Visualization of liquid-liquid phase transitions using a tiny G-quadruplex binding protein

Bikash R. Sahoo,[1,2] Xiexiong Deng,[1,2] Ee Lin Wong,[1,2] Nathan Clark,[1,2] Harry J. Yang,[1,2] Alexey Kovalenko,[1,2] Vivekanandan Subramanian,[3] Bryan B. Guzman,[4] Sarah E. Harris,[5] Budheswar Dehury,[6] Emi Miyashita,[7] J. Damon Hoff,[8] Vojč Kocman,[9] Hirohide Saito,[7,10] Daniel Dominguez,[4] Janez Plavec,[9] and James C.A. Bardwell[1,2*]

[1]Howard Hughes Medical Institute

[2]Department of Molecular, Cellular and Developmental Biology, University of Michigan, Ann Arbor, MI-48109, USA

[3]College of Pharmacy, University of Kentucky, Lexington, KY-40508, USA

[4]Department of Pharmacology, University of North Carolina, Chapel Hill, NC-27514, USA

[5]Department of Biochemistry and Biophysics, University of North Carolina, Chapel Hill, NC-27514, USA

[6] Department of Bioinformatics, Manipal School of Life Sciences, Manipal-576104, India

[7]Center for iPS Cell Research and Application, Kyoto University, Kyoto-6068507, Japan

[8] Department of Biophysics, University of Michigan, Ann Arbor, MI-48109, USA

[9] Slovenian NMR Centre, National Institute of Chemistry, Ljubljana, Slovenia

[10] Institute for Quantitative Biosciences, The University of Tokyo, Tokyo-1130032, Japan

Abstract

Liquid-liquid phase transitions govern a wide range of protein-protein and protein-RNA interactions. Although the importance of multivalency and protein disorder in driving these transitions is clear, there is limited knowledge concerning the structural basis of phase transitions or the conformational changes that accompany this process. In this work, we found that a small human protein, SERF2, is important for the formation of stress granules. We determined the solution NMR structure ensemble of SERF2. We show that SERF2 specifically interacts with non-canonical tetrahelical RNA structures called G-quadruplexes, structures linked to stress granule formation. The biophysical amenability of both SERF2 and RNA G4 quadruplexes have allowed us to characterize the multivalent protein-RNA interactions involved in liquid-liquid phase transitions, the role that protein disorder plays in these transitions, identify the specific contacts involved, and describe how these interactions impact the structural dynamics of the components enabling a detailed understanding of the structural transitions involved in early stages of ribonucleoprotein condensate formation.

42 **Main**

43 When cells experience stressors that dissociate the translation complex, the resulting
44 naked mRNA recruits proteins to form a liquid-liquid phase transition (LLPT) compartment
45 known as the stress granule[1]. These membraneless organelles have been implicated in storing
46 proteins and RNAs and protecting cells from stress-induced damage[2]. A key stress granule
47 protein, G3BP1, recruits G-quadruplexes (G4s) to stress granules[3]. RNA G-quadruplexes (rG4s)
48 are non-canonical RNA structures formed through Hoogsteen guanine (G-G) base pairs
49 stabilized by a centrally located monovalent cation distinguishing them from canonical Watson-
50 Crick base paired structures[4]. These non-canonical structures have also been detected in the
51 nucleus, cytoplasm, mitochondria, and the endoplasmic reticulum[5–7]. They have been linked to
52 telomere maintenance, initiation of DNA replication, control of transcription and translation and
53 protein folding[8–12]. Expansion of the rG4 forming G4C2 hexanucleotide RNA repeat sequences
54 within the *C9orf72* gene has been linked to neurodegenerative diseases including amyotrophic
55 lateral sclerosis (ALS) and frontotemporal dementia (FTD)[13]. rG4s are thought to be transiently
56 folded in vivo[14]. However, in cells experiencing stress conditions, rG4s are enriched and bind to
57 disordered proteins, including G3BP1[15–18]. Upon this binding, rG4s appear to be involved in the
58 formation and organization of stress granules[13,17,18].

59 Despite the emerging importance of rG4s, relatively little is known regarding the regulation
60 of their folding and function in vivo. Several proteins with disordered domains[19], including
61 nucleolin[20], helicases[21,22], FUS[23], hnRNPA1[24], TDP43[25], TRF2[26], and Znf706[27] have been shown
62 to interact with rG4s in vitro, to regulate their folding. Helicases, for example, are thought to
63 unwind rG4s to safeguard against G4-induced genome instability[28], while nucleolin acts to
64 enhance the stability of G4 structures[29]. Determining the mechanism and structural details of
65 rG4-protein interactions will aid in understanding how proteins can regulate the functions of
66 these non-canonical nucleic acid structures. However, our understanding of these interactions is
67 complicated by the often-dynamic nature of the protein and rG4 components of these
68 complexes and their tendency to undergo LLPT[19]. Recent studies have elucidated the structural
69 mechanisms of a few proteins such as helicases DHX36 and Pif1[30,31] and nucleolin[32] whose four
70 RNA-binding domains bind DNA G4s to modulate their function. Despite these advances, the
71 structural mechanisms by which proteins interact with rG4s to mediate ribonucleoprotein
72 condensation via LLPT, or how the dynamic behavior of individual LLPT components is
73 regulated, remain poorly understood. For example, though RNA molecules or intrinsically
74 disordered proteins are identified as key components in LLPT, how proteins that can bind rG4s

75 and mediate LLPT is unclear[33,34]. One way around this apparently intractable problem is to
76 focus on NMR tractable components of stress granules as NMR is unrivaled in its ability to
77 provide detailed dynamic information[35]. However, NMR has so far only shown limited utility in
78 analyzing LLPT [36]. Though >60% of proteins in the human proteome appear to contain
79 disordered regions[37], most of these proteins are large, making them a challenge to analyze
80 using NMR[36]. We decided to focus on small rG4s since several of them have detailed structural
81 information available and since some appear to be involved in stress granule formation[15–17,38].
82 On the protein side, we chose to study the small NMR tractable protein SERF2[39], which we
83 show here specifically binds to rG4s and is an important component of in vivo stress granules.
84 Through studying SERF2 and rG4 interactions, we were able to obtain detailed insights into how
85 SERF2 engages rG4 structures and potentially contributes to condensate formation.

86 **RESULTS**

87 **SERF2 binds to RNA G-quadruplexes**

88 SERF family proteins were initially identified as in vivo drivers of amyloid formation[40–43], a
89 process that has been linked to various age-related diseases[39–42]. However, their normal
90 physiological functions remain obscure. SERF related proteins share a conserved, highly
91 charged N-terminal domain and are remarkably small, ~60-80 amino acids in length[43].
92 Previously, we showed that the SERF-related human Znf706 protein binds with high affinity to
93 G4s primarily using N-terminal sequences that are conserved within the SERF protein family[27] .
94 This motivated us to explore if other SERF family proteins such as human SERF2 also bind G4
95 structures. Using fluorescence polarization binding assays, we found that SERF2 binds with
96 sub-micromolar affinities to various RNA sequences that are known to fold into rG4
97 structures[3,44], including polyG, $(GGGA)_4$, $(GGA)_7$, $(G4C2)_4$ and $(AGG)_5$. In contrast, no
98 interactions with SERF2 were observed with guanine rich RNA oligonucleotides that fail to form
99 rG4 structures, as well as the homo-polyribonucleotide sequences polyA, polyC, or polyU, or
100 single or double-stranded RNA sequences that form hairpin structures under physiological
101 buffer conditions (pH 7.4, 150 mM KCl) (Figs. 1a and S1). SERF2 shows ~10- to 20-fold
102 stronger binding to rG4 forming repeat-sequences such as four repeats of GGGGCC and
103 human telomeric UUAGGG (TERRA23) as compared to their DNA G4 counterparts (Fig. S2a
104 and S2b). In contrast, an identical experimental setup demonstrated that SERF2 does not bind
105 non rG4 forming structures or unstructured RNAs including HIV-1 TAR (5′-
106 GGCAGAUCUGAGCCUGGGAGCUCUCUGCC-3′)[45], four repeat CCCCGG[46] sequences and a

107 mutated telomeric four-repeat sequence UUACCG[47] that is known to not form G4 structures
108 (Fig. S2c).

109 **RNA binding sequence specificity of SERF2 in high-throughput sequencing**

110 To further investigate the RNA binding specificity of SERF related proteins, two
111 independent high-throughput binding assays were performed on SERF2, RNA bind-n-seq[48] and
112 FOREST[49] (Folded RNA Element profiling with STructure). Both assays have been previously
113 validated for known RNA binding proteins including those that specifically bind to rG4.The top
114 five 6-mers obtained using the RNA bind-n-seq enrichment analysis are all very rich in guanines
115 and are predicted to form stable rG4 structures[50] (Figs. 1c-d, S2d). Using an independently done
116 FOREST analysis[49] we found that 8 of the top 10 binding microRNA sequences were previously
117 known to form rG4 structures (Fig. 1e).

118 The binding affinity of SERF2 for the FOREST hit (UG4U)6 and other previously studied
119 rG4 sequences that are found in the cell, including some that are found in stress granules, were
120 next determined using a fluorescence polarization assay. These rG4s include human telomeric
121 TERRA repeat-containing RNA[51] TERRA23 (UAGGGUUAGGGUUAGGGUUAGGG) , the four-
122 repeat (GGGGCC)4 sequence (referred to as G4C2 hereafter), a hexanucleotide repeat found
123 within the *C9orf72* gene, whose expansion is associated with ALS[52], and the mRNA G4s
124 encoded by genes *Mark2* (5′-GAAGGGGAGGGGGCUGGGGGGGGGCAGGG-3′) and *Stxbp5*
125 (5'-GGGAAGGGAAGGGGAGUGGG-3') both of which are known to localize to stress
126 granules[53,54]. SERF2 binds to all of these sequences with low micromolar affinity with K_D values
127 for rG4s G4C2 the K_D was 0.88 ± 0.1 μM, for TERRA23: 0.30 ± 0.02 μM, for UG4U: 0.80 ± 0.3
128 μM, and for *Mark2*: the K_D was 4.5 ± 0.5 μM (Figs. 1f and S2e and S2f). These binding affinities
129 are similar to the affinities of other rG4 binding proteins for their rG4 binding partner including
130 cold-inducible RNA-binding protein[55], FUS[56], and FMRP[57] as well as G4 binding small
131 molecules such as pyridostatin, NMM, and BRACO-19[58]. We conclude that SERF2 specifically
132 binds to rG4 structures. We had previously found that the SERF related protein Znf706,
133 specifically binds to dG4 sequences[27].

134 **SERF2 forms liquid-liquid phase transition droplets with RNA G-quadruplexes**

135 In isolation, SERF2 remains soluble even at concentrations as high as 1 mM, in various salt
136 concentrations (0 - 200 mM KCl) and under crowding conditions (10 % w/v PEG8000) (Fig.
137 S3a). Under non-crowding conditions, SERF2 also shows no phase separation when mixed with
138 total HeLa cell RNA or rG4s (Figs. 2a-c and S3b). However, in the presence of crowding agent
139 PEG8000, we found that SERF2, even when present in concentrations as low as 1.5 to 3 μM,

140 form irregularly shaped particles when mixed with total HeLa cell RNA (Fig. 2a). To understand
141 the physical properties of these particles, fluorescence recovery after photobleaching (FRAP)
142 measurements were carried out. FRAP results show that particles containing SERF2 interacting
143 with total RNA are reversible and dynamic in nature with a recovery halftime of 13 seconds (Fig.
144 2b). This observation indicates that the irregular shaped SERF2-total RNA particles are similar
145 to mesh-like condensates that are observed in some RNA binding proteins[59,60]. In contrast,
146 SERF2 forms spherical phase separated droplets in the presence of different rG4 forming
147 sequences in crowding conditions that include TERRA23, (G4C2)4, and (UG4U)6 (Fig. 2c).
148 However, these rG4 sequences when tested in isolation do not phase separate on their own
149 (Fig. S3b) suggesting SERF2 and rG4 interaction mediates phase separation. Phase regime
150 studies revealed that SERF2-rG4 droplets form at a variety of concentrations of salt, PEG8000,
151 protein, and TERRA23 rG4 (Figs. 2d and S3c). The 1.5-3 μM concentrations of SERF2 that
152 allow phase separation in the presence of rG4s are below the SERF2 concentrations present in
153 human cells which range from ~ 7 μM in U2OS cells to ~ 23 μM in MCF7 breast cancer cells[61].
154 SERF2 also phase separate with a small 12-nucleotide long TERRA12 rG4 but does not phase
155 separate when mixed with the similarly sized non-rG4 forming 10- or 20-nucleotide polyA RNA
156 (Fig. S3d). We postulate that rG4 structures likely trigger SERF2 binding that drives LLPT.

157 **SERF2 and RNA G-quadruplexes form slowly exchanging droplets**

158 rG4s have recently been shown to regulate LLPT, and in neurodegenerative diseases there
159 is evidence that they accumulate[62] and may act as molecular scaffolds that help transition liquid
160 droplets to more solid gel-like aggregates[63]. We previously showed that G4s promote α-
161 synuclein aggregation, a process that is associated with Parkinson's disease[27], but suppress the
162 ability of the SERF related protein Znf706's to promote α-synuclein aggregation by promoting
163 Znf706's ability to form LLPT. Recently TERRA rG4s have also shown to promote α-synuclein
164 aggregation in vivo[47] indicating that G4s regulate protein aggregation. Our observation that
165 SERF2's interaction with rG4 can lead to LLPT is intriguing and suggestive of a link to the
166 possible roles of SERF2 and rG4 in neurodegenerative diseases. To explore how rG4 binding
167 may affect SERF2's transition to the liquid-liquid phase, we examined the dynamics of the
168 interactions between SERF2 and RNA or rG4 using FRAP experiment. A recovery with a
169 halftime of ~30 seconds was observed for SERF2 condensates formed with short (30
170 nucleotide) or long (700-3500 kDa) random length polyA (Figs. S3e and S3f). In contrast,
171 SERF2 droplets in the presence of structured RNA that include rG4s or hairpin HIV-1 TAR RNA
172 that bind weakly to SERF2 in solution were found to recover more slowly (Fig. S3h). In the

173 presence of rG4s, SERF2 recovery half-time was 49, 70 and 79 seconds for TERRA23,
174 (G4C2)4 and (UG4U)6 rG4s respectively, which is similar to recovery times observed for the
175 non-rG4 hairpin structure of HIV-1 TAR RNA (49 seconds) (Figs. 2e, S3g, and S3h). These
176 results suggest that SERF2 droplets are less fluid when formed with structured RNA binding as
177 compared to SERF2 droplets that form with unstructured RNAs.

178 **Structural characterization of SERF2 reveals partial disorder**

179 Studying the detailed structure of ribonucleoprotein condensates is challenging due to their
180 dynamic and heterogeneous nature[64], and often the large size and complex multi-component
181 composition of these condensates. One challenge is the structural flexibility of the components
182 involved, many of which contain regions of intrinsic disorder[65]. Though reviews have highlighted
183 the importance of intrinsically disordered proteins in many biological systems[66,67], there is an
184 understandable paucity of experimental approaches to examine how intrinsically disordered
185 proteins interact with other cellular components. To address this gap, we focused on the
186 interaction between the partially disordered protein SERF2 and rG4 structures. This system
187 provides a tractable model for structural studies, as both SERF2 and rG4 species such as the
188 G4C2 repeat, and the telomeric-repeat TERRA10 or TERRA12 are NMR compliant, and NMR is
189 well suited to study flexible interactions[68].

190 While TERRA12 and TERRA10 rG4 structures are very well-characterized[69–71], the
191 structural properties of SERF2 are relatively poorly characterized. Thus, to study SERF2-rG4
192 interactions, including those involved in LLPT, we first explored the structural properties of
193 SERF2. SERF2 presented a narrow 2D 1H-15N HSQC NMR chemical shift dispersion at 4 °C,
194 indicating that it contains a high degree of structural disorder even at this low temperature (Fig.
195 S4a). At 37 °C most of the 1H-15N peaks disappear from the HSQC spectrum, suggesting that
196 a higher degree of disorder exists for SERF2 at this temperature or that induction of different
197 conformational states in SERF2 occurs upon temperature upshift, a phenomenon that has been
198 demonstrated to occur in other disordered proteins[72,73] (Fig. S4a). Circular dichroism
199 spectroscopy analysis indicates that substantial helicity exists in the SERF2 structure at 4 °C,
200 with decreasing helicity as the temperature increases (Fig. S4b). The well resolved peaks
201 observed for SERF2 in multidimensional NMR experiments performed at low temperature (4 °C)
202 and at low pH (5.5) or both, enabled us to achieve NMR backbone assignment for 51 of the 58
203 non-proline residues[39] found in SERF2 and to perform torsion angle analysis on SERF2 using
204 TALOS-N[74]. Using a total of 558 restraints that include 40 dihedral angles and 24 hydrogen
205 bonds (Table 1) and 494 NOE restraints (intra - 205, short - 220 and medium - 69), we

206 calculated an ensemble of structures for SERF2 (Fig. 3a). The final ensemble shows good
207 structural quality in the ordered region (residues 32-48), with a backbone RMSD of 0.24 ± 0.09
208 Å, and no distance violations >0.5 Å or Ramachandran outliers. We found that the N-terminal
209 (1-32) and C-terminal (48-59) regions in SERF2 to be dynamic and disordered, while residues
210 33-47 adopt an ensemble-averaged helical conformation (Figs. 3a and S4c). The structural
211 ensemble of the helical core and disordered flanks observed in SERF2 is shared with its
212 sequence homologue, human Znf706 which also has a disordered N-terminal domain, though
213 Znf706 contains a C-terminal zinc finger that SERF2 lacks[27].
214
215 **Disorder in SERF2 mediates its interaction with RNA G-quadruplexes**
216 Disordered regions are commonly involved in protein-protein and protein-RNA
217 interactions[75], but the exact role that disorder plays in these interactions is often unclear.
218 Studying the molecular interactions between TERRA rG4 and SERF2 provides a well-suited
219 model to elucidate the mechanistic contributions of disorder to RNA recognition and condensate
220 formation. TERRA rG4 was selected because it has been previously structurally well
221 characterized,[76,77] and because of the evidence for its presence in cellular condensates like
222 stress granules, and its potential involvement in their formation and organization[53]. We
223 employed NMR to explore in detail the interaction between SERF2 and TERRA rG4, the role
224 played by SERF2's disordered domain and the structural impact of binding in both interacting
225 partners. The long repeat TERRA rG4s such as TERRA23 or TERRA60 binds stronger to
226 SERF2 as compared to short TERRA12 or TERRA10 repeats (Figs. 1f and S5a). However, the
227 short TERRA12 or TERRA10 rG4 sequences fold better and form multimeric structures suitable
228 for structural studies, as demonstrated earlier[69]. TERRA12 in 20 mM NaPi, 100 mM KCl, pH 7.4
229 buffer that presented a strong CD absorption at ~263 nm and well-resolved imino proton NMR
230 peaks (Fig. S5b) indicating that it is well folded compared to TERRA23[64]. Studying different
231 TERRA rG4 sequence lengths enabled us to investigate how sequence influences rG4 structure
232 and RNA-protein interactions, with some lengths being more useful to address certain specific
233 questions. For instance, we used both longer and shorter TERRA rG4 sequences to map RNA
234 binding sites in SERF2, but primarily used shorter sequences such as TERRA10 and TERRA12
235 to map protein binding sites on RNA due to their better NMR spectral resolution. In contrast,
236 longer TERRA sequences such as TERRA23 and TERRA60, though structurally more
237 heterogeneous and challenging to characterize in detail, are biologically relevant given their
238 intramolecular G4 structure[78], association with telomeric chromatin, involvement in scaffolding
239 amyloid aggregation and evidence suggesting their role in stress granule formation and genome

