SUPPLEMENTARY INFORMATION

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

Elsa Zacco*#[1], Martina Gilodi*[1], Abigail Dos Santos[2], Alexandros Armaos[1], Rachel James[2], Francesco Di Palma[1], Fergal M. Waldron[3], Marco Scotto[1],

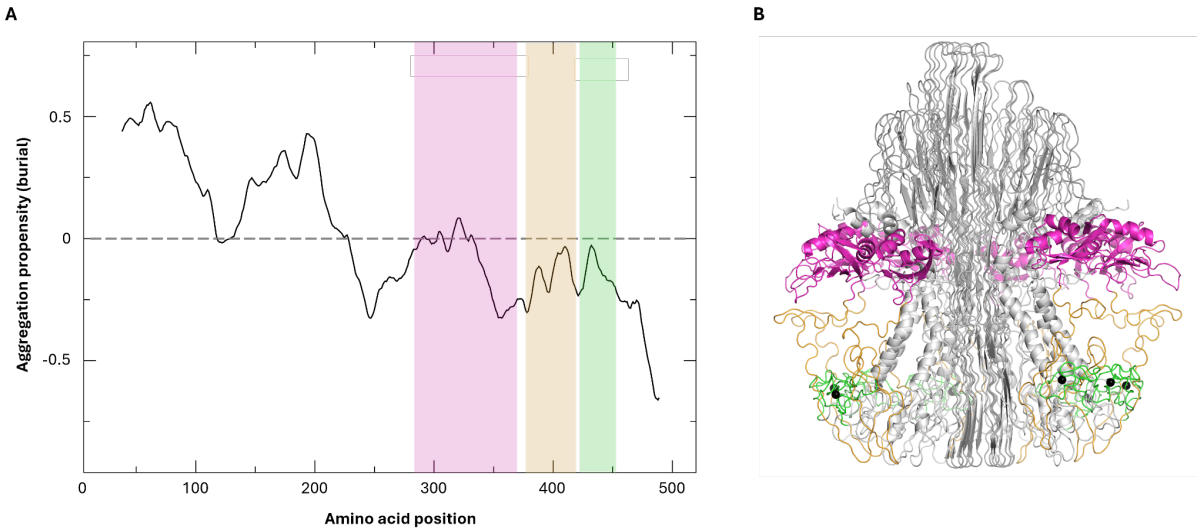
Jakob Rupert[1], Vlad Korobeynikov[4], Neil A. Shneider[4],

Mathew H. Horrocks[2], Jenna M. Gregory#[3] & Gian Gaetano Tartaglia#[1]

