
240
260
280
300
320
30C
70C
D
Neg24nt
50000
40000
30000
20000
10000
0
-10000
-20000
-30000
200
220
240
260
280
300
320
30C
70C

Figure 4.

A.

Control

FUS p.P525L untreated

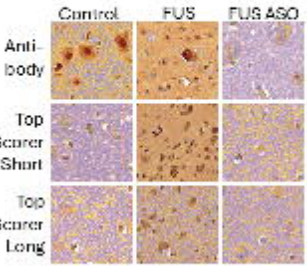
FUS p.P525L ASO treated

Antibody

TopScorer Short

TopScorer Long

B. Superpixel digital masks



C. Staining intensity

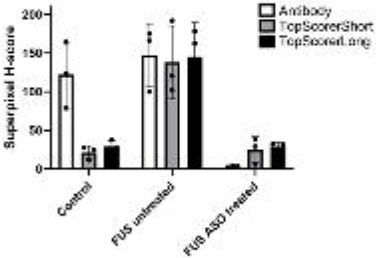
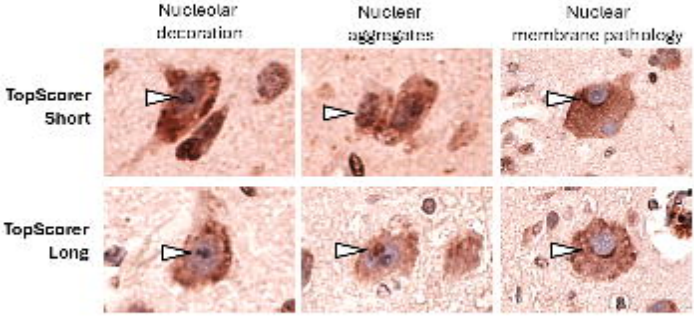
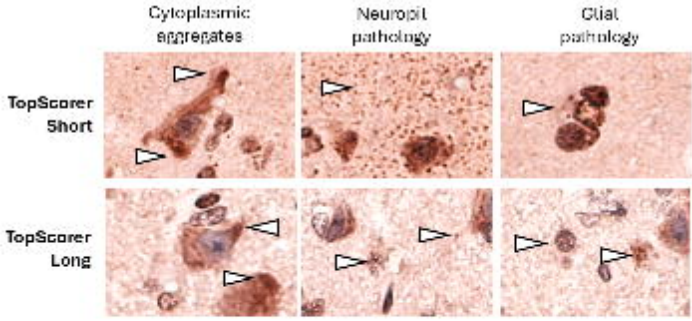


Figure 5.

A. Nuclear pathology



B. Cytoplasmic pathology



NAME	NORMALISED RRM-RGG-ZnF SCORE	K_d (μM)
BEST BINDERS		
Top scorer_long	0.84	0.023±0.002
Enriched	0.89	0.10±0.03
Top scorer_short	0.95	0.140±0.018
Enriched catRAPID	0.75	0.20±0.07
1278	0.69	0.53±0.095
YNY14	0.72	0.610±0.18
533	0.51	0.74±0.013
INTERMEDIATE BINDERS		
Apt1 catRAPID	0.5	1.10±0.16
U1SL34_catRAPID	0.5	1.90±0.40