240 regulation[47,53,79]. The TERRA10, TERRA12 and TERRA23 sequences bound to a similar region
241 of SERF2 at moderate to low RNA to protein ratios (Figs. 3c, S5c and S5d). The longest
242 TERRA rG4, TERRA60, showed signal broadening in several SERF2 residues (Figs. S5d and
243 S6) at very low concentration (1:40 RNA:protein) that increases upon increasing RNA
244 concentration (1:4 RNA:protein). This observation is similar to what observed for equimolar or
245 high RNA concentrations for short and long TERRA rG4 sequences such as TERRA23,
246 TERRA10 and (Fig. S6), consistent with the hypothesis that various length TERRA rG4 lengths
247 bind to SERF2, but with affinities that increase with TERRA length.
248
249 The residue-specific binding site analysis on SERF2 by TERRA rG4 was next probed using
250 TERRA23 a length that binds relatively stronger as compared to TERRA12, and shows a sub-
251 micromolar binding affinity similar to TERRA60. Further, at low RNA to protein ratio, it yields
252 small NMR signal broadening enabling 1H-15N HSQC titration experiments. TERRA23 rG4-
253 SERF2 interaction causes substantial chemical shift perturbations both in SERF2's N- (3-21)
254 and C-terminal (51-56) residues, as tracked by 1H-15N HSQC spectra (Fig. 3b). At low
255 TERRA23 rG4 concentrations (Fig. 3c, green), SERF2 residues R7, R11, K16, K23, R27, Q46,
256 K47 and K54 showed substantial chemical shift changes suggestive of direct binding to rG4 or
257 conformational alteration upon rG4 interaction. The N-terminal domain residues of SERF2
258 including T2, N5, R7, R11 and S21 showed increasing perturbation as the rG4 concentration
259 increased (Fig. 3b), implicating the N-terminal domain of SERF2 of involvement in rG4 binding.
260 Binding assays showed that the isolated N-terminal domain of SERF2 (residues 1-32) binds
261 TERRA23 rG4, though ~10-fold weaker than the full-length protein. Conversely, the C-terminal
262 domain of SERF2, residues 31-59, in isolation, showed no binding saturation upon addition to
263 TERRA23 rG4 quadruplex in a fluorescence polarization binding assay (Fig. S7a). These results
264 suggest the N-terminal domain of SERF2 drives rG4 recognition, but both domains coordinate
265 for high-affinity binding, forming a tight SERF2-TERRA rG4 complex.

266 To further explore the impact of TERRA rG4 binding on SERF2 dynamics at residue level
267 resolution, we compared the NMR relaxation data of SERF2 in its free and rG4 bound states.
268 The relaxation measurement results showed that TERRA23 rG4 binding significantly increases
269 R_2 values in both the N-terminal and C-terminal intrinsically disordered regions of SERF2,
270 suggesting that SERF2 undergoes segmental slow-motion (Fig. S7b). In contrast, R_1 and
271 heteronuclear NOE values remain largely unchanged, suggesting that SERF2 maintains local
272 flexibility after interaction with TERRA23 (Figs. S7c and d). The dynamic and disordered SERF2

273 regions spanning residues 3-24 and 48-56, identified by their low heteronuclear NOE values
274 (<0.3), showed significant chemical shift perturbations and reduced signal intensities upon
275 TERRA23 rG4 interaction (Figs. S7d and S7e). These dynamics suggest that flexibility in these
276 regions facilitates rG4 recognition. Binding also led to increased spin-spin relaxation rates (T2)
277 and higher backbone R2/R1 values, indicating restricted SERF2 dynamics in the SERF2-
278 TERRA23 rG4 complex compared to the unbound state (Fig. 3d). The α-helical core of SERF2
279 (residues 37-47) showed a marked increase in R_2 values but no significant changes in R_1 or
280 heteronuclear NOE values, indicating local conformational exchange or slow dynamics in the
281 intermediate regime occur upon RNA binding, rather than changes in backbone order or in
282 secondary structure. Protein secondary structure analysis using CD spectroscopy confirmed
283 this, no substantial change in SERF2 protein helicity was present in the presence of TERRA23
284 rG4 (Fig. S7f). The slowed dynamics and peak broadening in SERF2 upon TERRA23 rG4
285 binding (Fig. S6) could arise from an increase in size that occurs upon protein-TERRA23 rG4
286 association to form either dimers or higher oligomeric forms. Size-exclusion chromatography
287 analysis that showed that mixing SERF2 with TERRA23 rG4 generates multimeric species with
288 distinct populations do not present in solutions containing SERF2 or TERRA23 in isolation. A
289 1:1 SERF2-rG4 complex of ~30 kDa, and higher-order multimeric complexes that are >30 kDa)
290 (Fig. S7g). These complexes are detected by rG4 absorption at 260nm where SERF2 has no
291 absorption.
292
293 We next tested if SERF2 has preferential binding sites within the TERRA rG4 using the
294 structurally well characterized TERRA10 and TERRA12 rG4 sequences. These lengths were
295 tested because they exhibit well-resolved guanine imino protons in comparison to TERRA23's
296 poorly resolved spectra. Binding of SERF2 caused chemical shift changes in the imino protons
297 of G4, G5, G9, and G10 of TERRA12 and TERRA10 (Fig. S8). This observation indicates
298 SERF2 may either directly bind near TERRA10 or TERRA12 rG4's G-quartet core, or
299 alternatively that SERF2 binding may impact G-quartet conformation (Fig. S8). To further probe
300 the spatial proximity between SERF2 and the TERRA12 rG4, saturation transfer difference
301 (STD) NMR or experiments were carried out. STD NMR approach detects interactions between
302 protein and rG4 by transferring selective radiofrequency saturation from one binding partner to
303 the other, indicating close spatial proximity. When guanine imino proton (~11.11 ppm) of
304 TERRA12 rG4 was selectively saturated, magnetization transfer was observed across both
305 amino (~7-8.5 ppm) and aliphatic (~0.5-3 ppm) regions of SERF2-rG4 complex (Fig. S9a).
306 These transfers were absent when TERRA12 rG4 was examined in isolation by saturating

307 different imino peaks, indicating that the observed saturation arises from complex formation
308 (Fig. S9b). When the amino proton of SERF2, particularly the signal corresponding to residue
309 A33 at ~8.5 ppm (isolated from RNA amino peaks, Fig. 3b), was selectively saturated, transfer
310 to the RNA imino region was also observed (Fig. S9c). The STD effects were observed in the
311 ~11-11.5 ppm region, corresponding to the guanine imino protons of TERRA12 rG4, indicating
312 that magnetization was transferred from SERF2 to these protons upon complex formation.
313 Importantly, these effects were absent in control experiments using either SERF2 or TERRA12
314 alone (Figs. S9b and S9d). Together, these results support a model in which SERF2 binds in
315 close spatial proximity to the G-quartet core of TERRA12 rG4, enabling inter-molecular
316 magnetization transfer being detectable by STD NMR.

317 **SERF2 and RNA G-quadruplexes form multimeric complexes**

318 Since SERF2 forms multimeric protein-RNA assemblies and is involved in RNA-mediated
319 phase transitions, we next investigated the mechanisms of SERF2-RNA complex formation and
320 how multimerization occurs. To determine the stoichiometry of these complexes, we used
321 analytical ultracentrifugation which provides higher mass resolution as compared to size-
322 exclusion chromatography. Analytical ultracentrifugation of TERRA12 rG4 in isolation shows two
323 structures with sedimentation coefficients of 1.95 and 3.33, suggestive of globular, folded
324 structures (Fig. 3e). In the presence of SERF2, sedimentation coefficient values were observed
325 with predicted sizes of ~ 12.4, 17.4, and 27.1 kDa corresponding to monomeric rG4, 1:1, and
326 1:2 rG4-SERF2 complexes (Fig. 3f). These findings demonstrate that SERF2 forms multimeric
327 complexes with TERRA12 rG4, supporting a model in which SERF2-rG4 interactions drive
328 multimerization. Notably, similar multimerization behavior was observed when SERF2 was
329 mixed with the longer TERRA23 rG4 (Fig. S7g).
330

331 **Structural dynamics and stability of the SERF2-RNA G-quadruplex complex**

332 The TERRA rG4 quadruplex binding sites in SERF2 we obtained from NMR measurements
333 were used to build the TERRA-SERF2 complex structure and subjected to an all-atom
334 molecular dynamic (MD) simulation study. The goal of this analysis was to characterize the
335 structural changes and dynamic interactions that occur during the early stages of SERF2-rG4
336 complex formation and their potential involvement in liquid-liquid phase transitions. To model
337 the complex, the TERRA rG4 sequence length was selected based on available high-resolution
338 structures, TERRA12, which forms a dimeric quadruplex (PDB ID: 2KBP), and TERRA10, which
339 adopts a tetrameric rG4 conformation (PDB ID: 2M18)[80]. NMR experiments revealed that both

340 TERRA10 and TERRA12 rG4s bind to a similar surface on SERF2, involving conserved
341 guanine residues such as G4, G5, G9, and G10. Despite differences in higher-order folding with
342 TERRA12 forming a dimeric quadruplex and TERRA10 adopting a tetrameric conformation, the
343 local structural elements involved in SERF2 binding appear to be preserved. This observation
344 supports the use of either RNA construct to model protein-RNA interaction interfaces. Given that
345 in 100 ns all-atom MD simulation, TERRA10 adopts a more compact and stable conformation
346 as compared to TERRA12 a rG4 that showed a relatively distorted end (Fig. S10a). We
347 therefore chose to use TERRA10 for the long-range MD simulation studies. Notably, MD
348 simulation is well suited to probe the conformational flexibility and binding dynamics of protein-
349 RNA complexes, and remain informative regardless of differences in global topology. We
350 consider that the choice of TERRA10 does not compromise the broader relevance of the model.
351 Using the HADDOCK program[81], models of the TERRA rG4 and SERF2 1:1 and 1:2 complexes
352 were built. We identified residues in SERF2 that are involved in direct or indirect binding (due to
353 conformational alteration upon rG4 interaction) as evidenced from NMR chemical shift
354 perturbations performed under non-crowding conditions. We also identified rG4 guanines that
355 interact with SERF2 and show magnetization transfer in NMR saturation transfer difference
356 experiments (Figs. 3b and S9). These interacting protein residues and guanines were used as
357 input to HADDOCK to generate SERF2-rG4 complex structures (see methods). These models
358 were then subjected to microsecond long all-atom MD simulations to further examine the
359 stability and dynamic nature of the complex structures. After a 0.5 μs all-atom MD simulation,
360 the 1:1 rG4:SERF2 complex retained N-terminal contacts but lacked C-terminal interactions
361 (Fig. 3g). In contrast, the 1:2 rG4:SERF2 complex gained additional C-terminal contacts (Fig.
362 S10b), aligning with our NMR data that showed chemical shift perturbations for several C-
363 terminal residues upon rG4 binding. These observations illustrate specific multivalent
364 interactions that occur between SERF2 and rG4 during an oligomerization process that can lead
365 to LLPT.
366 Our analysis revealed that a hydrophilic core in the SERF2-rG4 complex is formed by
367 residues in regions that are dynamic in monomeric SERF2 (Figs. 3g and 3h). Specifically, N-
368 terminal residues in SERF2 including R3, R11, K16, K17, and K23 form hydrogen bonds with
369 TERRA10 rG4 nucleotides G1, G3, G7-G9 that are located in the bottom two G-quartets. This
370 interaction forms a planar interface in the SERF2-TERRA10 rG4 complex. It consists of SERF2
371 residues M1-T2 and R3 facing opposite to K16-K17 in one interface, and residues R11-R36
372 facing residues S19-S21- K23 in another interface, creating a quadrupole-like contact
373 architecture with alternating positive and negative charge residues distributed on the corners of

374 a square (Fig. 3h). The guanines in the TERRA10 rG4 structure (PDB ID: 2M18) form five
375 stacked G-quartets through an 'i' to 'i+6' nucleotide pattern[76] (Fig. S10c). SERF2 binding in the
376 1:2 complex disrupts two G-quartets, strongly distorting the rG4 structure. Residues R11 and R3
377 break the rG4 G9-G3 Hoogsteen bond, while S19-S21-K23 and M1-T2 interactions expose the
378 rG4 U4, which stabilizes the structure via rG4 G2-G3 hydrogen bonding (Figs. 3h and S10b).
379 Additionally, K16-K17 bonding with TERRA10 rG4's G7 and G8 nucleotides disrupts G7-G1 and
380 G7-G8 Hoogsteen pairs, further distorting the TERRA10 rG4 structure. Prior to binding SERF2,
381 TERRA10 rG4 forms a parallel topology, and after binding, TERRA10 rG4 structure is
382 noticeably distorted in the MD simulation.
383 The hydrophilic core of the 2:1 SERF2:TERRA10 rG4 complex is larger than that present in
384 the 1:1 complex due to charged C-terminal residues (R36, K50, K54, K55) that form hydrogen
385 bonds with TERRA10 rG4. Over the atomic simulation timescale, SERF2 shifts from interacting
386 with uracil in the 1:1 complex to guanine in the 1:2 complex (Fig. S10d). rG4 structure distortion
387 was also observed in the 1:2 complex bound to SERF2. Experimental evidence backing up
388 these MD simulation results includes a decrease in TERRA10 rG4 parallel structure with
389 increasing SERF2 concentrations as measured by CD spectroscopy, and a co-commitment loss
390 in imino NMR proton signal, indicating rG4 destabilization (Figs. S8 and S10e). Though these
391 findings correlate with the atomistic MD simulation results, the simulation appears to
392 overrepresent the structural distortion in solution as compared to the experimental results. This
393 is evidenced by the fact that, all-atom MD simulation of the TERRA10 rG4 in isolation in 150
394 mM KCl solution presented a slight distortion in rG4 structure in the presence of K^+ ions (Fig.
395 S10a). The stronger structural distortion seen in our MD simulation as compared to our CD
396 results could thus be contributed by multiple factors. The partition of K^+ ion into the tetrad-
397 quartets similar to as has previously been observed for the G4s[82] accompanied by K^+ ion
398 dissociation upon SERF2 binding. Alternatively, it could be an effect of the protein-RNA force
399 field we utilized, as force field effects on G4 folding have been previously demonstrated[80,83].
400
401 **Mutational study reveals the multivalent interactions that drive SERF2 phase transition**
402 MD simulations and NMR experiments provide the evidence that many of the
403 evolutionary conserved lysine residues in SERF2 spanning the region 16-25 are important in
404 rG4 binding. To further experimentally validate that these residues are involved in binding, we
405 performed mutagenesis studies on the conserved K16, K17, K23, G24, and K25 residues that
406 are predicted in our MD simulation to lie at the binding interfaces. Gel-shift assays show that
407 single mutations (K23A or K25A) minimally affect rG4 binding, but the double K16-K17 mutation

408 significantly impairs TERRA23 rG4 substrate binding (Fig. 3i). Interestingly, this same K16-K17
409 mutation in SERF2 also fails to bind its amyloid substrate[84]. Mutations disrupting another
410 interface in SERF2 (K23, G24, and K25) also weaken rG4 binding, while mutating all conserved
411 lysine residues in both binding interfaces further reduces rG4 binding. Mutations in the C-
412 terminal Q46, K47 residues show no effect on binding as evidenced by chemical shift
413 perturbations in the NMR experiments, indicating C-terminal contacts have minimal effects on
414 rG4 recognition. Under crowding conditions, LLPT is slightly affected by single-point mutations
415 (K23A or K25A), or disruption of single-binding interfaces (K16-K17 or K23- G24-K25), but LLPT
416 is impacted when both binding interfaces are disrupted (Fig. S10f). These findings suggest
417 SERF2 uses multivalent interactions via multiple interfaces to drive phase separation.

418 **Structural insights into ribonucleoprotein condensate formation**

419 Biomolecular condensation underpins key cellular processes and diseases[85], yet atomic-
420 level understanding of this process is limited, especially for RNA-protein condensates. Using
421 NMR and all-atom MD simulations, we have uncovered the detailed interactions that occur
422 between SERF2 and TERRA rG4 under non-crowded conditions. To overcome NMR limitations
423 that occur in crowded systems, which tend to broaden NMR signals, we added the molecular
424 crowder PEG to our all-atom MD simulations to investigate phase separation under more
425 physiologically relevant conditions. The simulation system contains 30 SERF2 and 30
426 TERRA10 rG4 molecules. Our all-atom MD simulations are predicated on the assumption that
427 the type of interactions between protein and RNA do not significantly change upon the addition
428 of crowding agents. Condensation is facilitated by an increase in the effective interaction
429 network density under molecular crowding, which enhances multivalent interactions without
430 altering the intrinsic valency of the individual components[86].

431 Simulation results revealed a spectrum of assemblies, from unbound TERRA10 rG4 to
432 SERF2-TERRA10 rG4 oligomers, culminating in a dominant ring-like higher order structure
433 enclosing a hydrated core (~6 nm) (Figs. 4a-c, S11a and Supporting Video SV1). The lower-
434 ordered 2:1 SERF2-TERRA10 rG4 complex (dilute-phase oligomers) retained a four-site binding
435 state (Fig. 4c, top) involving N-terminal residues M1-R3-R7, R11, K16-K18, and K25-R26-R27
436 consistent with what observed in a non-crowded system (Fig. 3h), suggesting early-phase
437 continuity in interaction networks. The 2:2 SERF2:rG4 oligomers added C-terminal contacts,
438 forming multivalent networks pivotal for phase separation (Fig. S11b).