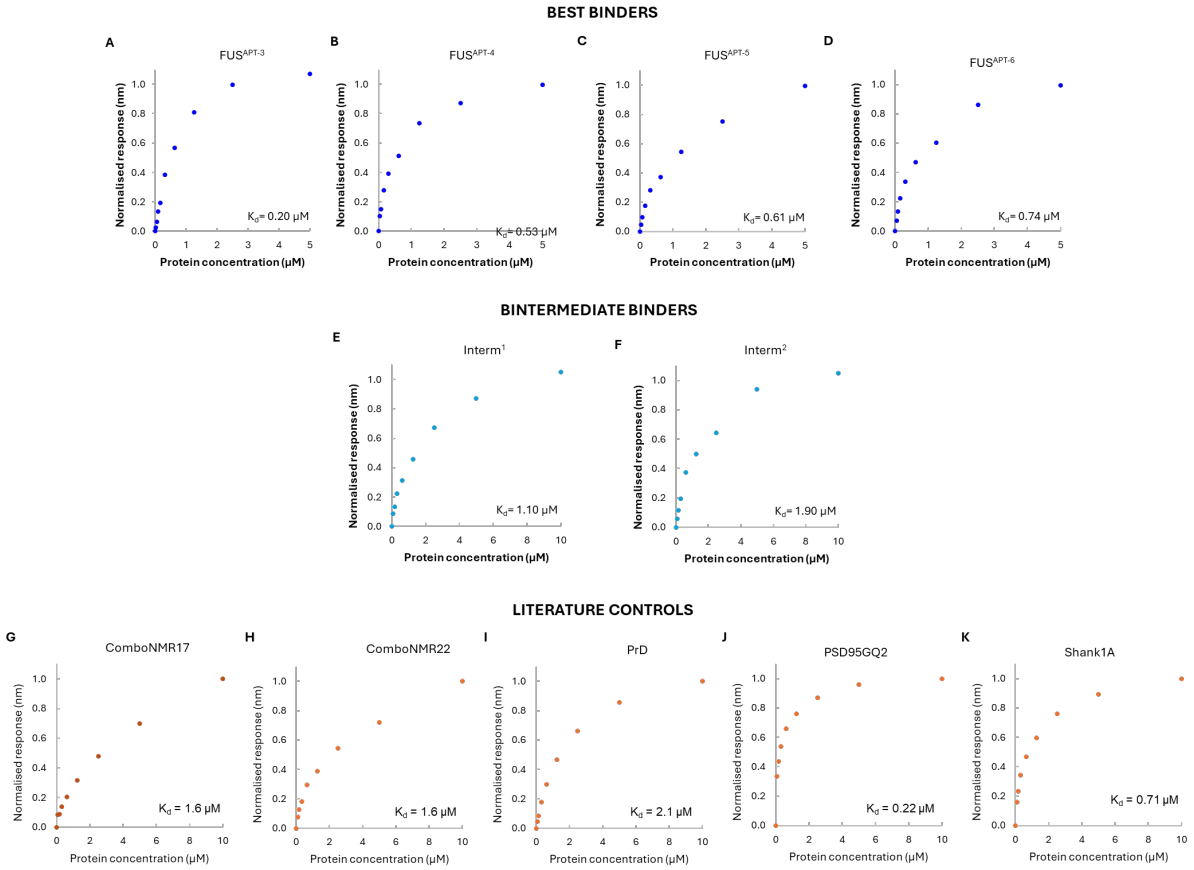
1. RNA Systems Biology, Istituto Italiano di Tecnologia, Genoa, Via Enrico Melen, 83, 16152, Genoa, Italy
2. School of Chemistry, University of Edinburgh, Joseph Black Building, David Brewster Road, Edinburgh, EH9 3FJ, UK
3. Institute of Medical Sciences, University of Aberdeen, Foresterhill, Aberdeen AB25 2ZD, United Kingdom
4. Department of Neurology, Columbia University Irving Medical Center, 710 West 168th Street, New York, NY 10032, USA

*These authors contributed equally

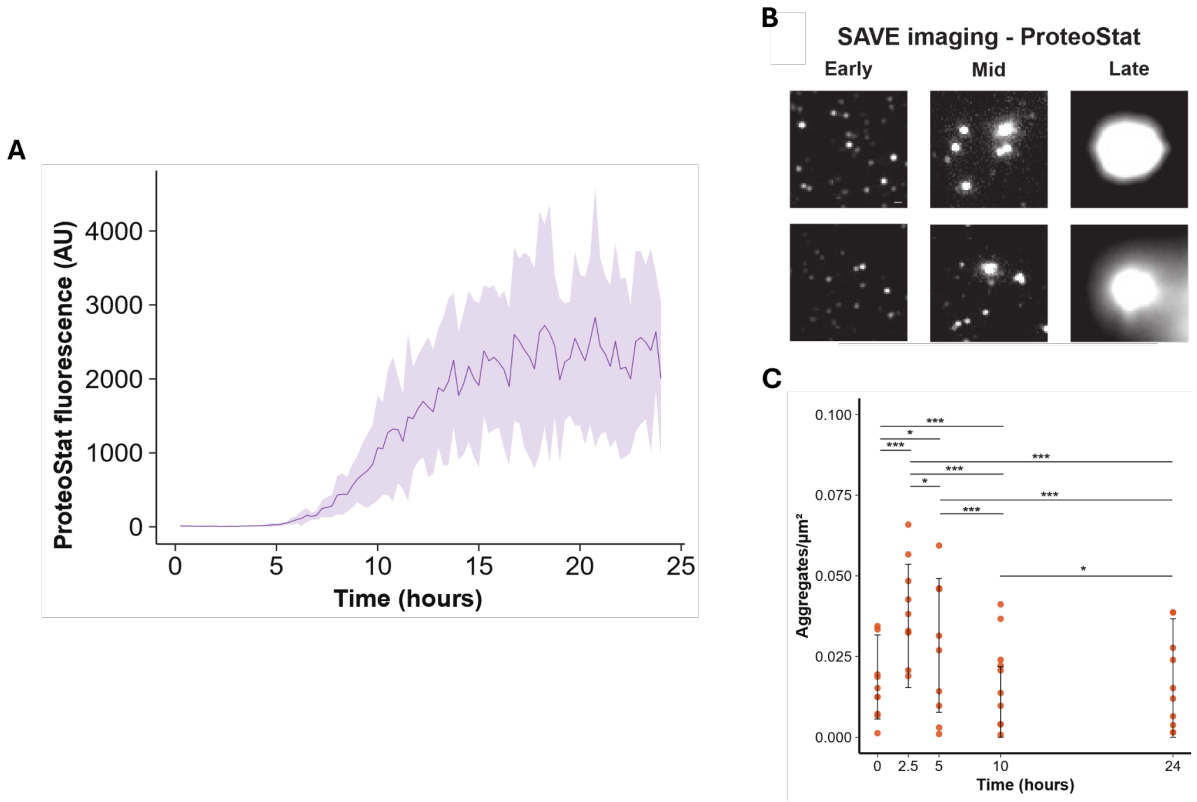
correspondence to: elsa.zacco@iit.it , jenna.gregory@abdn.ac.uk and gian.tartaglia@iit.it



