Supplementary Note

RGG domains form weak interactions with FUS SYGQ LC – dispersed phase NMR

We sought to understand the molecular underpinnings of SYGQ LC and RGG domain interactions that contribute to their co-phase separation and hence full-length FUS phase separation. To structurally characterize the direct contacts between the RGG domains and the SYGQ LC that give rise to LLPS, we used NMR experiments (^{1}H-^{15}N HSQC) in the dispersed phase that monitor the change in chemical environment due to binding/interactions with a residue-by-residue resolution at each non-proline positions[54,85–87]. To determine at what sites the RGG domains interact with the SYGQ LC, we performed the experiment on mixtures of ^{15}N-labeled SYGQ LC (30 μM) and natural abundance MBP-tagged RGG domains. We retained the MBP-tag on the RGG domains to prevent phase separation (Fig. 1) from obscuring measurement of chemical shifts and signal intensity changes arising from direct interactions[41]. We quantified the ^{15}N chemical shift perturbations and intensity differences as we have previously used them to measure weak, transient interactions[54,85–87]. The addition of increasing concentrations of MBP-tagged RGG domains (30, 90, 300 μM) produced chemical shift perturbations, intensity differences and elevated transverse relaxation rates above the MBP control for residue positions across the SYGQ LC domain (Extended Data Fig. 2A,C). These data suggest that the RGG domains interact with the SYGQ LC in a similar manner as chemical shifts and intensity differences were localized to the same regions. Particularly, the regions surrounding residues 40 and 100 have attenuated signal intensity even in the presence of MBP alone, suggesting that these regions may be prone to non-specific interactions. Indeed, these regions also showed attenuation in the presence of karyopherin β2[21]. To characterize the interaction between the two components on the RGG domains, we obtained the spectra of ^{15}N-labeled RGG1, RGG2, or RGG3 in the presence of varying concentrations of natural abundance MBP-tagged SYGQ LC. Due to extensive peak overlap, we were unable to resolve the resonances for all the RGG motifs within RGG1 and RGG2 (Extended Data Fig. 3A-C). Importantly, all RGG domains appear intrinsically disordered based on the narrow distribution of resonances within the ^{1}H dimension. We observed ^{15}N chemical shift perturbations and intensity differences for residues spread across all three RGG

domains in the presence of MBP-tagged SYGQ LC (Extended Data Fig. 2B), including in the glycine-rich region of RGG1 lacking arginine residues. The chemical shift perturbations were higher in the presence of MBP-SYGQ LC than an MBP control (Extended Data Fig. 2D). Because of the interaction between MBP and both FUS SYGQ LC and RGG domains in the control data, specific regions of interaction between the SYGQ LC and RGG domains cannot be determined from these data with confidence and may only reflect a slight increase in affinity compared to the MBP control. Additionally, given that MBP forms some transient contacts with FUS domains, the presence of MBP may change the distribution of possible interactions between FUS SYGQ LC and RGGs. To test if the RGG domain fused to MBP indeed contributes to direct SYGQ LC interaction, we altered the sequence of MBP-RGG3 and observed if these changes affect the chemical shifts and intensity attenuation of FUS SYGQ LC caused by addition of MBP-RGG3. We conducted experiments on mixtures of ^{15}N-SYGQ LC with three forms of MBP-RGG3: RGG wild-type and two RGG3 variants where the arginine residues are substituted by lysines or serines (R10xK and R10xS, respectively). These lysine or serine variants should disrupt specific contacts made by the arginine residues, and hence addition of these variants should show smaller perturbations than WT. We observed the largest average ^{15}N chemical shift and greatest intensity attenuation for MBP-RGG3 as compared to MBP-RGG3 R10xK or R10xS, suggesting that chemical shifts and intensity differences are directly reporting on interactions between SYGQ LC and the RGG domains (Extended Data Fig. 2D, E). However, these interactions must be weak as the perturbations are small and do not appear saturated even at 300 μM and the motions of FUS LC remain largely unperturbed in the presence of excess MBP-RGG (Extended Data Fig. 2F), which is also incompatible with tight binding. Though the caveats noted above regarding the contribution of MBP to the observed interactions are important to note, combined with the observations that RGG motifs contribute to phase separation of FUS (Fig. 1), these data support the view that phase separation of FUS is mediated by weak, distributed interactions between the SYGQ-LC and the RGG domains.

Detailed description of two chain molecular simulation results

We characterized the interactions between FUS SYGQ LC and RGG domains using two chain all-atom simulations, as in our previous work[19]. In simulations with SYGQ $LC_{11\text{-}54}$ or SYGQ $LC_{120\text{-}163}$ and RGG1, RGG2, or RGG3, we analyzed the free energy surface as a function of total contacts between the SYGQ LC and RGG fragments (Extended Data Fig. 3E). All three RGG domains show free energy minima above zero contacts with some variability in the location of the minimum, suggesting that all three RGG domains can interact with SYGQ $LC_{11\text{-}54}$ or SYGQ $LC_{120\text{-}163}$, albeit to different extents. We note that the amino acid composition between the three RGG domains is slightly different, with RGG1 containing the least amount of arginine residues (Extended Data Fig. 3B). Consistent with the compositional differences within the RGG domain simulation fragments, we observe differences in the single-chain behavior such as the radius of gyration of the RGG fragments (Extended Data Fig. 4F). These differences in single chain properties could account for the difference in contact propensities for certain fragment pairs such as SYGQ $LC_{11\text{-}54}$ with RGG2 and SYGQ $LC_{120\text{-}163}$ with RGG3 (Extended Data Fig. 3E,F). We analyze the nature of these contacts in detail in the main text.

Additional References

85. Conicella, A. E., Zerze, G. H., Mittal, J. & Fawzi, N. L. ALS Mutations Disrupt Phase Separation Mediated by α-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. *Structure* **24**, 1537–1549 (2016).
86. Libich, D. S., Fawzi, N. L., Ying, J. & Clore, G. M. Probing the transient dark state of substrate binding to GroEL by relaxation-based solution NMR. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 11361–11366 (2013).
87. Conicella, A. E. *et al.* TDP-43 α-helical structure tunes liquid-liquid phase separation and function. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 5883-5894 (2020).