439 The ring-structure emerges from electrostatically driven clustering (Fig. S11c) of the anionic
440 rG4 molecules, which are bridged by multivalent, positively charged SERF2 molecules. While

441 one might expect homogeneous mixing in condensates, the dynamic wrapping of SERF2
442 around clustered rG4s results in the emergence of a peripheral protein-RNA shell and a solvent-
443 rich core (Figs. 4b and 4d). This architecture is multivalent in nature, within which individual
444 SERF2 molecules bridge up to three different rG4 molecules through charged residues (Fig.
445 4b), creating a dynamic multivalent network that likely favors condensation. The presence of a
446 hydrated core may reflect the anisotropic nature of RNA-protein interactions[87], where multivalent
447 bridging and excluded volume effects combine to favor peripheral assembly.
448 The crowding molecule polyethylene glycol (PEG), plays an active but indirect role in this
449 process. Initially distributed throughout the system, PEG molecules gradually become excluded
450 from the dense core region of the condensate as it matures, likely due to steric hindrance and
451 limited PEG permeability through the RNA-protein network (Figs. 4d and 4e). However, PEG is
452 not completely excluded; transient PEG-biomolecule interactions are evident, particularly at the
453 condensate periphery, suggesting that PEG may be playing a stabilizing role (Fig. 4e). This
454 behavior supports PEG's role as a modulator of phase separation acting mainly by enhancing
455 molecular crowding and interaction valency, rather than acting mostly as a direct binding
456 partner.
457 In the higher order ring-like structure, SERF2 engages multiple TERRA10 rG4s via
458 extensive charged interfaces (Fig 4a-c). Contact map analysis results show the N-terminal
459 charged residues in SERF2 predominantly form contacts with TERRA10 rG4 in the large system
460 (Fig. S11d). A single SERF2 molecule interacts with three TERRA10 rG4s through a large
461 interface formed by residues R3, K16, K17, K23, and R27, while smaller interfaces are involved
462 with two other rG4 molecules through residues K25, K55 and K47, K49 (Fig. 4b, top). Similarly,
463 two SERF2 molecules are shown to bind two TERRA10 rG4s in a fashion where one tightly
464 engages its N- and C-terminus, while the other shows a weak contact map through region R11-
465 K16 and exposes its N- and C-terminus hinting at likely recruitment of additional TERRA10 rG4
466 partners (Fig. S11b). This multivalent interaction pattern allows the dynamic recruitment of rG4
467 and protein molecules, as observed in the dilute-phase 2:2 complex (Fig. 4c). Interaction
468 network analysis (see methods) showed that the charged residues are predominantly engaged
469 in rG4 interaction and their bonding occupancy and multivalency increases as the system size
470 increases (Fig. S11e). In the crowded state, most SERF2 molecules bind two TERRA10 rG4s,
471 forming the distinct ring-like structure. This ring-like structure was also observed at early
472 simulation times and in a smaller system comprising only six SERF2 and TERRA10 rG4
473 molecules simulated for 1 μs (Fig. S12a). To investigate whether the formation of this ring-like
474 structure depends on the stoichiometry, we analyzed systems with 1:2 and 2:1 protein to RNA

475 ratios in addition to the 2:2 ratio tested above. Our results show that the ring-like structure
476 observed in 1:1 large system (Fig. S12b) is also maintained in the 1:2 protein to RNA system
477 (Fig. S12c). The 2:1 system generated a partial ring-shaped elongated arc-shape structure
478 similar to one observed in the 1:1 small system (Fig. S12d). These results suggest SERF2-rG4
479 interactions in crowding conditions have a tendency to form ring-like structures that are
480 maintained at different protein to RNA ratios. We note that, the hollow-core feature observed
481 here is a simulation-based observation and has not yet been experimentally validated. However,
482 it is conceptually consistent with the sticker-and-spacer model of disordered protein phase
483 separation, where interaction hotspots (stickers) drive association and flexible regions (spacers)
484 permit structural plasticity[88].
485 The structural stability of TERRA10 rG4 molecules in the condensates were next analyzed
486 by calculating the formation of Hoogsteen hydrogen bonds within the rG4 molecules, with
487 guanine N1, N2 and O2 serving as donor atoms and N7 and O6 serving as acceptor atoms, as
488 a function of simulations time (Fig. 4f). The results showed an average ~40 Hoogsteen
489 hydrogen bonds for the 30 rG4 molecules, very similar to a ~40 Hoogsteen bonds present within
490 a tetrameric five-tetrad TERRA10 rG4 in isolation. These results suggested that the structural
491 integrity of rG4 remain largely unaffected in the simulation. These results are distinct as
492 compared to the non-crowding simulation which showed rG4 structure distortion upon SERF2
493 interaction. The structural stability of TERRA10 rG4 we observe in crowding conditions could be
494 contributed by the crowding agents as molecular crowders are shown to enhance the
495 stabilization of G4 structures[89]. Further, contact analysis presented an average ~3 SERF2
496 molecules are in contact (within 8Å) with one rG4 molecule (Fig. 4f).
497
498 **Single-molecule tracking of SERF2-RNA G-quadruplex liquid-liquid phase transition**
499 **droplets**
500 To validate our MD simulations, we compared the diffusion characteristics of SERF2 in
501 SERF2-rG4s condensates obtained from MD with single-particle tracking using single-molecule
502 fluorescence experiments (Fig. 5a). We hypothesize that the in silico computed diffusion
503 coefficients of SERF2 molecules in condensates likely correlate to experimentally derived
504 diffusion constant as has been previously demonstrated for protein-only condensates[90,91]. The
505 diffusion coefficients of SERF2 molecules in the condensed structure were computed from the
506 all-atom MD simulation by analyzing their mean-square displacement (MSD)(Fig.4g). These
507 analyses yielded average 3D diffusion constant of ~1.07 $\mu m^2/s$ (average MSD ~6.4 $\mu m^2/s$) for
508 SERF2, indicating that the protein molecules remain dynamically mobile within the condensate

environment. In contrast, diffusion in the dilute phase primarily involving dimers and small oligomers was significantly faster, with MSD ranging from ~20-40 μm²/s (Fig. S12e), suggesting rapid diffusion in these states. In contrast, the MSD of SERF2 molecules in higher-order condensates were much slower, with MSD ranging from ~0.3 to 5 μm²/s (Fig. S12e). This corresponds to an average experimental MSD value of 0.95 ± 0.23 μm²/s, measured in SERF-TERRA12 rG4 condensates using single-particle tracking (Fig. 5b), confirming that SERF2 molecules experience slower diffusion in the condensed phase as compared to in the dilute phase. These experimentally measured and simulation-calculated diffusion rates in SERF2-rG4 droplets (Fig. 5c) are also consistent with the ~0.1-0.6 μm²/s diffusion rates that have been observed for other biomolecular nucleoprotein condensates[92,93].

The slow diffusion of SERF2 in TERRA12 rG4 condensates may be due to strong electrostatic interactions of charged lysine and arginine residues with negatively charged nucleic acid bases as has been observed for other nucleic acid-binding proteins[92]. The stability of SERF2-TERRA12 rG4 droplets even under non-physiological salt concentrations (Fig. S13a), suggests that a robust electrostatic network exists between SERF2 and rG4. This is consistent with findings for other ribonucleoprotein droplets[92], which generally exhibit more viscous and rigid properties as compared to protein-only condensates. Indeed, the fusion time of SERF2-TERRA12 or TERRA23 rG4 droplets are very similar as measured by dual-trap optical tweezers which was ~2-4 seconds. This is one to two orders of magnitude slower than fusion rates observed in protein-only condensates[90,91] (Fig. S13b). Additionally, FRAP revealed a relatively slow recovery half-time of ~2 minutes for SERF2-TERRA12 (Fig. S13c) as compared to that observed in protein-only condensates which generally FRAP on the seconds timescale.

The end-to-end distance between the N- (T2C) and C- (A51C) termini of SERF2 was calculated using static fluorescence resonance energy transfer (FRET) measurements in both dilute and condensed phases. In the dilute phase, the mean distance between the N- and C-termini was 1.51 ± 0.5 nm (Fig. 5d), consistent with NMR data suggesting close proximity between the termini in solution[39]. In the condensed phase, the mean diameter of SERF2 was 0.99 ± 0.27 nm (Fig. 5d), indicating that SERF2 adopts a more compact structure within the condensates. We note that considering the experimental limitations (see methods) and assuming a random orientation of SERF2 within the condensate, an accurate distance estimation is not possible between a donor and acceptor. MD simulation results showed the mean distance between T2C and A51 in the condensed structure was slightly larger at 2.09 ± 0.2 nm as compared to the experimental results (Fig. S13d). Single-molecule distance analysis in the MD simulation however presented several SERF2 molecules having an end-to-end

543 distance within a range of ~1.1 to 1.3 nm (Fig. S13d), consistent with the FRET data. We
544 conclude that SERF2 retains a compact structure within the ribonucleoprotein condensates.
545
546 **SERF2 regulates stress granule formation and liquid phase dynamics**
547 To further investigate SERF2's involvement in RNA binding and in LLPT, we next explored
548 SERF2's cellular localization, function, and phase separation behavior in vivo.
549 Immunofluorescence analysis in U2OS cells showed that under non-stress conditions, SERF2 is
550 distributed in the cytoplasm but most intensely to the nucleolus as judged by its co-localization
551 with the nucleolar marker protein fibrillarin (Figs. 6a and 6b). Interestingly, upon exposure to
552 various stressors that include oxidative, osmotic, ER, mitochondrial, or proteasomal stress,
553 SERF2 is primarily localized to stress granules in the cytosol as evidenced by its colocalization
554 with five different stress granule marker proteins (Figs. 6c, 6d and S14a). Nucleolar size and
555 shape were not affected by SERF2 depletion (Fig. S14b). We found some SERF2 colocalization
556 with rG4 structures under stress conditions, as demonstrated by G4-specific BG4 antibody
557 staining (Figs. 6e and 6f). Although rG4s are distributed across multiple cellular compartments,
558 their enrichment in stress granules has been demonstrated recently[16,18,38]. Thus, SERF2's
559 specific interaction with rG4s in vitro suggests a functional interplay between SERF2 and rG4s
560 may also exist in vivo.

561 The depletion of SERF2 using siRNA in U2OS and BJ fibroblast cells markedly reduced the
562 number of stress granules observed, the number of stress granule-positive cells dropped from
563 ~94% to 20% upon sodium arsenite treatment (Figs. 7a, 7b, and S15). SERF2-depleted HeLa
564 Kyoto live cells expressing EFGP-FUS or SERF2-null HEK293T cells also showed markedly
565 fewer stress granules than cells expressing normal levels of SERF2 (Figs. 7c, 7d and S15). The
566 G3BP1 and G3BP2 proteins are thought to play a key role in stress granule formation[94]. We find
567 that SERF2 depletion alone was as impactful as the combined knockdown of G3BP1 and
568 G3BP2[95] in reducing the number of EGFP-FUS granules (Fig. S15). In addition to contributing to
569 stress granule formation, we found that SERF2 also influences their fluidity. FRAP
570 measurements revealed significant fluorescence recovery occurs in about the 60 second
571 timescale in the EGFP-FUS stress granules formed in response to sorbitol treatment control
572 HeLa Kyoto live cells, while granules formed in SERF2-depleted HeLa Kyoto live cells showed
573 very little recovery within this time frame (Fig. 7e). A similar, although slightly less dramatic
574 result is seen for stress granules formed in response to sodium arsenite treatment (Fig. 7f).
575 These results indicate that SERF2 is not only impacting the formation, size, and distribution of

576 stress granules, but also their liquid properties. Our data suggests that SERF2 promotes stress
577 granule formation in vivo and may enhance their fluidity likely through LLPT with rG4 as we
578 observed in vitro (Fig. 2b). We conclude for these observations that SERF2 containing
579 condensates are liquid-like in vitro and in vivo, unlike gels or aggregates which are
580 characterized by extremely slow or no measurable diffusion and exchange[96].
581

582 **SERF2 and RNA G-quadruplexes modulate the G3BP1-RNA phase transitions**

583 Our earlier findings which indicated SERF2 depletion could impact stress granule fluidity
584 motivated us to investigate whether SERF2 influences the phase behavior of the key stress
585 granule protein G3BP1[97] in vitro. We observed that SERF2 condensates are colocalized with
586 G3BP1 and promotes G3BP1-RNA condensation, with condensates of size up to 4.2 μm
587 diameter forming in the presence of SERF2, G3BP1 and total cellular RNA (Figs. 8a, 8b, and
588 S16a). Notably, rG4s or short poly RNA do not induce G3BP1 LLPT in the absence of SERF2 in
589 a non-crowding environment even when they are present at high concentrations (100 μM). In
590 contrast, under crowding conditions (2.5% PEG) where G3BP1 does not phase separate, rG4s
591 combined with SERF2 do trigger G3BP1 condensation even at low (20 μM) protein
592 concentrations (Figs. 8b, 8c, and S16b). FRAP analysis revealed that G3BP1 condensates
593 formed in crowding conditions have a slower recovery halftime of ~450 seconds. On the
594 contrary, G3BP1 showed a fast ~13-fold (34 seconds) halftime recovery in condensates
595 containing SERF2, indicative of enhanced dynamics and fluidity (Figs. 8d, 8e and S16c). The
596 fast recovery of G3BP1 is also observed in condensates containing SERF2 and rG4 mixture
597 with recovery halftime of ~67 seconds, but slowed to around 360 seconds in the presence HeLa
598 cell total RNA. These observations suggest SERF2 and rG4 structures enhance G3BP1
599 dynamics and fluidity within these tricomponent condensates. SERF2 recovery was also
600 enhanced (halftime ~10 seconds) in these tricomponent condensate mixtures containing G3BP1
601 and rG4 as compared to those condensates that are formed in the absence of G3BP1 which
602 show recovery halftimes of ~50-80 seconds (Figs. 8e, 2e and S3g). This ability of SERF2 or
603 G3BP1 to enhance the condensate fluidity in combination with rG4 aligns with the dynamic
604 nature of stress granules measured in vivo, where SERF2 depletion significantly reduces fluidity
605 of the few remaining granules (Figs. 7e and 7f).

606 **Discussion**

607 There is a large range of liquid-liquid compartments in the cell. Though the precise function
608 of this form of compartmentalization in health and diseases is unclear[98–103], one common

609 feature is that they act to concentrate the molecules present within them. Perhaps the most
610 well-characterized of these compartments is the nucleolus which is important for an efficient
611 ribosome assembly[104]. Another well-characterized compartment is the stress granule which may
612 be involved in RNA or protein storage during stress[100]. Many of the proteins within liquid-liquid
613 phase compartments contain elements of disorder and understanding how they function and
614 interact with other molecules is limited. This is due to a lack of experimentally determined
615 conformational ensembles for disordered proteins making it difficult to study their interaction
616 with RNA species. We have a broad overview of factors affecting LLPT including the importance
617 of multivalency and disorder. However, the lack of detailed structural information leaves both the
618 protein and RNA components of these membraneless compartments often diagramed as vague
619 spaghetti-like lines. A handful of structural models for condensates have been reported
620 recently[90,91], but these have been limited to protein only condensates and our understanding of
621 ribonucleoprotein condensates remains largely unexplored. Our replacing these vague lines,
622 with the structural and biophysical details diagrammed in Figures 3 and 4, helps us understand
623 the process of LLPT and the interactions that take place within these cellular compartments.

624 Here we have obtained detailed structural information for two important components of
625 stress granules namely SERF2 and rG4 both separately and in combination helping us to
626 understand the process of LLPT and the interactions that likely take place within these
627 membraneless compartments. We show that SERF2, [37]is a component of the nucleolus and
628 stress granules and appears to be important for stress granule formation in vivo. We also show
629 that SERF2 specifically binds to rG4s in high throughput and in vitro assays. These non-
630 canonical RNA structures have been implicated in translational repression and transcriptional
631 termination[19], mRNA processing, telomere maintenance, RNA translocation and can effectively
632 influence protein aggregation[19,47,51,105–108]. In findings likely relevant to ours, it has recently been
633 found that stress enhances the number of rG4 present[17]. rG4 are abundant in stress granules,
634 and stress promotes rG4 folding[17]. SERF2 binds to known rG4 structures including some that
635 are known to be present in stress granules with sub-micromolar binding affinities. As SERF2
636 and TERRA rG4 both in isolation are shown to promote amyloid formation[47,109], we speculate
637 that under unfavorable conditions like stress, their relatively strong interactions could suppress
638 their ability to bind and promote aggregation of amyloid substrates, a phenomenon we
639 demonstrated for the SERF2 homologous protein Znf706[27].

640 Despite the lack of direct intermolecular NOEs or side-chain-specific assignments, which
641 remain a major technical challenge in NMR studies of ligand or dynamic protein-rG4

642 complexes[110], our integrative approach combining chemical shift perturbations guided docking
643 and MD simulations offers a structurally and biophysically plausible model[111–114]. This strategy
644 reflects current best practices for mapping challenging or transient interactions involving
645 intrinsically disordered regions and provides experimentally testable hypotheses that we further
646 validated through mutagenesis. Our studies on SERF2 and TERRA rG4 provide a structural
647 framework for understanding protein-RNA condensates and how multiple transient electrostatic
648 interactions can drive oligomer formation. Upon interaction of SERF2 with rG4, a planar G-
649 quartet RNA-protein interaction forms through a quadrupole-like interaction. The SERF2 protein
650 dynamics are constrained upon rG4 binding, and the rG4 structure becomes distorted (Fig. 3).
651 The conformational adaptability of SERF2 enables it to interact with the TERRA rG4 structure
652 which induces destabilization in the TERRA rG4 structure indicative of SERF2 being involved in
653 rG4 unfolding in non-crowding conditions. Several RNA-binding proteins, such as FMRP,
654 nucleolin, CNBP, eIF4A, hnRNPA1, and DHX36, have also been shown to bind rG4 and
655 modulate their folding[20–24,26]. Such interactions seem to be able to regulate their cellular
656 functions, for instance, yeast Rap1 interacts with G4 using planar G-quartets to stabilize the
657 complex[115]. In contrast, the DHX36 helicase unfolds G4[31] by forming a flat non-polar surface on
658 the G-quartet[104], in a way similar to that observed for SERF2 interacting with TERRA rG4.
659 However, under macromolecular crowding conditions, which mimic the dense intracellular
660 environment, the TERRA rG4 structure remains stable even in the presence of SERF2. This
661 suggests that the crowded milieu promotes and stabilizes rG4 folding, an effect consistent with
662 known biophysical principles[116]. Crowding has been shown to favor compact nucleic acid
663 structures, including G4s, by shifting the equilibrium toward folded states due to excluded
664 volume effects and dehydration of the RNA backbone[89,116,117]. These environmental constraints
665 may counteract SERF2-induced destabilization, thus preserving rG4 architecture despite protein
666 binding. Our data therefore suggest that SERF2 may function as an environment-dependent
667 modulator of rG4 structure, where it is capable of perturbing rG4 conformation in dilute
668 environments, but its destabilizing influence is diminished or buffered in crowded conditions.
669 This suggests a potential regulatory switch wherein SERF2's impact on RNA structure may be
670 tightly controlled by the physicochemical environment, such as that found within stress granules.