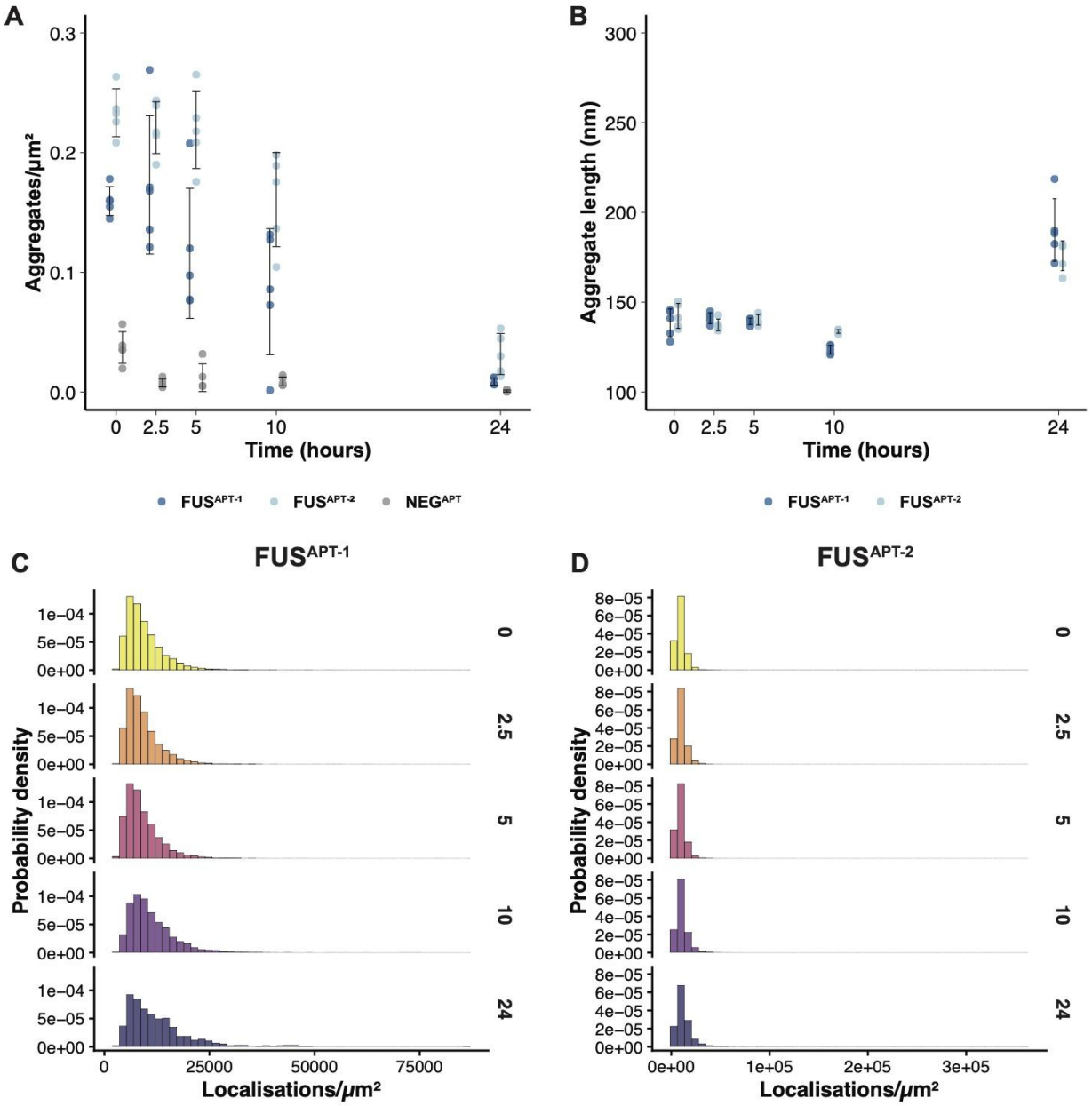
Suppl. Fig. 1: The RNA-binding regions of FUS are predicted to be external to the aggregation core and thus available for aptamer binding. A: Zyggregator calculation [1,2] of the residues most prone to aggregate and, therefore, to be buried in the core. The protein region starting from amino acid 210 to the C-terminus is predicted to be exposed to the solvent and ready for interaction. **B**: AlphaFold reconstruction [3] of an aggregate formed by eight copies of FUS, showing that indeed the RRM (magenta), RGG2 (yellow) and zinc finger (green) are exposed to the surface of the aggregate and available to interact with RNA binding partners.