671 The pattern of charged residues in disordered proteins and the charge interactions have
672 been demonstrated to play an important role in LLPT as described in the simplified stickers and
673 spacers model[88]. Excitingly, we can now start to define stickers at an atomic level based on the
674 frequency with which specific protein residues appear in our condensate model to be interacting

675 with RNA (Fig. 4). Our simulation results suggest that charge interactions and nucleic acid
676 conformational changes may help stabilize the dynamic condensates and facilitate their
677 selective recruitment of specific proteins and nucleic acids during phase transition. We find that
678 distortion of nucleic acid structures, including unwinding or bending, enhance interactions by
679 exposing charged regions, and shielding RNA-RNA interactions, and promoting multivalent
680 intermolecular interactions may be important for phase separation as previously predicted[102].
681 Unexpectantly, we frequently observed ring-like structures that bring multiple SERF2 and
682 TERRA rG4 molecules together through multivalent interactions. In support of a physiological
683 role for macromolecular crowding, we find that the PEG crowding agent, becomes partially
684 excluded from the ring-like condensates while being involved in interactions at the periphery of
685 the rings. As this is the only structure that we are aware of a liquid-liquid phase separated
686 protein-RNA complex, we have no way of knowing how common these ring-like structures will
687 prove to be in phase separated complexes. However, their discrete size represents an attractive
688 mode of interaction that could act to maintain an appropriate degree of exchangeability of the
689 condensate components while avoiding higher-order interactions that could form insoluble
690 complexes like aggregates.

691 Though the current work provided structural insights into SERF2 binding to TERRA rG4
692 and how it distorts the rG4 structure, future work will be necessary to probe if it interacts
693 similarly or differently with in vivo rG4 binding partners, in stress granules or elsewhere.
694 Determining if the ring-shaped structures we have observed are a common feature within
695 condensates likely awaits the development of other protein-RNA pairs that can be analyzed in
696 the same detail as our SERF2-rG4 system. A better understanding of the structure-function
697 relationships present within quadruplexes will aid us in understanding rG4-linked biological
698 functions such as gene regulation and stress granule formation.

699 **Methods**

700 **Cell Culture, Treatment, and Transfection**

701 U2OS cells (human) were purchased from Sigma-Aldrich (Cat. no. 92022711-1VL), and HEK
702 293T/17 cells (ATCC® CRL-11268™) used in this study are derived from human embryonic
703 kidney tissue. The HeLa Kyoto eGFP-FUS cells (human) were obtained from Dr. Antony Hyman
704 at Max Planck Institute of Molecular Cell Biology and Genetics, Germany. All cells were cultured
705 in Dulbecco's Modified Eagle Medium (Fisher Scientific, 11-995-073) supplemented with 10%
706 heat-inactivated fetal bovine serum (Sigma, F4135) and 1X Penicillin-Streptomycin-Glutamine

707 (Fisher Scientific, 10-378-016). Stress treatments were performed on cells with vehicles alone
708 or with different stress inducers. Cells were then treated with a final concentration of 0.4M
709 sorbitol or 0.5 mM sodium arsenite to induce osmotic or oxidative stress, respectively. For ER
710 stress, the cells were additionally treated with a final concentration of 2 mM dithiothreitol for 30
711 minutes in a culture medium at 37 °C. Cells were treated with 10 μM of the proteasome inhibitor
712 MG132 (Cas No- 133407-82-6, Sigma) diluted in culture medium and incubated for 30 minutes
713 at 37 °C to inhibit protease function. For mitochondrial stress, cells were treated with 75 mM
714 NaN_3 for 30 minutes at 37 °C. For heat shock, cells were incubated at 43 °C for 1 hour. For
715 plasmid transfection, cells were transiently transfected using Lipofectamine™ LTX and PLUS™
716 reagent (Fisher Scientific, 15338030) according to the manufacturer's instructions. For
717 knockdown, 13 μM of control (Horizon Discovery, D-001206-13-20) or SERF2 (Horizon
718 Discovery, M-016317-01-0010) siRNAs were transfected using RNAiMAX reagent (Thermo
719 Fisher, 13778150). After 48 hours of incubation, the transfected cells were harvested for
720 western blot analysis and RT-qPCR.

721 **Immunofluorescence**

722 Cells were grown on coverslips with the conditions described above and then fixed with 4%
723 paraformaldehyde (Electron Microscopy Sciences, 157-8-100) in 1X DPBS (Fisher Scientific,
724 14-190-144) at room temperature for 10 minutes. The fixed cells were washed three times with
725 1X DPBS and then permeabilized with 0.1 % Triton X-100 at room temperature for 30 minutes.
726 The permeabilized cells were next blocked at room temperature for 1 hour using UltraCruz™
727 blocking reagent (Santa Cruz, sc-516214), followed by incubation with primary antibody
728 solutions for 1 hour at room temperature. Primary antibody solutions were prepared in
729 UltraCruz™ blocking reagent using the following dilution factors: 1:200 rabbit anti-SERF2
730 (Proteintech, 11691-1-AP), 1:1000 mouse anti-G3BP1 (BD Biosciences, 611127), 1:1000
731 mouse anti-Fibrillarin (Boster Bio, M03178-3), 1:100 mouse anti-Nopp140 (Santa Cruz, sc-
732 374033), 1:100 mouse anti-USP10 (Santa Cruz, sc-365828), 1:200 mouse anti-FUS (Thermo
733 Fisher, 50-554-337), 1:500 mouse anti-BG4 (Absolute antibody, Ab00174-1.1), 1:100 mouse
734 anti-TIA1 (Santa Cruz, sc-398372), and 1:100 mouse anti-eIF2α (Santa Cruz, sc-133132). The
735 cells were then washed three times with 1X DPBS and incubated in secondary antibody
736 solutions in the dark at room temperature for 1 hour. Secondary antibody solutions were
737 prepared in a blocking reagent using a 1:1000 dilution, goat-anti-rabbit secondary Alexa Fluor
738 Plus 488 (Thermo Scientific, A32731), and goat-anti-mouse secondary Alexa Fluor 647 (Thermo
739 Scientific, A21235), goat-anti-mouse secondary Alexa Fluor 488 (Invitrogen, A11001). After

740 three washes with 1X DPBS, the cells were stained with 0.25 µg/mL DAPI (Thermo Scientific,
741 D1306) in 1X DPBS for 3 minutes. The cells were washed three more times with 1X DPBS, air
742 dried, and mounted with ProLongGlod™ antifade reagent (Cell Signaling, 9071S) on glass
743 slides. After 24 hours of mounting, the slides were sealed with nail polish and imaged using a
744 Leica SP8 confocal or Thunder™ microscope.

745 **Oligonucleotide synthesis**

746 High-performance liquid chromatography (HPLC) purified unlabeled, TAMRA- and 6-FAM
747 fluorescent labeled RNA nucleotides were either purchased from Integrated DNA Technology
748 (IDT) or synthesized at Slovenian NMR Center using previously described methods[118,119].
749 Different length TERRA RNA sequences that include TERRA10, TERRA12, TERRA23 and
750 TERRA60; 4-repeat GGGGCC, six-repeat UGGGGU, short poly RNA, HIV-1 TAR RNA,
751 CCCCGG and mutated TERRA repeat RNA were synthesized from IDT. Random length polyA
752 was purchased from Sigma. All the oligonucleotides were suspended in nuclease-free water or
753 buffer prepared using nuclease-free water. All oligonucleotides were desalted using a 3 kD
754 cutoff filter (Amicon ®Ultra 0.5 mL) by treating with 20 mM Tris-HCl, pH >11, heated at 95 °C for
755 3 minutes following 10-time buffer exchange in 20 mM Tris-HCl, pH 7.4. The oligonucleotide
756 concentration was next measured using the extinction coefficients obtained using the IDT
757 OligoAnalyzer™ tool. G4 sequences were folded by cooling the samples prepared in KCl
758 (20mM sodium phosphate (referred to as NaPi hereafter), pH7.4 and 100 mM KCl) or LiCl
759 buffers (20 mM Tris-HCl, pH7.4 and 100 mM LiCl) using a thermocycler with 1 °C/min. The
760 folded G4s were stored at 4 °C for immediate use or at -20 °C for future use. All the chemicals
761 and reagents used in this study were commercially purchased with >98% purity and used
762 without further purification.

763 **Purification of recombinant proteins**

764 The plasmids, containing a codon-optimized human SERF2 gene, were synthesized
765 commercially by GenScript and subcloned to a pER28a-SUMO vector as reported
766 elsewhere[42].[39] The expression and purification of the G3BP1, SERF2 wild-type or different
767 lysine-alanine mutant proteins was like that previously reported for yeast SERF[42]. Briefly, the
768 plasmids were transformed into competent BL21(DE3) *Escherichia coli* cells, incubated
769 overnight in 10 mL of cell culture medium, and then transferred to freshly prepared 1 L of PMEM
770 medium containing 50 mg/L Kanamycin. The cells were further grown at 37 °C, under shaking,
771 until the OD_{600} reached 1.0. Then they were transferred to a 20 °C shaker for 1 hour, protein

772 expression was induced by adding 0.1 mM IPTG, and the cells were incubated overnight at 20
773 °C with continued shaking.
774 Isotope-labeled 15N and 15N-13C SERF2 proteins for NMR studies were produced by
775 growing cells in M-9 minimal media supplemented with 100% 15N NH_4Cl (1g/L) for the 15N
776 labeling or 15N NH_4Cl and D-Glucose-$13C_6$ (4 g/L) for the 15N-13C labeling. Cells were
777 subsequently harvested and lysed by sonication in ice-cold lysis buffer (40 mM Tris-HCl pH 8.0,
778 10 mM NaPi, 400 mM NaCl, 20 mM imidazole, 10% glycerol, 1 tablet of cOmplete protease
779 inhibitor (Roche), and 1.25 µg/mL DNase I (Roche). The lysate was centrifuged at 36,000 g for
780 30 minutes and the supernatant was passed through a HisTrap column (Cytiva, 17-5248-02).
781 Lysis buffer containing 0.5 M imidazole was used to elute the His-SUMO tagged SERF2
782 proteins, the elution was then supplemented with beta-mercaptoethanol to a final concentration
783 of 5mM. The sample was next incubated overnight at 4°C with 10 µL of homemade SUMO
784 protease 6His-ULP1 for His-SUMO cleavage. The digestion mixture was dialyzed in 40 mM
785 Tris-HCl pH8.0 and 300 mM NaCl overnight at 4°C, using a 3.5 kD cutoff dialysis membrane
786 (Repligen, 132724). The dialyzed proteins were run through a 5 mL HisTrap column to remove
787 the cleaved His-ULP1 and the His-SUMO. The flow-through SERF2 protein was further purified
788 by an ion exchange HiTrap SP column (Cytiva, 17-5161-01) using buffer A (50 mM NaPi and
789 125 mM NaCl pH6.0) and buffer B (50 mM NaPi and 1 M NaCl, pH 6.0). A final purification of
790 the ion-exchange purified SERF2 protein was conducted using a size-exclusion
791 chromatography column Hiload75 (Cytiva, 28989333) in 20 mM NaPi pH7.5 and 150 mM NaCl
792 or 40 mM HEPES pH 7.5 and 100 mM NaCl. The protein samples used in the biophysical and
793 biochemical studies were prepared in indicated buffers as needed via buffer exchange using a 3
794 kD cutoff filter (EMD Millipore, UFC503024). The SERF2 protein concentration was determined
795 using a Pierce TM BCA assay calibrated with the SERF2 A51W mutant serving as a standard. A
796 similar expression and purification protocol, as described above, was used to produce different
797 SERF2 cysteine and alanine mutants.

798 For the high throughput screening assay SERF2 protein was prepared containing GST and
799 streptavidin binding peptide (SBP) tags. SERF2 was cloned into a pGEX-6P-1 construct
800 containing GST and SBP tags and transformed into Rosetta-competent cells. Cell cultures were
801 grown in LB at 37 °C until OD_{600} reached ~0.6, thereafter induced with 0.5 mM IPTG overnight
802 at 16 °C. Cells were harvested at 4000×g for 13 minutes and resuspended in lysis buffer (1%
803 triton x-100, 5 mM DTT, 4 mM $MgCl_2$, 200 mM NaCl, 20 mM HEPES and 1 tablet/L culture
804 Pierce™ Protease Inhibitor Mini Tablets EDTA-free). Lysates were sonicated and incubated

805 with 3 units/L culture RQ1 DNAse (Promega) and 500 units/L culture Benzonase Nuclease
806 (Sigma-Aldrich) for 15 minutes at room temperature. Following incubation, lysates were
807 centrifuged at 36000 × g for 30 minutes, and the supernatant was passed through a 0.45 µm
808 filter. GST-SBP-SERF2 was purified using Pierce™ Glutathione Agarose (Thermo Fisher
809 Scientific). Protein was further cleaned up using HiTrap® Heparin High Performance (Sigma
810 Aldrich) and concentration was assessed using Pierce 660 nm assay (Thermo Fisher Scientific).
811 Protein purity was determined by SDS-PAGE and Coomassie blue staining.

812 **Protein labeling**

813 Cy-5 (Cytiva, PA25031) or Alexa Fluor 488 (AF488) (Invitrogen, A10254) labeling of protein
814 was done by incubating 200 µM of SERF2 (T2C) or G3BP1 with a 10-molar excess of C_5
815 maleimide mono-reactive dye in 20 mM Tris-HCl pH7.4 and 100 mM KCl buffer overnight, at 25
816 °C (SERF2) or at 4 °C (G3BP1), under continuous shaking at 300 rpm. The free excess label
817 was then removed by passing the sample through a PD-10 desalting column in a dark room.
818 The samples were concentrated using a 3 kD Amicon Ultra-15 Centrifugal Filter Unit (EMD
819 Millipore, UFC800324), and any residual free dyes were removed by resuspending the proteins
820 in a working buffer after centrifugation at 7000 × g (8 times) for 15 minutes using an Amicon
821 ®Ultra 0.5 mL 3 kD cutoff filter.

822 **High-throughput screening assays**

823 *RNA bind-n-seq assay and analysis*

824 For the RNA bind-n-seq assay, a single-strand DNA library containing a randomized 40
825 nucleotide region was obtained from IDT, gel-purified and the RNA library was prepared
826 following a previously described method[48] using a T7 promoter and in vitro transcription. A
827 second pool of RNA was made by replacing the guanines with 7dG (Trilink) to eliminate rG4
828 folding while preserving the sequence. Residual DNA was removed with DNase I (Promega)
829 followed by a phenol-chloroform extraction. The RNA was resolved in a 6% TBE-Urea gel, the
830 expected size band was excised, and gel purification was conducted as previously described[48].

831 The RNA bind-n-seq method was modified from a previous study as described below[48].
832 Briefly, 60 µL of recombinant GST- SERF2 at different concentrations (250 nM and 50 nM) in
833 binding buffer (25 mM Tris-HCl pH 7.5, 150 mM KCl or LiCl, 3 mM $MgCl_2$, 500 µg/mL Ultrapure
834 BSA and SUPERase-In RNase Inhibitor) were equilibrated with 60 µL of pre-washed (binding
835 buffer) magnetic beads (Dynabeads MyOne Streptavidin T1, Invitrogen) for 30 minutes at 4 °C.
836 RNA pools were heated in the presence of 150 mM KCl or LiCl for 5 minutes at 100 °C and
837 cooled down to room temperature for at least 10 minutes. 60 µL of 3 RNA pools (KCl, LiCl, and

838 7dG (in LiCl)) were then mixed with GST-SBP-SERF2 and further incubated for 1 hour at 4 °C.
839 The final concentrations of protein in the binding reaction were 250 nM and 50 nM and the
840 RNA's concentration was 1 μM. The protein-RNA complexes were washed with buffer (25 mM
841 Tris-HCl pH 7.5, 150 mM KCl or LiCl, and SUPERase-In RNase Inhibitor (Invitrogen)). The
842 complexes were magnetically isolated, and RNA was eluted with buffer (4 mM biotin and 25 mM
843 tris-HCl pH 7.5) for 30 minutes at 37 °C. The elution was performed twice, the eluates were
844 combined, and RNA was purified by a phenol-chloroform method[48]. Before reverse transcription,
845 RNAs were heated for 5 minutes at 100 °C in the presence of 150 mM LiCl to facilitate rG4
846 unfolding. Following reverse transcription, the samples were prepared for sequencing as
847 described elsewhere[48].
848 We performed sequence enrichment analysis as previously described[48,120]. Briefly, we
849 analyzed k-mer (k=6) enrichments to determine 'R' values. R was defined as the frequency of a
850 given k-mer in the protein-associated pool divided by the frequency of that *k*mer in the input
851 pool. To specifically determine the enrichment of rG4s, we searched for strong G4 patterns ($G_{(3-}$
852 $_{6)}N_{(0-7)})_4$. As a control, we removed all sequences that matched the G4 pattern but still had
853 greater than 8 guanines in the randomized region. The G4 pattern analysis was performed on
854 the randomized region plus the adapters (as described in Reference [48] for RNA structure
855 analysis), as they are part of the RNAs presented to the protein in the binding reaction. The
856 enrichment of each pattern was the frequency of the pattern in the protein-bound sample
857 divided by its frequency of the input pool.
858 *FOREST assay*
859 Library-1 from the previously published paper[49] which includes 1800 pre-miRNA and 10 rG4
860 sequences was used for FOREST screening. A detailed method for oligonucleotide template pool,
861 DNA barcode microarray design, in vitro transcription, RNA fluorophore labeling, hybridization,
862 and microarray scanning has been provided in the supporting information. Briefly, the templates
863 used were synthesized by oligonucleotide library synthesis and the size was limited to 170
864 nucleotides (OLS, Agilent Technologies). The in vitro transcribed RNA structure library was
865 labeled with Cy5, at the 3' end, to detect and quantify RNA probes on a microarray. Lastly, the
866 library was prepared in K^+ folding buffer (10 mM Tris-HCl pH 7.5, 100 mM KCl), heated at 95 °C
867 and cooled to 4 °C at a rate of −6 °C/s on a ProFlex Thermal Cycler (Thermo Fisher Scientific) to
868 allow for G4 folding.
869 His-tagged SERF2 recombinant protein was used for the FOREST binding assay. For this
870 purpose, the target protein (100 pmol of SERF2), 20 μL of TALON magnetic beads (Clontech),
871 and 1 μg of the refolded RNA structure library were mixed in 1 mL of protein-binding buffer (10 mM

872 Tris-HCl pH 7.5, 100 mM KCl, 10% glycerol, and 0.1 μg/μL BSA). A mixture containing no protein
873 was also prepared as a control. The mixtures were incubated on a rotator at 4 °C for 30 min and
874 washed three times with the protein-binding buffer. Then, 200 μL of elution buffer was added to
875 the magnetic beads, and the mixture was heated at 95 °C for 3-min. The RNA was collected from
876 the supernatant by removing the magnetic beads. The RNA structure library was extracted with
877 phenol and chloroform and an ethanol precipitation was conducted for RNA purification. The
878 enriched RNA sample was hybridized for microarray analysis as detailed in the supporting
879 methods. To determine the protein-binding intensities of each RNA probe, we subtracted the
880 fluorescence intensities of the negative control sample (samples without protein) from those of
881 the enriched protein samples. To account for any undesired interactions with the barcode region,
882 we calculated the average fluorescence intensity of each structure by averaging the intensities of
883 the RNA probes that had the same RNA structure but differing barcodes.

884 **NMR experiments**

885 Unless indicated, all NMR measurements were done with protein or protein-RNA mixture
886 samples that had been suspended in 20 mM NaPi and 100 mM KCl (pH 7.4) containing D_2O at
887 4 °C. Isotopically unlabeled protein and oligonucleotide (TERRA10, TERRA12, and TERRA23)
888 samples were used to collect 1D proton NMR spectra with a recycle delay of 1 second.
889 Multidimensional NMR data were collected using uniformly isotope-labeled 15N and 15N-13C
890 SERF2 samples. A series of 2D and 3D NMR experiments, that included 15N-HSQC,13C-
891 HSQC, HNCO, HNCA, HNCOCA, HNCACO, CBCA(CO)NH, HNCACB, 15N-HSQC-NOESY
892 and 15N-HSQC-TOCSY, were conducted for backbone assignments of 1.2 mM proteins
893 dissolved in 20 mM NaOAc-d_3 with pH 5.5 and 8% D_2O at 4 °C. 2D 15N-1H HSQC experiments
894 were recorded for 100 μM 15N labeled SERF2 titrated with an increasing concentration of
895 different length TERRA (10, 12, 23 or 60) rG4 for binding site analysis. The effect of pH and
896 temperature on SERF2's chemical exchange was studied by acquiring the 2D spectrum at 4
897 and 37 °C as demonstrated elsewhere for disordered proteins[72], in buffers containing either 20
898 mM NaPi and 100 mM KCl (pH7.4) or 20 mM d_3-NaAc (pH 5.5). 15N relaxation NMR
899 measurements, including heteronuclear NOE, T1, and T2, were done using 200 μM 15N SERF2
900 mixed with and without 100 μM of TERRA rG4 at 4 °C in 20 mM NaPi and 100 mM KCl (pH7.4).
901 The relaxation delays used for the T1 experiments were 20, 50, 90, 130, 200, 320, 450, 580,
902 750, 900, 1200, and 2005 ms. T2 relaxation delays used were 16.96, 50.88, 84.80, 135.68,
903 169.60, 220.48, 305.28, 373.12, 457.92, 542.72, 678.40, and 1356.80 ms. 1D proton NMR
904 experiments were carried out to map the chemical shift changes in the TERRA rG4 imino

905 protons (200 μM TERRA12 or 100 μM TERRA10) at increasing SERF2 concentration.
906 Saturation transfer difference (STD) NMR spectra were recorded with 512 scans and a
907 saturation time of 4 s. On-resonance saturation was applied at selected well-resolved chemical
908 shift regions, and off-resonance saturation was applied at -40.0 ppm to serve as a control. NMR
909 samples used for STD NMR data collection consisted of 100 μM TERRA12 rG4, 100 μM
910 SERF2, or a 1:1 mixture of SERF2 and TERRA12 rG4, prepared in 20 mM NaPi, 100 mM KCl,
911 pH 7.4. To test for direct magnetization transfer and to exclude potential contributions from
912 solvent exchange, we selected isolated and well-resolved imino proton peaks from TERRA12
913 rG4 and an isolated backbone amide proton peak (~8.5 ppm) from SERF2. These peaks were
914 first saturated in the individual samples (RNA or protein alone) with off-resonance irradiation
915 serving as the reference. The same peaks were then saturated in the protein-RNA complex
916 sample under identical conditions. Differences in saturation transfer between the free and mixed
917 states were analyzed to verify that the observed transfer is due to direct protein-RNA
918 interactions and not mediated by solvent exchange. NMR data were collected on a Bruker 800
919 MHz spectrometer equipped with a triple resonance cryoprobe, while 15N relaxation data were
920 collected on a Bruker 600 MHz NMR spectrometer equipped with a triple resonance cryoprobe.
921 The NMR data were processed using Bruker's Topspin 4.1.4 and spectra assignment and
922 analysis was done using an NMRFAM-Sparky 1.47.

923 **Fluorescence polarization and anisotropy assay**

924 Fluorescently labeled 6-FAM RNA or AF488 labelled SERF2 probes were prepared in
925 nuclease-free water containing 20 mM NaPi (pH 7.4) and 100 mM KCl. Fluorescence
926 polarization assays were done either using 20 nM RNA 6-FAM rG4 or non rG4 RNA probe
927 mixed with increasing concentrations of SERF2, ranging from low nM to high μM, or using 200
928 nM AF488 labelled SERF2 with increasing concentration of structured or unstructured RNA. All
929 fluorescence polarization (FP) assays were done in samples dissolved in 20 mM NaPi and 100
930 mM KCl (pH7.4). In addition to unlabeled TAR RNA binding to AF488 SERF2, FP assay was
931 also done using TAMRA-TAR RNA to reveal SERF2 binding specificity for rG4 structured RNA.
932 FP or gel-shift EMSA binding assay was also carried out to test the binding specificity of SERF2
933 to two rG4 sequences derived from the 3' UTR regions of *Mark2* and *Stxbp5* genes. The sample
934 mixtures for FP assay were incubated for 30 minutes at room temperature. Fluorescence
935 polarization data were recorded on a TECAN Infinite M1000 microplate reader at 25 °C with the
936 excitation/emission wavelengths set at 470/530 nm, for FAM or AF488 probes; 530/575 nm for
937 TAMRA probes. Fluorescence anisotropy measurements were done by titrating SERF2 to 200

938 nM of Cy3 (excitation/emission, 550/570 nm) or FAM (excitation/emission, 493/517 nm) labeled
939 oligonucleotides, in a 1 mL quartz cuvette (Hellma,101-QS), using a Cary Eclipse
940 spectrofluorometer (Agilent) at 25 °C. Slit bandwidths were set at 5 nm and 10 nm for excitation
941 and emission, respectively. The binding constant (K_D) was calculated from the change in
942 polarization or anisotropy values, in GraphPad Prism 9.5.1, using non-linear regression for
943 curve fitting with a one-site specific binding model.

944 **Circular dichroism spectroscopy**

945 The secondary structure of the G4s (15 μM) or SERF2 (50 μM), suspended in 20 mM NaPi
946 and 100 mM KCl (pH7.4), were studied by collecting circular dichroism (CD) spectra using a
947 JASCO J-1500 spectropolarimeter. For folding analysis, the CD spectra of SERF2 were
948 collected at different temperatures (4, 25, and 37 °C). CD titration was done for 20 μM
949 TERRA23 rG4 against an increasing SERF2 concentration (10 to 100 μM) to monitor rG4
950 secondary structure change at 25 °C. The buffered CD spectrum was subtracted from the
951 average CD spectrum obtained from 8 scans. The structural change in 40 μM SERF2 mixed
952 with or without 20 μM TERRA12 rG4 was measured, that corresponds to 2:1 protein to RNA
953 ratio used in NMR relaxation experiments. CD spectra of rG4 alone were used as a baseline
954 and subtracted from the spectra of the SERF2-rG4 complex to monitor secondary structural
955 change in SERF2.

956 **In vitro liquid-liquid phase transition assay**

957 16-well Culture-Well chamber slips (Grace Bio-Labs) or 384-well plates (Cellvis, P384-
958 1.5H-N), pre-treated with 5% (w/v) Pluronic™ F-127 (Sigma, P2443) overnight, were used to
959 study in vitro phase separation. The well chambers were washed three times with 20 mM NaPi,
960 100 mM KCl, pH 7.4 buffer and air dried. 50 or 20 μl aliquots of reaction sample mixtures were
961 incubated for 30 minutes at room temperature in various conditions (varying protein, total RNA
962 from HeLa cells, poly-ribonucleotides, and various rG4 concentrations) and a buffer with or
963 without 10% of PEG8000 (Sigma, P5413). The effect of salt and PEG on LLPT was tested for
964 50 μM SERF2 mixed with equimolar TERRA12 or TERRA23. LLPT was measured in a non-
965 crowding (20 mM NaPi, 100 mM KCl, pH 7.4) or crowding condition (20 mM NaPi, 100 mM KCl,
966 pH 7.4 containing 2.5% or 5% PEG8000) at varying G3BP1, HeLa cells extracted total RNA,
967 SERF2, unstructured homo polyA RNA, structured HIV-1 TAR RNA and rG4 concentrations. For
968 fluorescence imaging, 1/200th fluorescence labeled (AF488, 6-FAM or Cy-5) protein/rG4
969 samples were mixed into unlabeled sample mixtures. Sample imaging and FRAP
970 measurements were done, using a Nikon Ti2-E motorized, inverted microscope. This

971 microscope is controlled by NIS Elements software and contains a SOLA 365 LED light source
972 and a 100X oil immersion objective. Recovery half-life analysis was done using GraphPad
973 Prism and the image processing was done using Fiji ImageJ.

974 **Mutational analysis**

975 5 μM of TERRA23 rG4s, dissolved in NaPi buffer, were mixed with equimolar
976 concentrations of wild-type and different lysine-alanine mutant SERF2 and incubated for 30
977 minutes at room temperature. Gel shift mobility assays were next performed by loading a 5 μM
978 TERRA rG4 sample mixture, containing SERF2 and 20% glycerol, to a 4-20% TBE gel
979 (Invitrogen, EC6225). For LLPT assay, 50 μM of TERRA rG4s was dissolved in 20 mM NaPi,
980 100 mM KCl, pH 7.4 containing 10% PEG8000 and mixed with different lysine-alanine SERF2
981 mutants at equimolar concentration. The samples were incubated for 30 minutes at room
982 temperature and DIC images were taken on a Nikon Ti2-E motorized, inverted microscope.

983 **Size-distribution analysis**

984 *Size-exclusion chromatography*

985 20 μM of TERRA23 rG4s, dissolved in NaPi buffer, were mixed without or with 40 μM
986 SERF2 and incubated for 30 minutes at room temperature. The mixture was then injected into a
987 Superdex 200 Increase 10/300 GL size-exclusion chromatography column (Cytiva, 28-9909-44).

988 *Analytical Ultra Centrifugation (AUC)*

989 The AUC measurements were done for a SERF2 (9.4 μM) and TERRA12 rG4 (4.7 μM)
990 sample mixture that was dissolved in 20 mM NaPi, 100 mM KCl (pH 7.4) at 22 °C. AUC
991 measurements were done at 260 nm, where SERF2 has no absorption, and with an intensity
992 mode of 42,000 rpm. 420 μL samples were loaded into a two-channel epon-charcoal
993 centerpiece, with a 1.2 cm path length, in an An60Ti rotor of a Beckman Optima XI-I AUC. The
994 data were analyzed with Ultrascan III software (version 4) and the LIMS server using the
995 computing clusters available at the University of Texas Health Science Center and XSEDE sites
996 at the Texas Advanced Computing Center.

997 **Structure calculation and all-atom MD simulations**

998 SERF2 backbone assignments were done using standard 3D experiments including, HNCA,
999 HNCACB, CBCAcoNH and HNCO experiments along with the 2D HSQC collected at 4 °C. The
1000 side chain assignment was carried out using a 3D ^{15}N-resolved [^{1}H,^{1}H] TOCSY and a 2D
1001 TOCSY measured with a mixing time of 100 ms. The data were processed using NMRFAM-
1002 Sparky[121] and the dihedral angles were obtained utilizing the backbone carbon chemical shifts

1003 by the TALOS-N program[74]. Distance restraints for the structure calculation were obtained from
1004 3D ^{15}N-resolved [^{1}H,^{1}H] NOESY and a 2D NOESY measured with a mixing time of 200ms and
1005 150ms respectively. A total of 494 (intra - 205, short - 220 and medium - 69) NOE distance
1006 constraints and 40 dihedral angles and 24 hydrogen bonds were used for the multiple-state
1007 ensemble calculation in CYANA 3.98.15. The NOE's are calibrated using the standard
1008 calibration in Cyana where it uses the CH2 distances as 1.7 angstrom and the rest is calibrated
1009 with respect to that. One hundred initial conformers were calculated using 10,000 torsion-angle
1010 dynamic steps and energy minimized. The 20 conformers, with the lowest target function
1011 values, were used to represent the calculated SERF2 structure. The 20 conformers were
1012 subjected to refinement using Crystallography & NMR System or CNS constrained energy
1013 minimization in an explicit water model. The CYANA-generated SERF2 PDB coordinates, NOE
1014 constraints, dihedral angles, and hydrogen bond constraints were input to CNS for structure
1015 refinement. The 20 coordinates of the energy minimized structures are deposited to the protein
1016 data bank as 9DT0 and the chemical shifts along with the restraints are deposited to the BMRB
1017 as 52141.
1018
1019 The MD simulations were performed using GPU nodes running in parallel. To gauge the
1020 intrinsic structural dynamics of these SERF2-TERRA rG4 systems, we performed all-atom MD
1021 simulations, that were defined in a structure-based balanced forcefield, i.e., CHARMM36m[122]
1022 using GROMACS 2022.4[123]. The SERF2 structure was initially subjected to a 100 ns all-atom MD
1023 simulation in CHARMM36m force-field at pH 7.4 and 0.1M KCl for structural refinement. All-atom
1024 MD simulations for 100 ns production run using CHARMM36m was initially done to test the
1025 structural stability of TERRA12 (PDB ID: 2KBP) and TERRA10 (PDB ID: 2M18) in solution
1026 containing 0.1M KCl. The TERRA10 structure was next considered for protein-RNA complex
1027 simulation. The SERF2-TERRA rG4 complex structure was next built, using the HADDOCK
1028 program[81], by parsing the NMR distance restraints obtained from chemical shift perturbations and
1029 saturation transfer NMR data analysis. While side-chain chemical shifts or NOEs can further refine
1030 models, HADDOCK was explicitly designed to handle sparse experimental data, including
1031 backbone-only constraints, which have been shown to be sufficient for generating reliable models
1032 of protein-RNA and protein-protein complexes. The 15N-1H CSP-guided docking in principle can
1033 generate reasonable starting model[124], demonstrated in protein-protein or protein-RNA
1034 systems[111–114], that can further be subjected to long-timescale all-atom MD simulations, allowing
1035 to probe the complex's stability and interaction mechanism in atomic detail. A set of ambiguous
1036 active site SERF2 residues (T3, N5, R7, R11, Q12, K16, S19, S21, K23, A33, Q46, K47, A51,

1037 N52, K55 and E56) and TERRA rG4 guanines (G5, G9, and G10) obtained from chemical shift
1038 perturbation and saturation transfer difference NMR experiments were provided to the HADDOCK
1039 program to allow it to build the initial model structure of the complex. The energetically best cluster
1040 complex structure was next used for refinement and dynamic analysis. The complex structure
1041 was solvated in the TIP3P water model in an orthorhombic water box. These model systems were
1042 electro-neutralized using 100 mM KCl followed by energy minimization, using the steepest
1043 descent algorithms in less than 5000 steps, to remove steric clashes. The energy-minimized
1044 systems were then subjected to two-step equilibrations using NVT and NPT ensembles for 50 ns
1045 and keeping the temperature to 37 °C and pressure to 1 bar using a V-rescale thermostat and
1046 Berendsen barostat. These equilibrated systems were subjected to a final production MD of 0.5
1047 μs. All the bonds involving hydrogen atoms were restrained using the SHAKE algorithm. The
1048 atomic coordinates of each system were saved every 100 ps resulting in 5000 snapshots for post-
1049 dynamics analysis.

1050 **MD simulation of SERF2-TERRA RNA G4 quadruplex liquid-liquid phase separation**

1051 We designed four densely packed, multi-component systems, containing 1:1 or 1:2 o4 2:1
1052 protein:RNA molecules, to study SERF2 and TERRA10 rG4 condensates using atomistic MD
1053 simulation. We designed two 1:1 MD systems, a large one comprising of 30 protein and rG4
1054 molecules, and a small one that include 6 protein and rG4 molecules. The 1:2 and 2:1 system
1055 consists of 6 or 12 SERF2 or RNA molecules. The MD systems comprised randomly distributed
1056 multiple copies of SERF2, TERRA10, and 10% PEG8000 molecules solvated in a TIP3P water
1057 model in an orthorhombic water box with the dimensions of 20nm×20nm×20nm. To mechanize
1058 the complex molecular assembly and preparation of the input files for the MD simulation under
1059 periodic boundary conditions, we used a Multicomponent Assembler in CHARMM-GUI[125]. The
1060 small 1:1 system consisted of 6 SERF2, 6 TERRA, 60 units of PEG, 234 K^+ and 60 Cl^-, and 23185
1061 water molecules. The 1:2 system has 6 SERF2, 12 TERRA10 rG4, 97 units of PEG, 506 K^+, 134
1062 CL^- and 35961 water; and the 2:1 system contains 12 SERF2, 6 TERRA10 rG4, 98 units of PEG,
1063 230K^+, 134 CL^-, and 36793 water molecules. The large system contained 30 units of SERF2, 30
1064 TERRA10 rG4, 150 PEG, 1266 K^+ and 396 Cl^-, and 216499 water molecules. The input files for
1065 the MD simulations were generated using a CHARMM36m force field compatible with
1066 GROAMCS. The same MD simulation procedure involving energy minimization, two-step
1067 equilibration, and production was adopted, as described above, for both densely packed systems.
1068 The small system was subjected to 1 μs production MD, while the large systems were set to 0.5

1069 μs. The trajectories of both systems were harvested at 100 ps intervals for post-dynamics
1070 analysis.

1071 **Post-dynamic analysis of Molecular Dynamic trajectories**

1072 We employed the Gromos Clustering Algorithm method, with a cut-off of 0.2 nm, to obtain
1073 a representative conformation of the entire trajectory with maximum population. The hydrogen
1074 bond occupancy of each residue was computed using the hydrogen bonding analysis toolkit of
1075 Visual Molecular Dynamics (VMD 1.9.4a53). An angle cutoff of 30° and donor-acceptor distance
1076 of 3.5 Å was set to compute hydrogen bonds between protein and nucleic acids. The residue
1077 pairs (amino-acid and nucleic acid) having higher percentages of hydrogen-bond occupancy were
1078 used for the generation of interaction plots. Structural representations were created using PyMOL
1079 and the BIOVIA Discovery Studio Visualizer. To probe the diffusion of SERF2 molecules, in the
1080 small and large MD systems, the mean square displacement of individual SERF2 molecules was
1081 calculated. The average end-to-end distance between the nitrogen atoms of T2 and A51 in
1082 SERF2 was computed using the 0.5 μs MD trajectory. The mean square displacement analysis
1083 was carried out using gromacs *gmx msd* command. Cluster analysis, PEG molecule distribution
1084 within 6 Å from protein or RNA, time-resolved Hoogsteen H-bonds (donor atoms: guanine N1, N2
1085 and O2; acceptor atoms: N7 and O6) formation and contact map analysis in the large MD system
1086 comprising 30 SERF2 and 30 TERRA10 rG4 was performed using MDanalysis program[126].