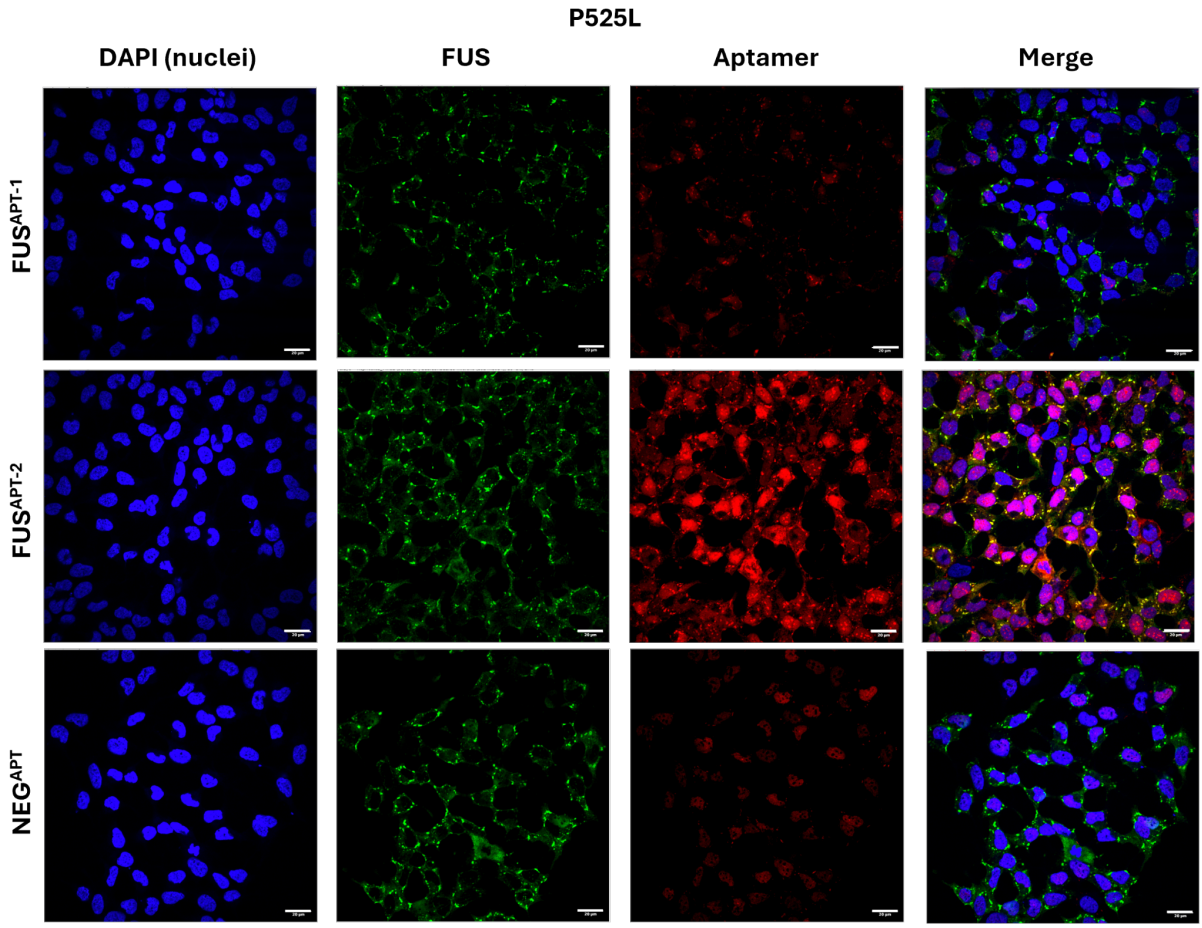
Suppl. Fig. 2: Biolayer interferometry-derived binding curves of selected aptamers and literature controls with FUS. Each graph reports the binding response as a function of the protein concentration and the K_d values calculated at steady state.