1087 **Single-molecule fluorescence microscopy and droplet fusion by optical tweezers**

1088 Single particle tracking (SPT) and fluorescence resonance energy transfer (FRET)
1089 experiments were conducted using a double cysteine SERF2 mutant (A2C and A51C) labeled
1090 with Cy3 and Cy5 fluorophores. 200 μM of double-mutant SERF2 (A2C, A51C) was fluorescently
1091 labeled by incubating it with 5x molar excess of Cy3 and Cy5, applying the same protocol
1092 described above in the section 'Protein labeling'. The occupancy of the Cy3 and Cy5 in each
1093 SERF2 molecule was not determined because the SPT and FRET measurements do not rely on
1094 a uniform 1:1 Cy3:Cy5 labeling for studying SERF2-TERRA12 rG4 condensates. For single-
1095 molecule experiments, droplets were prepared by mixing 100 μM unlabeled SERF2 and
1096 TERRA12 rG4 containing a picomolar concentration of Cy3-Cy5 SERF2 or diluted as needed to
1097 obtain a sparse concentration where single molecules are distinguishable for SPT. The sample
1098 chamber was prepared by mounting #1.5 coverslips on a standard glass microscope, using
1099 double-sided tape placed about 5 mm apart as spacers. The chamber was coated with ~15 μL of
1100 0.1 mg BSA (5 minutes) followed by rinsing thrice with NaPi buffer containing 10% PEG8000.
1101 SERF2-TERRA12 rG4 droplets were then injected into the sample chamber and droplets were

1102 allowed to settle for ~20 minutes. An O_2-scavenging system (1 mM trolox, 2.5 mM protocatechuic
1103 acid (PCA), and 7 μg/mL protocatechuate-3,4-dioxygenase (PCD)) prepared in NaPi, PEG8000
1104 buffer was then applied to the sample chamber immediately before performing the SPT and FRET
1105 measurements at room temperature. Fluorescent images were acquired with total internal
1106 reflectance fluorescence (TIRF) illumination (using an Olympus CellTIRF module). Acceptor and
1107 donor fluorescent images were collected simultaneously on dual Andor Ultra EMCCD cameras at
1108 frame rates between 10Hz and 100Hz. The data were analyzed using an in-house MATLAB script
1109 to obtain the diffusion constants via a linear fit of the mean squared displacement for each particle
1110 and FRET efficiency.

1111 The fusion of SERF2 and TERRA rG4 droplets was studied using a bespoke dual-trap optical
1112 tweezers instrument equipped with a bright-field camera and a 60X oil immersion objective (Nikon
1113 APO TIRF 60X). Two optical traps are generated by splitting a 10W 1064nm laser by polarity,
1114 with each polarized beam steered by an acousto-optic deflector (IntraAction DTD-274HD6). Each
1115 particle was trapped using approximately 250 mW laser power. 100 μM of SERF2 was mixed with
1116 equimolar TERRA23 or TERRA12 rG4s in 20 mM NaPi, 100 mM KCl, pH 7.4 buffer containing
1117 10% PEG8000 (v/v). The glass slide chamber was prepared as described above and coated with
1118 0.1 mg BSA (5 minutes), followed by rinsing. 15 μL of SERF2-rG4 droplet samples were passed
1119 into a glass slide chamber. Particle fusion was observed by trapping two condensate droplets in
1120 separate optical traps and steering the laser foci together until particle fusion was observed.
1121 Movies of fusion events were recorded within ~5 minutes at room temperature to avoid droplet
1122 settlement on the glass slide and droplet aging. The droplet fusion time was extracted using the
1123 ImageJ program.
1124

1125 **Data availability**

1126 The atomic coordinate generated in this study has been deposited in the Protein Data Bank (PDB)
1127 under the accession code 9DT0 (Human SERF2).
1128

1129 **Acknowledgments**

1130 The authors acknowledge the CERIC-ERIC consortium for access to experimental facilities and
1131 financial support. NMR data reported in this study was supported by the University of Michigan
1132 BioNMR Core Facility. We thank Dr Anthony Vecchiarelli for providing the instrumentation facility
1133 with fluorescence imaging. We thank National Brain Research Centre (NBRC), India for providing
1134 a High-Performance Research Computing facility. We thank the grants from the Howard Hughes
1135 Medical Institute - HHMI to J.C.A.B., from the National Institutes of Health (R35GM142864) to

1136 D.D., JSPS KAKENHI (20H05626) to H.S., and Slovenian Research and Innovation Agency -
1137 ARIS grants P1-0242 and J1-60019 to J.P.
1138

1139 **Author contributions**

1140 B.R.S. and J.C.A.B. conceived the original idea; J.C.A.B. supervised the research; B.R.S., X.D.,
1141 E.W., N.W.C., B.B.Z., and E.M. performed the experiments in this study. B.R.S., X.D., E.J., E.M.,
1142 B.B.G., H.Y., V.S., V.K., J.P., S.E.H, B.D., A.K., J.D.H., H.S., and D.D., performed data analysis.
1143 B.R.S., X.D., and J.C.A.B. wrote the manuscript. The manuscript was written through the
1144 contributions of all authors. All authors have approved the final version of the manuscript.
1145

1146 **References**

1147 1. Buchan, J. R. & Parker, R. Eukaryotic Stress Granules: The Ins and Outs of Translation.
1148 *Mol Cell* **36**, 932–941 (2009).
1149 2. Hofmann, S., Kedersha, N., Anderson, P. & Ivanov, P. Molecular mechanisms of stress
1150 granule assembly and disassembly. *Biochimica et Biophysica Acta (BBA) - Molecular Cell*
1151 *Research* **1868**, 118876 (2021).
1152 3. Turner, M. *et al.* rG4detector, a novel RNA G-quadruplex predictor, uncovers their impact
1153 on stress granule formation. *Nucleic Acids Res* **50**, 11426–11441 (2022).
1154 4. Lane, A. N., Chaires, J. B., Gray, R. D. & Trent, J. O. Stability and kinetics of G-
1155 quadruplex structures. *Nucleic Acids Res* **36**, 5482–5515 (2008).
1156 5. Chen, X. C. *et al.* Tracking the Dynamic Folding and Unfolding of RNA G-Quadruplexes
1157 in Live Cells. *Angewandte Chemie - International Edition* **57**, 4702–4706 (2018).
1158 6. Laguerre, A. *et al.* Visualization of RNA-Quadruplexes in Live Cells. *J Am Chem Soc*
1159 **137**, 8521–8525 (2015).
1160 7. Biffi, G., Di Antonio, M., Tannahill, D. & Balasubramanian, S. Visualization and
1161 selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human
1162 cells. *Nat Chem* **6**, 75–80 (2014).
1163 8. Kharel, P., Becker, G., Tsvetkov, V. & Ivanov, P. Properties and biological impact of
1164 RNA G-quadruplexes: From order to turmoil and back. *Nucleic Acids Res* **48**, 12534–
1165 12555 (2020).
1166 9. Lyu, K., Chow, E. Y.-C., Mou, X., Chan, T.-F. & Kwok, C. K. RNA G-quadruplexes
1167 (rG4s): genomics and biological functions. *Nucleic Acids Res* **49**, 5426–5450 (2021).
1168 10. Dumas, L., Herviou, P., Dassi, E., Cammas, A. & Millevoi, S. G-Quadruplexes in RNA
1169 Biology: Recent Advances and Future Directions. *Trends Biochem Sci* **46**, 270–283
1170 (2021).
1171 11. Simone, R., Fratta, P., Neidle, S., Parkinson, G. N. & Isaacs, A. M. G-quadruplexes:
1172 Emerging roles in neurodegenerative diseases and the non-coding transcriptome. *FEBS*
1173 *Lett* **589**, 1653–1668 (2015).
1174 12. Cammas, A. & Millevoi, S. RNA G-quadruplexes: emerging mechanisms in disease.
1175 *Nucleic Acids Res* **45**, 1584–1595 (2017).
1176 13. Fay, M. M., Anderson, P. J. & Ivanov, P. ALS/FTD-Associated C9ORF72 Repeat RNA
1177 Promotes Phase Transitions In Vitro and in Cells. *Cell Rep* **21**, 3573–3584 (2017).

1178 14. Yang, S. Y. *et al.* Transcriptome-wide identification of transient RNA G-quadruplexes in
1179 human cells. *Nat Commun* **9**, 4730 (2018).
1180 15. Byrd, A. K. *et al.* Evidence that G-quadruplex DNA accumulates in the cytoplasm and
1181 participates in stress granule assembly in response to oxidative stress. *Journal of*
1182 *Biological Chemistry* **291**, 18041–18057 (2016).
1183 16. Yang, X. *et al.* RNA G-quadruplex structure contributes to cold adaptation in plants. *Nat*
1184 *Commun* **13**, 6224 (2022).
1185 17. Kharel, P. *et al.* Stress promotes RNA G-quadruplex folding in human cells. *Nat Commun*
1186 **14**, 205 (2023).
1187 18. Asamitsu, S. *et al.* RNA G-quadruplex organizes stress granule assembly through
1188 DNAPTP6 in neurons. *Sci Adv* **9**, eade2035 (2023).
1189 19. Fay, M. M., Lyons, S. M. & Ivanov, P. RNA G-Quadruplexes in Biology: Principles and
1190 Molecular Mechanisms. *J Mol Biol* **429**, 2127–2147 (2017).
1191 20. Santos, T., Salgado, G. F., Cabrita, E. J. & Cruz, C. Nucleolin: a binding partner of G-
1192 quadruplex structures. *Trends Cell Biol* **32**, 561–564 (2022).
1193 21. Tippana, R., Chen, M. C., Demeshkina, N. A., Ferré-D'Amaré, A. R. & Myong, S. RNA
1194 G-quadruplex is resolved by repetitive and ATP-dependent mechanism of DHX36. *Nat*
1195 *Commun* **10**, 1855 (2019).
1196 22. Sauer, M. *et al.* DHX36 prevents the accumulation of translationally inactive mRNAs
1197 with G4-structures in untranslated regions. *Nat Commun* **10**, 2421 (2019).
1198 23. Ishiguro, A., Lu, J., Ozawa, D., Nagai, Y. & Ishihama, A. ALS-linked FUS mutations
1199 dysregulate G-quadruplex-dependent liquid–liquid phase separation and liquid-to-solid
1200 transition. *Journal of Biological Chemistry* **297**, 101284 (2021).
1201 24. Flynn, R. L. *et al.* TERRA and hnRNPA1 orchestrate an RPA-to-POT1 switch on
1202 telomeric single-stranded DNA. *Nature* **471**, 532–538 (2011).
1203 25. Ishiguro, A., Katayama, A. & Ishihama, A. Different recognition modes of G-quadruplex
1204 RNA between two ALS/FTLD-linked proteins TDP-43 and FUS. *FEBS Lett* **595**, 310–323
1205 (2021).
1206 26. Biffi, G., Tannahill, D. & Balasubramanian, S. An intramolecular G-quadruplex structure
1207 is required for binding of telomeric repeat-containing RNA to the telomeric protein TRF2.
1208 *J Am Chem Soc* **134**, 11974–11976 (2012).
1209 27. Sahoo, B. R. *et al.* Protein G-quadruplex interactions and their effects on phase transitions
1210 and protein aggregation. *Nucleic Acids Res* **52**, 4702–4722 (2024).
1211 28. Mendoza, O., Bourdoncle, A., Boulé, J.-B., Brosh, R. M. & Mergny, J.-L. G-quadruplexes
1212 and helicases. *Nucleic Acids Res* **44**, 1989–2006 (2016).
1213 29. Tosoni, E. *et al.* Nucleolin stabilizes G-quadruplex structures folded by the LTR promoter
1214 and silences HIV-1 viral transcription. *Nucleic Acids Res* **43**, 8884–8897 (2015).
1215 30. Hong, Z. *et al.* Eukaryotic Pif1 helicase unwinds G-quadruplex and dsDNA using a
1216 conserved wedge. *Nat Commun* **15**, 6104 (2024).
1217 31. Chen, M. C. *et al.* Structural basis of G-quadruplex unfolding by the DEAH/RHA helicase
1218 DHX36. *Nature* **558**, 465–483 (2018).
1219 32. Chen, L. *et al.* Structural basis for nucleolin recognition of *MYC* promoter G-quadruplex.
1220 *Science (1979)* **388**, (2025).
1221 33. Asamitsu, S. & Shioda, N. Potential roles of G-quadruplex structures in RNA granules for
1222 physiological and pathological phase separation. *The Journal of Biochemistry* **169**, 527–
1223 533 (2021).

1224 34. Luige, J., Armaos, A., Tartaglia, G. G. & Ørom, U. A. V. Predicting nuclear G-quadruplex
1225 RNA-binding proteins with roles in transcription and phase separation. *Nat Commun* **15**,
1226 2585 (2024).
1227 35. Alderson, T. R. & Kay, L. E. NMR spectroscopy captures the essential role of dynamics
1228 in regulating biomolecular function. *Cell* **184**, 577–595 (2021).
1229 36. Schiavina, M. *et al.* Intrinsically disordered proteins studied by NMR spectroscopy. *J*
1230 *Magn Reson Open* **18**, 100143 (2024).
1231 37. Tsang, B., Pritišanac, I., Scherer, S. W., Moses, A. M. & Forman-Kay, J. D. Phase
1232 Separation as a Missing Mechanism for Interpretation of Disease Mutations. *Cell* **183**,
1233 1742–1756 (2020).
1234 38. Kharel, P. & Ivanov, P. RNA G-quadruplexes and stress: emerging mechanisms and
1235 functions. *Trends Cell Biol* **34**, 771–784 (2024).
1236 39. Sahoo, B. R., Subramanian, V. & Bardwell, J. C. A. Backbone 1H, 13C, and 15N
1237 chemical shift assignments for human SERF2. *Biomol NMR Assign* **18**, 51–57 (2024).
1238 40. van Ham, T. J. *et al.* Identification of MOAG-4/SERF as a regulator of age-related
1239 proteotoxicity. *Cell* **142**, 601–612 (2010).
1240 41. Falsone, S. F. *et al.* SERF Protein Is a Direct Modifier of Amyloid Fiber Assembly. *Cell*
1241 *Rep* **2**, 358–371 (2012).
1242 42. Meinen, B. A., Gadkari, V. v, Stull, F., Ruotolo, B. T. & Bardwell, J. C. A. SERF engages
1243 in a fuzzy complex that accelerates primary nucleation of amyloid proteins. **116**, 23040–
1244 23049 (2019).
1245 43. Sahoo, B. R. & Bardwell, J. C. A. SERF, a family of tiny highly conserved, highly
1246 charged proteins with enigmatic functions. *FEBS J* **290**, 4150–4162 (2023).
1247 44. Palumbo, S. L. *et al.* A novel G-quadruplex-forming GGA repeat region in the c-myb
1248 promoter is a critical regulator of promoter activity. *Nucleic Acids Res* **36**, 1755 (2008).
1249 45. Mitra, R. *et al.* Molecular insights into the interaction between a disordered protein and a
1250 folded RNA. *Proceedings of the National Academy of Sciences* **121**, (2024).
1251 46. Reddy, K., Zamiri, B., Stanley, S. Y. R., Macgregor, R. B. & Pearson, C. E. The Disease-
1252 associated r(GGGGCC) Repeat from the C9orf72 Gene Forms Tract Length-dependent
1253 Uni- and Multimolecular RNA G-quadruplex Structures. *Journal of Biological Chemistry*
1254 **288**, 9860–9866 (2013).
1255 47. Matsuo, K. *et al.* RNA G-quadruplexes form scaffolds that promote neuropathological α-
1256 synuclein aggregation. *Cell* **187**, 6835-6848.e20 (2024).
1257 48. Dominguez, D. *et al.* Sequence, Structure, and Context Preferences of Human RNA
1258 Binding Proteins. *Mol Cell* **70**, 854–867 (2018).
1259 49. Komatsu, K. R. *et al.* RNA structure-wide discovery of functional interactions with
1260 multiplexed RNA motif library. *Nat Commun* **11**, 6275 (2020).
1261 50. Kikin, O., D'Antonio, L. & Bagga, P. S. QGRS Mapper: a web-based server for predicting
1262 G-quadruplexes in nucleotide sequences. *Nucleic Acids Res* **34**, W676–W682 (2006).
1263 51. Bettin, N., Oss Pegorar, C. & Cusanelli, E. The Emerging Roles of TERRA in Telomere
1264 Maintenance and Genome Stability. *Cells* **8**, 246 (2019).
1265 52. Zhang, K. *et al.* The C9orf72 repeat expansion disrupts nucleocytoplasmic transport.
1266 *Nature* **525**, 56–61 (2015).
1267 53. Larini, L. *et al.* Telomeric lncRNA TERRA localizes to stress granules in human ALT
1268 cells. *bioRxiv* 2024.06.18.599513 (2024) doi:10.1101/2024.06.18.599513.

1269 54. Rossi, S. *et al.* Nuclear accumulation of mRNAs underlies G4C2 repeat-induced
1270 translational repression in a cellular model of *C9orf72* ALS. *J Cell Sci* **128**, 1787–1799
1271 (2015).
1272 55. Huang, Z. L. *et al.* Identification of G-Quadruplex-Binding Protein from the Exploration
1273 of RGG Motif/G-Quadruplex Interactions. *J Am Chem Soc* **140**, 17945–17955 (2018).
1274 56. Yagi, R., Miyazaki, T. & Oyoshi, T. G-quadruplex binding ability of TLS/FUS depends
1275 on the β-spiral structure of the RGG domain. *Nucleic Acids Res* **46**, 5894–5901 (2018).
1276 57. Zhang, Y., Gaetano, C. M., Williams, K. R., Bassell, G. J. & Mihailescu, M. R. FMRP
1277 interacts with G-quadruplex structures in the 3 ′-UTR of its dendritic target shank1
1278 mRNA. *RNA Biol* **11**, 1364–1374 (2014).
1279 58. Awadasseid, A., Ma, X., Wu, Y. & Zhang, W. G-quadruplex stabilization via small-
1280 molecules as a potential anti-cancer strategy. *Biomedicine & Pharmacotherapy* **139**,
1281 111550 (2021).
1282 59. Ma, W., Zhen, G., Xie, W. & Mayr, C. In vivo reconstitution finds multivalent RNA–
1283 RNA interactions as drivers of mesh-like condensates. *Elife* **10**, (2021).
1284 60. Guillén-Boixet, J. *et al.* RNA-Induced Conformational Switching and Clustering of G3BP
1285 Drive Stress Granule Assembly by Condensation. *Cell* **181**, 346-361.e17 (2020).
1286 61. Wang, M. *et al.* PaxDb, a database of protein abundance averages across all three domains
1287 of life. *Molecular and Cellular Proteomics* **11**, 492–500 (2012).
1288 62. Kallweit, L. *et al.* Chronic RNA G-quadruplex accumulation in aging and Alzheimer's
1289 disease. *Elife* **14**, (2025).
1290 63. Posey, A. E., Holehouse, A. S. & Pappu, R. V. Phase Separation of Intrinsically
1291 Disordered Proteins. in *Methods in Enzymology* vol. 611 1–30 (2018).
1292 64. Dar, F. *et al.* Biomolecular condensates form spatially inhomogeneous network fluids. *Nat*
1293 *Commun* **15**, 3413 (2024).
1294 65. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates:
1295 organizers of cellular biochemistry. *Nat Rev Mol Cell Biol* **18**, 285–298 (2017).
1296 66. Dyson, H. J. & Wright, P. E. Intrinsically unstructured proteins and their functions. *Nat*
1297 *Rev Mol Cell Biol* **6**, 197–208 (2005).
1298 67. Holehouse, A. S. & Kragelund, B. B. The molecular basis for cellular function of
1299 intrinsically disordered protein regions. *Nat Rev Mol Cell Biol* **25**, 187–211 (2024).
1300 68. Camacho-Zarco, A. R. *et al.* NMR Provides Unique Insight into the Functional Dynamics
1301 and Interactions of Intrinsically Disordered Proteins. *Chem Rev* **122**, 9331–9356 (2022).
1302 69. Martadinata, H. & Phan, A. T. Structure of propeller-type parallel-stranded RNA G-
1303 quadruplexes, formed by human telomeric RNA sequences in K+ solution. *J Am Chem*
1304 *Soc* **131**, 2570–2578 (2009).
1305 70. Collie, G. W., Haider, S. M., Neidle, S. & Parkinson, G. N. A crystallographic and
1306 modelling study of a human telomeric RNA (TERRA) quadruplex. *Nucleic Acids Res* **38**,
1307 5569–5580 (2010).
1308 71. Collie, G. W., Sparapani, S., Parkinson, G. N. & Neidle, S. Structural Basis of Telomeric
1309 RNA Quadruplex−Acridine Ligand Recognition. *J Am Chem Soc* **133**, 2721–2728 (2011).
1310 72. Yamaguchi, T., Matsuzaki, K. & Hoshino, M. Transient formation of intermediate
1311 conformational states of amyloid-β peptide revealed by heteronuclear magnetic resonance
1312 spectroscopy. *FEBS Lett* **585**, 1097–1102 (2011).

1313 73. Whitehead, R. D., Teschke, C. M. & Alexandrescu, A. T. NMR Mapping of Disordered
1314 Segments from a Viral Scaffolding Protein Enclosed in a 23 MDa Procapsid. *Biophys J*
1315 **117**, 1387–1392 (2019).
1316 74. Shen, Y. & Bax, A. Protein structural information derived from nmr chemical shift with
1317 the neural network program talos-n. *Methods in Molecular Biology* **1260**, (2015).
1318 75. Zeke, A. *et al.* Deep structural insights into RNA-binding disordered protein regions.
1319 *WIREs RNA* **13**, e1714 (2022).
1320 76. Martadinata, H. & Phan, A. T. Structure of human telomeric RNA (TERRA): Stacking of
1321 two G-quadruplex blocks in K+ solution. *Biochemistry* **52**, 2176–2183 (2013).
1322 77. Collie, G. W., Haider, S. M., Neidle, S. & Parkinson, G. N. A crystallographic and
1323 modelling study of a human telomeric RNA (TERRA) quadruplex. *Nucleic Acids Res* **38**,
1324 (2010).
1325 78. Biffi, G., Tannahill, D. & Balasubramanian, S. An Intramolecular G-Quadruplex Structure
1326 Is Required for Binding of Telomeric Repeat-Containing RNA to the Telomeric Protein
1327 TRF2. *J Am Chem Soc* **134**, 11974–11976 (2012).
1328 79. Montero, J. J. *et al.* TERRA recruitment of polycomb to telomeres is essential for histone
1329 trymethylation marks at telomeric heterochromatin. *Nat Commun* **9**, 1548 (2018).
1330 80. Castelli, M., Doria, F., Freccero, M., Colombo, G. & Moroni, E. Studying the Dynamics
1331 of a Complex G-Quadruplex System: Insights into the Comparison of MD and NMR Data.
1332 *J Chem Theory Comput* **18**, 4515–4528 (2022).
1333 81. De Vries, S. J., Van Dijk, M. & Bonvin, A. M. J. J. The HADDOCK web server for data-
1334 driven biomolecular docking. *Nat Protoc* **5**, 883–897 (2010).
1335 82. Reshetnikov, R. V. *et al.* Cation binding to 15-TBA quadruplex DNA is a multiple-
1336 pathway cation-dependent process. *Nucleic Acids Res* **39**, 9789–9802 (2011).
1337 83. Kim, H. & Pak, Y. Improving All-Atom Force Field to Accurately Describe DNA G-
1338 Quadruplex Loops. *J Phys Chem B* **126**, 6199–6209 (2022).
1339 84. Pras, A. *et al.* The cellular modifier MOAG-4/SERF drives amyloid formation through
1340 charge complementation. *EMBO J* **40**, e107568 (2021).
1341 85. Wang, B. *et al.* Liquid–liquid phase separation in human health and diseases. *Signal*
1342 *Transduct Target Ther* **6**, 290 (2021).
1343 86. André, A. A. M. & Spruijt, E. Liquid–Liquid Phase Separation in Crowded Environments.
1344 *Int J Mol Sci* **21**, 5908 (2020).
1345 87. Erkamp, N. A. *et al.* Differential interactions determine anisotropies at interfaces of RNA-
1346 based biomolecular condensates. *Nat Commun* **16**, 3463 (2025).
1347 88. Ginell, G. M. & Holehouse, A. S. An Introduction to the Stickers-and-Spacers Framework
1348 as Applied to Biomolecular Condensates. in *Methods in Molecular Biology* vol. 2563 95–
1349 116 (2023).
1350 89. Gao, C. *et al.* Effects of Molecular Crowding on the Structure, Stability, and Interaction
1351 with Ligands of G-quadruplexes. *ACS Omega* **8**, 14342–14348 (2023).
1352 90. Galvanetto, N. *et al.* Extreme dynamics in a biomolecular condensate. *Nature* **619**, 876–
1353 883 (2023).
1354 91. Rekhi, S. *et al.* Expanding the molecular language of protein liquid–liquid phase
1355 separation. *Nat Chem* **16**, 1113–1124 (2024).
1356 92. Kamagata, K., Kusano, R., Kanbayashi, S., Banerjee, T. & Takahashi, H. Single-molecule
1357 characterization of target search dynamics of DNA-binding proteins in DNA-condensed
1358 droplets. *Nucleic Acids Res* **51**, 6654–6667 (2023).

1359 93. Shen, Z. *et al.* Biological condensates form percolated networks with molecular motion
1360 properties distinctly different from dilute solutions. *Elife* **12**, e81907 (2023).
1361 94. Matsuki, H. *et al.* Both G3BP1 and G3BP2 contribute to stress granule formation. *Genes*
1362 *to Cells* **18**, 135–146 (2013).
1363 95. Yang, P. *et al.* G3BP1 Is a Tunable Switch that Triggers Phase Separation to Assemble
1364 Stress Granules. *Cell* **181**, 325–345 (2020).
1365 96. Babinchak, W. M. & Surewicz, W. K. Liquid–Liquid Phase Separation and Its
1366 Mechanistic Role in Pathological Protein Aggregation. *J Mol Biol* **432**, 1910–1925 (2020).
1367 97. Protter, D. S. W. & Parker, R. Principles and Properties of Stress Granules. *Trends Cell*
1368 *Biol* **26**, 668–679 (2016).
1369 98. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates:
1370 organizers of cellular biochemistry. *Nat Rev Mol Cell Biol* **18**, 285–298 (2017).
1371 99. Mitrea, D. M., Mittasch, M., Gomes, B. F., Klein, I. A. & Murcko, M. A. Modulating
1372 biomolecular condensates: a novel approach to drug discovery. *Nat Rev Drug Discov* **21**,
1373 841–862 (2022).
1374 100. Ivanov, P., Kedersha, N. & Anderson, P. Stress granules and processing bodies in
1375 translational control. *Cold Spring Harb Perspect Biol* **11**, a032813 (2019).
1376 101. Dignon, G. L., Best, R. B. & Mittal, J. Biomolecular Phase Separation: From Molecular
1377 Driving Forces to Macroscopic Properties. *Annu Rev Phys Chem* **71**, 53–75 (2020).
1378 102. Roden, C. & Gladfelter, A. S. RNA contributions to the form and function of
1379 biomolecular condensates. *Nat Rev Mol Cell Biol* **22**, 183–195 (2021).
1380 103. Alberti, S. & Hyman, A. A. Biomolecular condensates at the nexus of cellular stress,
1381 protein aggregation disease and ageing. *Nat Rev Mol Cell Biol* **22**, 196–213 (2021).
1382 104. Lafontaine, D. L. J., Riback, J. A., Bascetin, R. & Brangwynne, C. P. The nucleolus as a
1383 multiphase liquid condensate. *Nat Rev Mol Cell Biol* **22**, 165–182 (2021).
1384 105. Song, J., Perreault, J.-P., Topisirovic, I. & Richard, S. RNA G-quadruplexes and their
1385 potential regulatory roles in translation. *Translation* **4**, e1244031 (2016).
1386 106. Shao, X. *et al.* RNA G-quadruplex structures mediate gene regulation in bacteria. *mBio*
1387 **11**, e02926-19 (2020).
1388 107. Zhang, R. *et al.* Detecting and profiling endogenous rna g-quadruplexes in the human
1389 transcriptome. *Int J Mol Sci* **22**, (2021).
1390 108. Bugaut, A. & Balasubramanian, S. 5'-UTR RNA G-quadruplexes: translation regulation
1391 and targeting. *Nucleic Acids Res* **40**, 4727–4741 (2012).
1392 109. Stroo, E. *et al.* Deletion of SERF2 in mice delays embryonic development and alters
1393 amyloid deposit structure in the brain. *Life Sci Alliance* **6**, e202201730 (2023).
1394 110. Bessi, I., Wirmer-Bartoschek, J., Dash, J. & Schwalbe, H. Targeting G-quadruplex with
1395 Small Molecules: An NMR View. in *Modern Magnetic Resonance* 1–22 (Springer
1396 International Publishing, Cham, 2017). doi:10.1007/978-3-319-28275-6_126-1.
1397 111. Ahuja, S. *et al.* A Model of the Membrane-bound Cytochrome b5-Cytochrome P450
1398 Complex from NMR and Mutagenesis Data. *Journal of Biological Chemistry* **288**, 22080–
1399 22095 (2013).
1400 112. van Ingen, H. & Bonvin, A. M. J. J. Information-driven modeling of large macromolecular
1401 assemblies using NMR data. *Journal of Magnetic Resonance* **241**, 103–114 (2014).
1402 113. Zeng, D., Abzhanova, A., Brown, B. P. & Reiter, N. J. Dissecting Monomer-Dimer
1403 Equilibrium of an RNase P Protein Provides Insight Into the Synergistic Flexibility of 5'
1404 Leader Pre-tRNA Recognition. *Front Mol Biosci* **8**, (2021).

1405 114. Nshogoza, G. *et al.* NMR Fragment-Based Screening against Tandem RNA Recognition
1406 Motifs of TDP-43. *Int J Mol Sci* **20**, 3230 (2019).
1407 115. Traczyk, A., Liew, C. W., Gill, D. J. & Rhodes, D. Structural basis of G-quadruplex DNA
1408 recognition by the yeast telomeric protein Rap1. *Nucleic Acids Res* **48**, 4562–4571 (2020).
1409 116. Kilburn, D., Roh, J. H., Guo, L., Briber, R. M. & Woodson, S. A. Molecular Crowding
1410 Stabilizes Folded RNA Structure by the Excluded Volume Effect. *J Am Chem Soc* **132**,
1411 8690–8696 (2010).
1412 117. Miyoshi, D., Karimata, H. & Sugimoto, N. Hydration Regulates Thermodynamics of G-
1413 Quadruplex Formation under Molecular Crowding Conditions. *J Am Chem Soc* **128**,
1414 7957–7963 (2006).
1415 118. Brčić, J. & Plavec, J. ALS and FTD linked GGGGCC-repeat containing DNA
1416 oligonucleotide folds into two distinct G-quadruplexes. *Biochim Biophys Acta Gen Subj*
1417 **1861**, 1237–1245 (2017).
1418 119. Kocman, V. & Plavec, J. Tetrahelical structural family adopted by AGCGA-rich
1419 regulatory DNA regions. *Nat Commun* **8**, 15355 (2017).
1420 120. Goering, R. *et al.* FMRP promotes RNA localization to neuronal projections through
1421 interactions between its RGG domain and g-quadruplex RNA sequences. *Elife* **9**, e52621
1422 (2020).
1423 121. Lee, W., Tonelli, M. & Markley, J. L. NMRFAM-SPARKY: Enhanced software for
1424 biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327 (2015).
1425 122. Huang, J. *et al.* CHARMM36m: An improved force field for folded and intrinsically
1426 disordered proteins. *Nat Methods* **14**, 71–73 (2016).
1427 123. Abraham, M. J. *et al.* Gromacs: High performance molecular simulations through multi-
1428 level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
1429 124. Hennig, J. *et al.* Structural Analysis of Protein–RNA Complexes in Solution Using NMR
1430 Paramagnetic Relaxation Enhancements. *Methods Enzymol* **558**, 333–362 (2015).
1431 125. Kern, N. R., Lee, J., Choi, Y. K. & Im, W. CHARMM-GUI multicomponent assembler
1432 for modeling and simulation of complex multicomponent systems. *Biophys J* **121**, 529a
1433 (2022).
1434 126. Michaud-Agrawal, N., Denning, E. J., Woolf, T. B. & Beckstein, O. MDAnalysis: A
1435 toolkit for the analysis of molecular dynamics simulations. *J Comput Chem* **32**, 2319–
1436 2327 (2011).
1437

1438

1439 **Figure Legends**

1440 **Fig. 1| High-throughput screening for SERF2 binding specificity. a** Fluorescence polarization assay
1441 used to measure the binding affinity of SERF2 to 6-FAM labeled poly ribo-polynucleotides and rG4
1442 quadruplex forming sequences, as indicated. Data are presented as mean values ± standard deviations
1443 obtained from three independent replicates. **b** A schematic representation of RNA Bind-n-Seq
1444 experiments using RNA pools. A randomized DNA oligonucleotide pool was transcribed to RNA and
1445 folded in a buffer containing either KCl which favors rG4 formation or LiCl which disfavors rG4 folding. An
1446 additional RNA pool was made by replacing guanines (G) with 7-deaza (7dG) to eliminate rG4 quadruplex
1447 folding. These pools were mixed with GST-SBP-SERF2, and bound RNA was isolated and sequenced
1448 with ~2 X 10^6 reads. **c** RNA Bind-n-Seq analysis of the 6-mers enrichment in KCl versus RNA that was
1449 made with 7dG in a sample mixture containing 50 nM SERF2. The guanine-rich 6mers in the top 5 kmers
1450 are labeled. **d** Enrichment of rG4 quadruplex patterns in different conditions were classified by G4
1451 quadruplex strength in buffer containing 50 nM SERF2. Sequences containing 3 or more guanines in the
1452 G-tetrad are referred to as strong rG4s, while sequences with ≥8 guanines but lacking a defined G4
1453 forming motif as defined by QGRS mapper program are referred to as non rG4s. Data are presented as
1454 mean values ± standard deviations for each condition obtained from two different RNA library
1455 preparations. **e** FOREST analysis: average binding intensities of SERF2 to a library containing 1800
1456 folded human pre-miRNAs (orange curve) and to 10 defined rG4 structures (grey curve). The p-value
1457 shown was determined by the two-tailed Brunner-Munzel test. **f** The fluorescence polarization plot that
1458 shows binding affinity measurements between SERF2 and three different 6-FAM labeled rG4s. The
1459 binding assays in **a** and **f** were done with varied protein concentrations mixed with 20 nM rG4 quadruplex
1460 or polynucleotides at room temperature in 20 mM NaPi (pH 7.4), 100 mM KCl. The standard deviations
1461 shown were calculated from three independent replicates, and data are presented as mean values.

1462

1463 **Fig. 2| RNA interaction drives liquid-liquid phase transition in SERF2. a** Phase regime illustrating
1464 phase transitions for SERF2 as a function of total RNA concentration using RNA extracted from HeLa
1465 cells (left panel). Fluorescence imaging shows gel-like structures present in solutions containing 50 μM
1466 SERF2 mixed with 200 ng of total HeLa cell RNA containing 10% (w/v) PEG8000, incubated for 30
1467 minutes at room temperature (right panel). **b** Dynamics and recovery of Cy5-labeled SERF2 in total HeLa
1468 cell RNA droplets, obtained by FRAP analysis, suggest that the mesh-like condensates observed in **a**,
1469 right panel are dynamic and reversible. Standard deviations were calculated by analyzing 6 isolated
1470 droplets subjected to FRAP, and data are presented as mean values. **c** Fluorescence images show 50
1471 μM SERF2, dissolved in 20 mM NaPi (pH 7.4),100 mM KCl, readily undergoes phase transition (right)
1472 when mixed with equimolar concentrations of three different rG4s, TERRA23, (G4C2)4, and (UG4U)6.
1473 Similar results were obtained in three independent repeated experiments. The sample mixture contains
1474 1/200th Cy-5 labeled SERF2 (purple) and 6-FAM rG4 (green), as indicated in the figure inset. Phase
1475 regimes illustrating phase transitions for SERF2 and TERRA23 rG4 in 10% PEG8000, at varying protein
1476 and RNA concentrations, are shown on the right panel. **d** Phase diagram showing SERF2 and TERRA23
1477 rG4 sample mixtures undergoing phase separation, in varying salt and PEG8000 concentrations. **e** Two-
1478 component FRAP analysis was done to measure the recovery rates of SERF2 (purple) and TERRA23
1479 rG4 (green) in SERF2-rG4 droplets. Data are presented as mean values, and standard errors were
1480 calculated by analyzing 8 isolated droplets subjected to FRAP. The pre-bleached, after-bleached (0 s),
1481 and recovered droplets (300 s) are shown above the FRAP plot. The FRAP data were fitted in GraphPad
1482 Prism, using a non-linear regression, one-phase association model, to obtain the recovery halftime (t1/2)
1483 reported in the text.
1484
1485

1486 **Fig. 3 | High-resolution NMR structure of SERF2 reveals that its disordered and dynamic N-**
1487 **terminal domain binds TERRA rG4**. **a** 20 best NMR ensemble model structures of SERF2. The average
1488 converged helical structure spanning residues 33-47 in SERF2 is shown as the helix in orange. **b** 1H-15N
1489 NMR assignment (red spectrum) of 100 μM human SERF2 mixed with 20 μM (yellow) and 50 μM (purple)
1490 TERRA23 rG4 at 4 °C. Spectral zooms on the right illustrate chemical shift changes in several N- and C-

terminal residues at 2:1 protein: rG4 ratio. **c** The 1H-15N chemical shift perturbations (left y-axis) were calculated from (**b**) and plotted as color-shaded peaks for each assigned residue in SERF2 at increasing TERRA23 rG4 concentrations. The yellow and purple colors in the graph correlate to their corresponding spectrum shown in (**b**), unassigned peaks are denoted with an asterisk. The heteroNOE values of SERF2 in the absence of TERRA23 rG4 were plotted to highlight the dynamic regions in SERF2. **d** 15N relaxation rates R2/R1 demonstrating that a significant change in dynamics occurs for 200 μM SERF2 on its own (pink dots) and in the presence of 100 μM interacting TERRA23 rG4 (grey squares), as a function of residue number. Error bars represent the standard error from per-residue exponential fitting of peak intensities in NMRFAM-Sparky. **e-f** 2D analysis plots derived from analytical ultracentrifugation experiments for 4.7 μM TERRA12 rG4 without (**e**) or mixed with a 2-molar excess SERF2 (**f**). The partial concentration shown in color on the right y-axes represents the abundance of individual species in the sample solution. **g-h** Cartoon shows the top (**g)** and side-view (**h)** of SERF2 and TERRA rG4 complex. Note the quadrupole-like interactions shown as ellipses in (**g)** and vertical slabs in **(h).** Residues generating the quadrupole-like interactions and distorting TERRA rG4 structure (PDB ID: 2M18) are labeled, and hydrogen bonds are indicated with dashed lines. **(i)** EMSA gel-shift assay of 5 μM TERRA23 rG4 mixed with an equimolar amount of wild-type SERF2 and different lysine to alanine SERF2 mutants. Similar results were obtained in two independent repeated experiments.