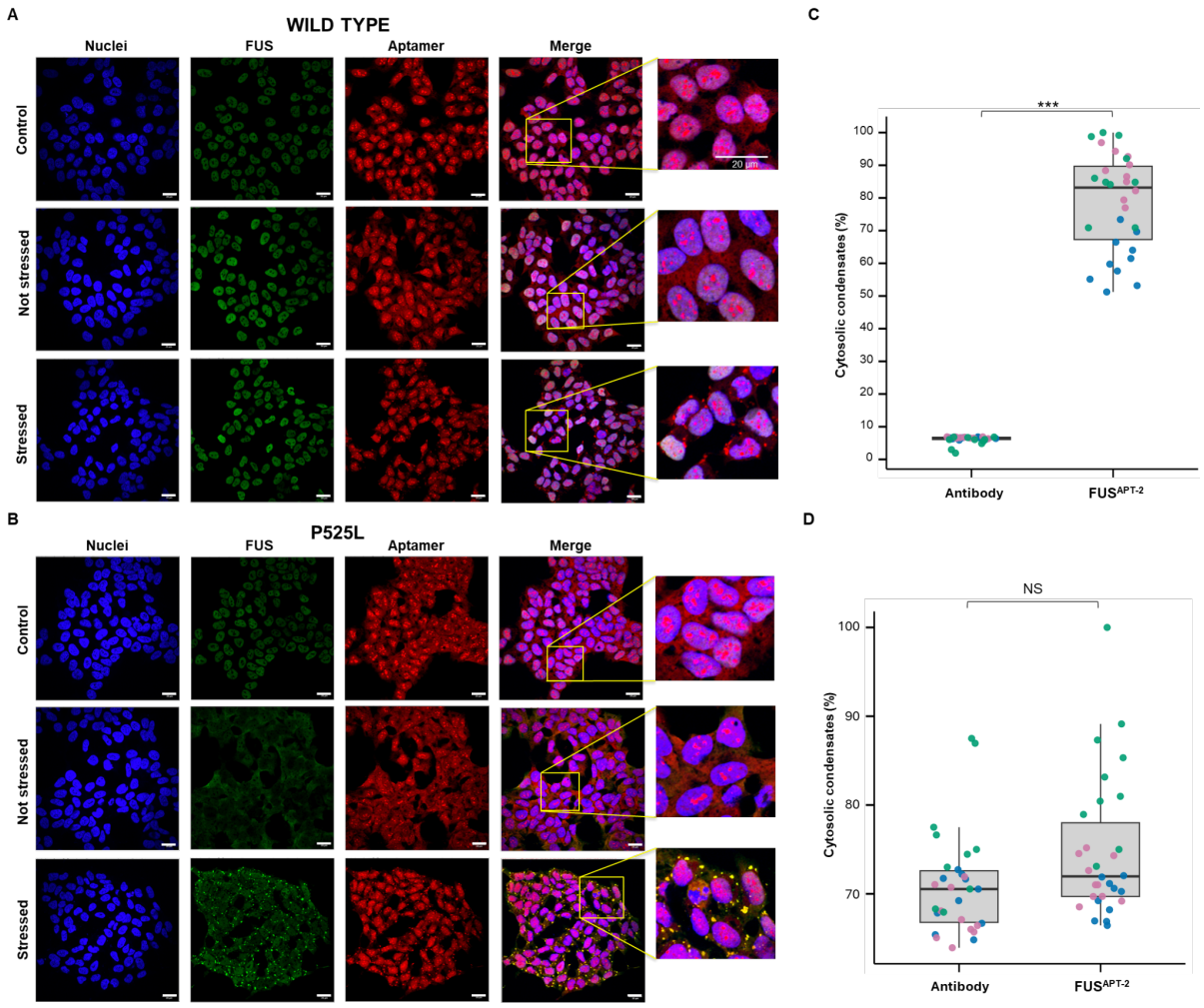
Suppl. Fig. 3: ProteoStat fluorescence demonstrates that aggregates are formed over the 24-hour period of aggregation. A. ProteoStat fluorescence as measured across 5 wells over time, showing an increase in average fluorescence (the solid line represents the mean, and the shaded region, the standard deviation). **B**. Representative diffraction-limited SAVE (Single-Aggregate Visualisation by Enhancement) images and **C**. quantification of the imaged aggregates formed over time. Aggregates are shown to decrease in number as they increase in size. In B, the 0 hour baseline is the “Early” timepoint, 5 hours is “Mid”, and 24 hours is “Late”. Scale bars = 500 nm.



Suppl. Fig. 4: Numbers, lengths, and localisations per unit area as detected by super-resolution imaging with the aptamers. Quantification of **A**: numbers of aggregates detected by all 3 aptamers and **B**: lengths of the aggregates detected by the detection aptamers used in this study. Each point is the average for one well of the aggregation, for a total of 5 technical replicates per aptamer. Statistical evaluation for aggregate lengths (Kruskal-Wallis test, with Dunn post-hoc and Benjamini-Hochberg correction) resulted in multiple significant comparison, reported in **Suppl. Table 2**. Normalised histograms of the probability density distribution of localisations per unit area as detected by **C**: FUS^{APT-1} or **D**: FUS^{APT-2}.



Suppl. Fig. 5: **Overview of immunofluorescence images of FUS P525L–expressing cells subjected to sodium arsenite stress and stained with an anti-FUS antibody together with one of the tested aptamers.** Top row: FUS^{APT-1}; middle row: FUS^{APT-2}; bottom row: NEG^{APT}. Nuclei are labeled with DAPI (blue); the anti-FUS primary antibody is detected using an Atto488-conjugated secondary antibody (green); aptamers are labeled with an Atto590 fluorophore (red). Scale bar= 20 µm.



Suppl. Fig. 6: Immunofluorescence overview of SK-N-BE cells permeabilized with Triton X-100 and stained with an anti-FUS antibody together with FUS^{APT-2}. A: Cells expressing wild-type FUS. **B**: Cells expressing the FUS P525L variant. For each condition, cells are shown untreated ('Control', top rows), following doxycycline-induced FUS expression ('Not stressed', middle rows), and after exposure to sodium arsenite ('Stressed', bottom rows). Nuclei are labeled with DAPI (blue); the anti-FUS primary antibody is detected using an Atto488-conjugated secondary antibody (green); aptamers are labeled with an Atto590 fluorophore (red). Scale bar = 20 µm. **C–D**: Quantification of the percentage of cytosolic condensates detected by the antibody and by FUS^{APT-2} in cells expressing wild-type FUS (**C**) or FUS P525L (**D**). Points represent data from triplicate experiments (blue: replicate 1; pink: replicate 2; green: replicate 3). Statistical significance was assessed using a one-tailed *t*-test: *** $p = 0.0001$.

NUM.	SEQUENCE
Seq. 1	CAGCCCGGGGGGCGGCCCGCGGCGGUAGGAGCU
Seq. 2	CAGCCCGGGGGGCGGCCCGCGGCGGUAGG
Seq. 3	GCCUCGCGGGCGGGGGGGGUGGGGG
Seq. 4	GGGCCCCGGGCGGGGCGGGGGUACG
Seq. 5	GGGCGCUGGGGGGGAGCCUAGUGGCUGGUGUUUGGGC
Seq. 6	GUGCCGGGGGAGGGACGGCGCGCGGGGUU
Seq. 7	CGCCUGAGGGAGGGCCUGGGGCAGGGCACAAGGGGUUGAUC
Seq. 8	CCAGGCCGGGGGGGCCGGAGUGCAGGUGGGUGU
Seq. 9	CCAGGCCGGGGGGGCCGGAGUGCAGGUGG
Seq. 10	GGGCGCUGGGGGGGAGCCUAGUGGCUGGUGUUU
Seq. 11	GCCUGGCGGGCGGGGCCGGUG
Seq. 12	GGGCCCCGGGCGGGGCGGGGGUACGUGGA
Seq. 13	GGCCGGAGGGCGGGUGGCGCGGGAG
Seq. 14	GCCUCGCGGGCGGGGGGGGUG
Seq. 15	GCGGGCUGGGGGCGCCCGCCGUGGUGGAC
Seq. 16	AGGGGCCGGGAGGGAGUCGGCUGAGGUGG
Seq. 17	GGGCGCUGGGGGGGAGCCUAGUGGCUGGU
Seq. 18	CAGGCAGGGGAGGGGGCUUGGACUGAAGGUGAG
Seq. 19	GCGGGCUGGGGGCGCCCGCCGUGGU
Seq. 20	CGCUGUGGGGCGGGGGUCGCG
Seq. 21	GGCUCCCGGGCGGGCAGGGUAGGGC
Seq. 22	CCUGAAAGGGAGGGAAAGGGGAGGGUAGCGGGAGGGU
Seq. 23	GGGCUGUGGGAGGGGCAAGCUACGGCCUGGGCUAGGU
Seq. 24	GGCCGGAGGGCGGGUGGCGCG
Seq. 25	CCCUGGGGGGAGGGUGGUGGC
Seq. 26	AGCCCACGGGCGGGGGUGCGG
Seq. 27	CGCCUGAGGGAGGGCCUGGGGCAGGGCACAAGGGGUU
Seq. 28	GGCAGAUUACAAUUCUAUUUGCCUGUGGGGAGGGGGCGUGGUG
Seq. 29	GAGUAGCGGGAGGGAAGUUGUAGUACGGGUGGG
Seq. 30	CUGCGGGGGGCGGGGUGAAGC
Seq. 31	GGCCGUGGGGCGGGGGGUCUU
Seq. 32	GGCCUGAGGGAGGGACCAUGGUGGC
Seq. 33	GGCAGAUUACAAUUCUAUUUGCCUGGGAGGAGGAAGAGGAGGAUGGUG
Seq. 34	GCGCCUCGGGGGCGCGGGUAG
Seq. 35	UGUCAAUGGGAGGGGCUGUUCCAUGCAGGGUGG
Seq. 36	GAGUAGCGGGAGGGAAGUUGUAGUACGGGUGGGGAGA
Seq. 37	CUGCCUGGGGCGGGAGGGUCC
Seq. 38	GUCCGAGGGGCGGGGGUGCUC
Seq. 39	UCUCUAGGGGAGGGGCAUGGGUAAGGGUG
Seq. 40	GUCUGGGGGGCGGGUGUGGUA
Seq. 41	CUGCGGGGGGCGGGGUGAAGCUGUG
Seq. 42	CUGCGGGGGGCGGGGUGAAGCUGUGAUGGACAU
Seq. 43	GCCAGGAGGGCGGGAGGUGGC
Seq. 44	GGGGACGGGGGGCGCUCACAGAAGGUAGG
Seq. 45	GGCUCCCGGGCGGGCAGGGUA
Seq. 46	GCCUCUGGGGAGGGCGGUUCU
Seq. 47	GCGGCCCAGCAGUCAGUGCCCGCGGAGCAGGUA

Seq. 48 GGGGAACUGGUGUGGAGGUGAGGGG
Seq. 49 CUCCUCAGGGCGGGGGUGACAGUGGGUGU
Seq. 50 GGCAGAUUACAAUUCUAUUUGCCUAGGGAGAGGGAGGGAGCAGAGUGGUG
Seq. 51 UCCUUAGGGGAGGGGUCCAGGUAUGCAGG
Seq. 52 GGCAGAUUACAAUUCUAUUUGCCUUGGGAUGGGGAGGGUGGUG
Seq. 53 CAGCCAGGGGAGGGCUGAAGCUAGAAUGGUCAAGGUG
Seq. 54 CUCUUUGGGGAGGGGGUAGGU
Seq. 55 GGCAGAUUACAAUUCUAUUUGCCUGAACAGAGGGAGGGAAUGGUG
Seq. 56 UUCUGGAGGGAGGGGGUUGUG
Seq. 57 CUGCGGGGGGCGGGGUGAAGCUGUGAUGG
Seq. 58 GCCUCUGGGGAGGGCGGUUCUGUGU
Seq. 59 UCCUUAGGGGAGGGGUCCAGGUAUG
Seq. 60 CCUGAAAGGGAGGGAAAGGGGAGGGUAGCGGGA
Seq. 61 CUCUUUGGGGAGGGGGUAGGUUGGA
Seq. 62 GCCUCGCGGGCGGGGGGGGUGGGGG
Seq. 63 GGGCCCCGGGCGGGGCGGGGGUACG
Seq. 64 GUGCCGGGGGAGGGACGGCGCGCGGGGUU
Seq. 65 CGCCUGAGGGAGGGCCUGGGGCAGGGCACAAGGGGUUGAUC
Seq. 66 CCAGGCCGGGGGGGCCGGAGUGCAGGUGGGUGU
Seq. 67 CCAGGCCGGGGGGGCCGGAGUGCAGGUGG
Seq. 68 GGGCCCCGGGCGGGGCGGGGGUACGUGGA
Seq. 69 GCCUGGCGGGCGGGGCCGGUG
Seq. 70 GCGGGCUGGGGGCGCCCGCCGUGGUGGAC
Seq. 71 GCCUCGCGGGCGGGGGGGGUG
Seq. 72 AGGGGCCGGGAGGGAGUCGGCUGAGGUGG
Seq. 73 GCGGGCUGGGGGCGCCCGCCGUGGU
Seq. 74 CAGGCAGGGGAGGGGGCUUGGACUGAAGGUGAG
Seq. 75 CGCUGUGGGGCGGGGGUCGCG
Seq. 76 GGGCUGUGGGAGGGGCAAGCUACGGCCUGGGCUAGGU
Seq. 77 CCUGAAAGGGAGGGAAAGGGGAGGGUAGCGGGAGGGU
Seq. 78 CGCCUGAGGGAGGGCCUGGGGCAGGGCACAAGGGGUU
Seq. 79 CUGCGGGGGGCGGGGUGAAGC
Seq. 80 GCGCCUCGGGGGCGCGGGUAG
Seq. 81 CUGCCUGGGGCGGGAGGGUCC
Seq. 82 UGUCAAUGGGAGGGGCUGUUCCAUGCAGGGUGG
Seq. 83 UCUCUAGGGGAGGGGCAUGGGUAAGGGUG
Seq. 84 CUGCGGGGGGCGGGGUGAAGCUGUG
Seq. 85 CUGCGGGGGGCGGGGUGAAGCUGUGAUGGACAU
Seq. 86 GCCUCUGGGGAGGGCGGUUCU
Seq. 87 CUCCUCAGGGCGGGGGUGACAGUGGGUGU
Seq. 88 UCCUUAGGGGAGGGGUCCAGGUAUGCAGG
Seq. 89 CUCUUUGGGGAGGGGGUAGGU
Seq. 90 CUGCGGGGGGCGGGGUGAAGCUGUGAUGG
Seq. 91 CCUGAAAGGGAGGGAAAGGGGAGGGUAGCGGGA
Seq. 92 GCCUCUGGGGAGGGCGGUUCUGUGU
Seq. 93 UUCUGGAGGGAGGGGGUUGUG
Seq. 94 UCCUUAGGGGAGGGGUCCAGGUAUG
Seq. 95 CUCUUUGGGGAGGGGGUAGGUUGGA

Seq. 96 GCCUCUGGGGAGGGCGGUUCUGUGUCUUG
Seq. 97 CCCUCUUGGGGGCGUGGUUUUGGGG
Seq. 98 CCUGAAAGGGAGGGAAAGGGGAGGGUAGC
Seq. 99 CUCCUGAGGGAGGGAAGGUAC
Seq. 100 UCUCUAGGGGAGGGGCAUGGGUAAG
Seq. 101 AAGUUCCAGAGGCCUAGACUUGGAACUGGUGUC
Seq. 102 UCUGAAAGGGAGGGUAGGGUU
Seq. 103 CCCUCUUGGGGGCGUGGUUUU
Seq. 104 UCUGAAAGGGAGGGUAGGGUUUAUG
Seq. 105 UCCCAAUGGGAGGGCCAGGUC
Seq. 106 UCUGAAAGGGAGGGUAGGGUUUAUGAGGC

Suppl. Table 1: List of aptamers generated in this study.

Suppl. Table 2: Summary of statistical analyses for super-resolution imaging experiments (Fig. 4 and Suppl. Figs. 3–4).

References

1. Tartaglia, G. G. *et al.* Prediction of aggregation-prone regions in structured proteins. *J. Mol. Biol.* **380**, 425–436 (2008).
2. 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).
3. Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).