Fig. 4| All-atom MD simulation approach to study the structure of SERF2-TERRA rG4 phase separated condensate. a A cubic all-atom MD simulation box encapsulating randomly distributed 30 molecules of SERF2 (orange), 30 molecules of the TERRA10 rG4 (PDB ID: 2M18, blue), molecules of PEG are shown in pink, Cl^- in green, and K^+ in grey. **b-c** Surface representation of the structure of SERF2-TERRA rG4 ring-shaped droplet-like structure found in the condensed phase (**b**), and lower-ordered oligomers (**c**), that were obtained at time 0.5 μs in the MD simulation. The enlarged all-atom cartoon structures of the condensed droplet-phase and the lower-ordered 1:2 SERF2:TERRA rG4 dilute-phase oligomers are shown on the top. The three distinct contact sites in SERF2 in the lower-ordered complex structure (**c,** top) is shown and the TERRA rG4 interacting residues are labeled. The high-resolution images in (**b**, top) show a representative interaction network that involves the three critical binding sites in SERF2 located in the disordered N-terminus. The SERF2 interacting residues are represented with an uppercase letter and TERRA rG4 nucleotide in a smaller-case letter (e.g. R11-∪14 denotes Arg3 and uracil 14 in SERF2 and TERRA rG4, respectively). **(d)** Representative all-atom MD snapshot at 0.5 μs reveals the emergence of a droplet-like condensate featuring a ring-shaped assembly of SERF2 (orange) and TERRA rG4 (blue) complexes, stabilized in the presence of PEG crowding agents (magenta). **(e)** Quantitative analysis of PEG localization reveals time-dependent enrichment of PEG molecules within 6 Å of either SERF2 or TERRA rG4 over the course of the 0.5 μs simulation. **(f)** Time-resolved quantification of Hoogsteen hydrogen bonds within the TERRA rG4 structure confirms high structural integrity of the quadruplex core during the entire 0.5 μs trajectory. Color bar on right indicates the number of protein molecules simultaneously in contact with the rG4, demonstrating that even under high protein association, the G4 core remains structurally preserved. **(g)** Mean squared displacement or MSD analysis of individual components reveals diffusion of SERF2 in the condensates. The 3D diffusion constant (D) extracted from the linear regime of MSD (t) using equation $D = \frac{\mathrm{MSD(t)}}{6\mathrm{t}}$ is calculated to be 1.07 $\mu m^2/s$.

Fig. 5| Single-molecule imaging and tracking of SERF2 interactions with TERRA rG4s. **a** Single-molecule fluorescence microscopy shows a single droplet of 50 μM SERF2 (spiked with picomolar Cy5 labeled SERF2) mixed with equimolar TERRA12 rG4s (left panel). Scale bar is 3.2 μm. Similar results were obtained in three independent repeated experiments. The corresponding single-particle tracking image of SERF2 molecules within the single droplet generated using ImageJ is shown on the right. **b** MSD plot of SERF2 (average MSD ≈ 0.95±0.23 $\mu m^2/s$) obtained from a single particle tracking fluorescence microscopy experiment. The shaded area represents the standard deviation across tracked particles. Data are presented as mean values ± standard error. **c** Distribution of diffusion coefficients (D) extracted from tracking trajectories of individual SERF2-Cy5 molecules inside droplets shows a skewed and heterogeneous mobility profile centered around D ≈ 0.4-0.6 $\mu m^2/s$, consistent with prior reports of protein diffusion in crowded, phase-separated compartments[92,93]. **d** FRET distance histogram between

1545 donor and acceptor was determined based on the FRET efficiency in SERF2 (T2C and A51C) and
1546 TERRA12 rG4 phase-separated droplets containing picomolar Cy3-Cy5 labeled SERF2 in dilute-phase
1547 (left) and condensed-phase (right) samples separated by centrifugation (see methods).
1548
1549
1550 **Fig. 6| SERF2 colocalizes with stress granules upon various stress conditions**. **a**
1551 Immunofluorescence images show that endogenous SERF2 is predominantly distributed in the nucleolus
1552 of fixed U2OS cells, as evidenced by staining with the nucleolar marker fibrillarin. **b** Colocalization
1553 analysis of fibrillin and SERF2 obtained from (**a**). Thirteen foci from nine cells in three biological replicates
1554 were subjected to Pearson's correlation coefficient analysis. **c** SERF2 forms cytoplasmic foci and co-
1555 localizes with the core stress granule marker protein G3BP1 in different stress conditions. **d** This plot
1556 shows the quantification of stress granules retrieved from (**c**) under various stress conditions containing
1557 both SERF2 and G3BP1. At least twenty-six foci from three biological replicates were subjected to
1558 Pearson's correlation coefficient analysis. **e** Fixed U2OS immunofluorescence cell images showing the
1559 DAPI stained nucleus (blue) and oxidative (sodium arsenite) stress-induced granules containing
1560 G4s(red), FUS (purple) and SERF2 (green) as detected by the BG4 antibody and the FUS and SERF2
1561 proteins. **f** Plot showing SERF2 colocalization with FUS and G4sretrieved from (**e**) as measured by the
1562 Pearson's coefficient**.** Twenty foci from three biological replicates were analyzed. The scale bars in (**a, c,**
1563 and **e**) indicate 10 µm.
1564
1565
1566 **Fig. 7| SERF2 regulates stress granule formation and dynamics**. **a** Immunofluorescence of SERF2
1567 and G3BP1 in fixed U2OS cells treated with 0.5 mM sodium arsenite for 1 hour. **b** Plot shows the
1568 percentage of stress granule positive cells after sodium arsenite treatment calculated from images shown
1569 in (**a**). Error bars were calculated from four independent experiments, and data are presented as mean
1570 values. **c** Live-cell imaging of EGFP-FUS HeLa Kyoto cells treated either with a control RNA (siCTRL) or
1571 an RNA targeting SERF (siSERF2), treated with different stressors (0.5 mM Sodium arsenite, 0.4 M
1572 Sorbitol, or 10 µM MG132) for 1 hour. Scale bars in (**a** and **c**) are 10 µm. **d** Plot shows percentage of
1573 stress granule positive cells under different stress treatments calculated from images shown in (**c**). At
1574 least 50 cells from three biological replicates were analyzed, and data are presented as mean values. ****
1575 shown in (**b** and **d**) indicates $P < 0.0001$. **e-f** These graphs show FRAP recovery curves in EGFP-FUS
1576 HeLa Kyoto cells with siCTRL or siSERF2 conditions treated with 0.4 M sorbitol and (**e**) 0.5 mM sodium
1577 arsenite (**f**). Standard deviations were calculated by analyzing four different foci in three replicates
1578 subjected to FRAP, and data are presented as mean values.
1579
1580
1581 **Fig. 8| SERF2 facilitates the phase transition of the stress granule core protein G3BP1 and**
1582 **enhances G3BP1 fluidity. a** DIC and fluorescence images showing co-phase separation of SERF2
1583 (purple) and G3BP1 (unlabeled) with 12.5 ng/µL HeLa total RNA. Similar results were obtained in three
1584 independent repeated experiments. **b** DIC and fluorescence images show SERF2 facilitates G3BP1-RNA
1585 condensation in samples containing the indicated SERF2 and G3BP1 concentrations and 12.5 ng/µL
1586 HeLa total RNA. Similar results were obtained in three independent repeated experiments. Scale bars in
1587 (**a** and **b)** are 10 µm. **c** Phase diagram showing G3BP1 phase transition in a non-crowding condition and
1588 a crowding condition that contained 2.5% PEG8000 in a 20 mM NaPi, pH 7.4, 100 mM KCl buffer. G3BP1
1589 phase transition was measured in the absence or presence of SERF2 or rG4 or mixtures of SERF2 and
1590 rG4s as indicated. **d** Time-series images of 50 µM G3BP1 condensates after photobleaching in the
1591 absence or presence of SERF2, total RNA, and TERRA23 rG4s prepared in 20 mM NaPi, 100 mM KCl
1592 (pH 7.4) containing 5% PEG8000. **e** FRAP recovery plots of G3BP1-Cy5 (left) and SERF2-AF488 (right)
1593 obtained from (**d**), the graph colors correspond to the plus and minus signs of the sample mixture shown
1594 in (**d**). Recovery halftime is calculated by non-linear regression fit and one-phase association using
1595 GraphPad prism. Scale bar is 5 µm. Standard deviations were calculated by analyzing 8 isolated droplets
1596 subjected to FRAP, and data are presented as mean values.
1597

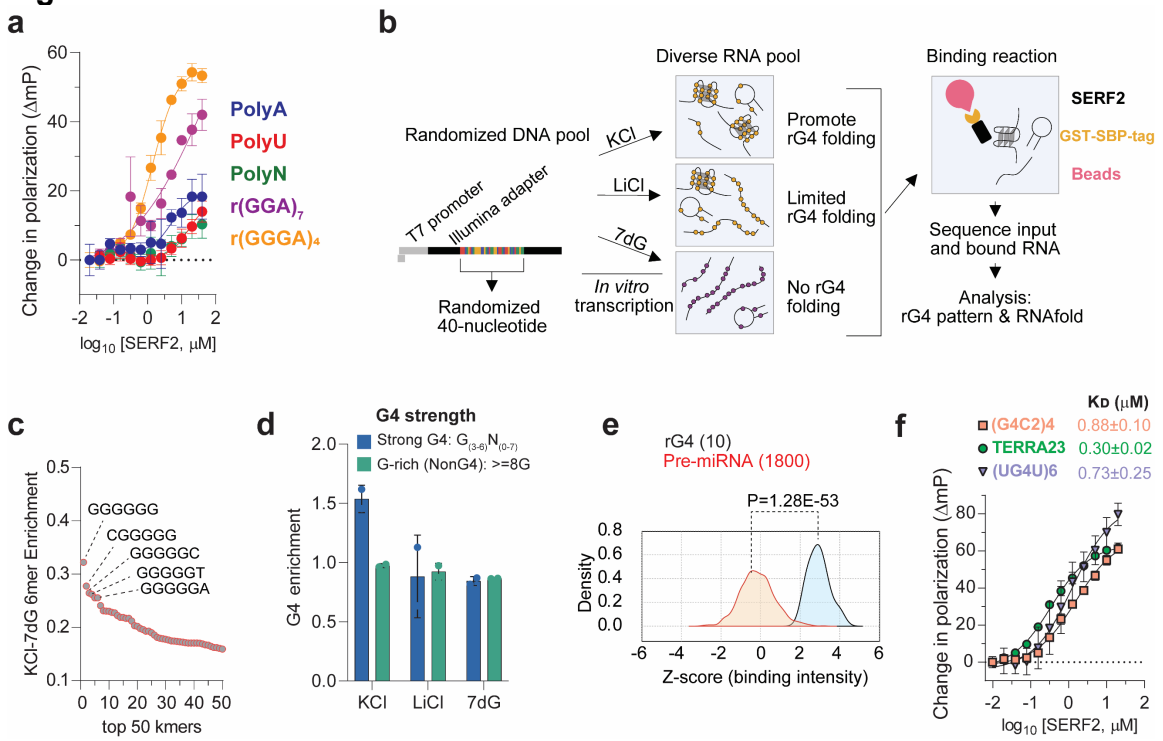
1598 **Table 1. NMR structural statistics of human SERF2**
1599

Parameters	Apo-form
Restraints	
NOE (Total)	494
Intra-residue (|i-j| = 0)	205
Short (|i-j| < 1)	220
Medium range (2 < |i-j| < 5)	69
Long range (|i-j| > 5)	-
Hydrogen bond	24
Dihedral angle	40
Total restraints	558
Restraints per residue (total/long range)	9.5
Residual distance restraint violations	
Target function	0.45 ± 0.17
Average distance violations/structure (Å)	0.02 ± 0.01
Average angle violation/structure (°)	0.25 ± 0.08
Maximum distance violation (Å)	0.67 ± 0.12
Model Quality (ordered region 32-48)	
RMSD from average coordinates (Å)	
All Backbone atoms (ordered/all)	0.24 ± 0.09
All Heavy atoms (ordered/all)	1.15 ± 0.20
Molprobity Ramachandran plot	
Most favored regions (%)	75.2
Additionally allowed regions (%)	24.3
Generously allowed regions (%)	0.5
Disallowed regions (%)	0.0

1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613

1614
1615

Fig. 1



1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636

Fig. 2

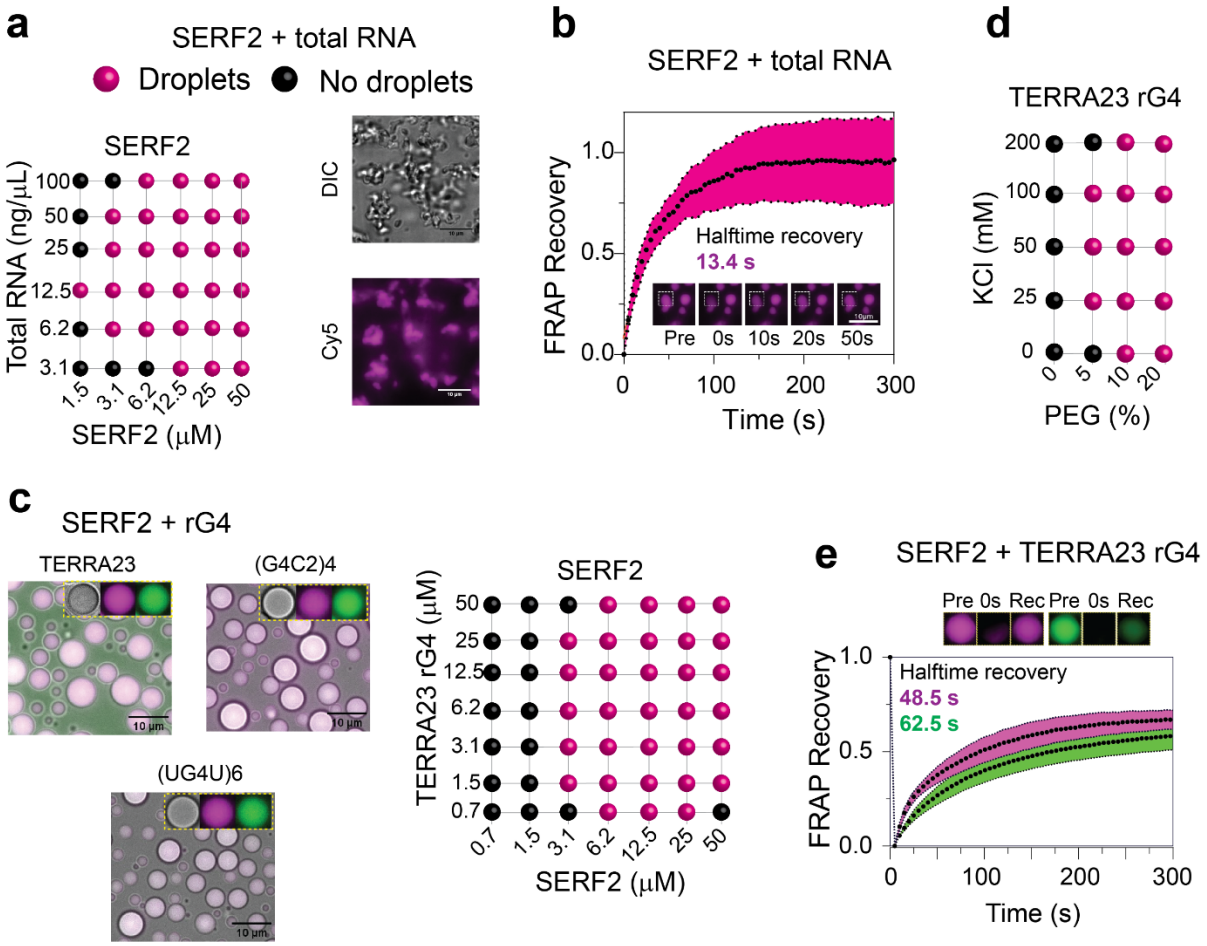

a
SERF2 + total RNA
Droplets
No droplets
SERF2
Total RNA (ng/μL)
100
50
25
12.5
6.2
3.1
1.5
3.1
6.2
12.5
25
50
SERF2 (μM)
DIC
Cy5
b
SERF2 + total RNA
FRAP Recovery
1.0
0.5
0.0
Halftime recovery
13.4 s
Pre
0s
10s
20s
50s
0
100
200
300
Time (s)
d
TERRA23 rG4
KCl (mM)
200
100
50
25
0
0
5
10
20
PEG (%)
c
SERF2 + rG4
TERRA23
(G4C2)4
(UG4U)6
SERF2
TERRA23 rG4 (μM)
50
25
12.5
6.2
3.1
1.5
0.7
0.7
1.5
3.1
6.2
12.5
25
50
SERF2 (μM)
e
SERF2 + TERRA23 rG4
Pre 0s Rec Pre 0s Rec
FRAP Recovery
1.0
0.5
0.0
Halftime recovery
48.5 s
62.5 s
0
100
200
300
Time (s)

Fig. 3

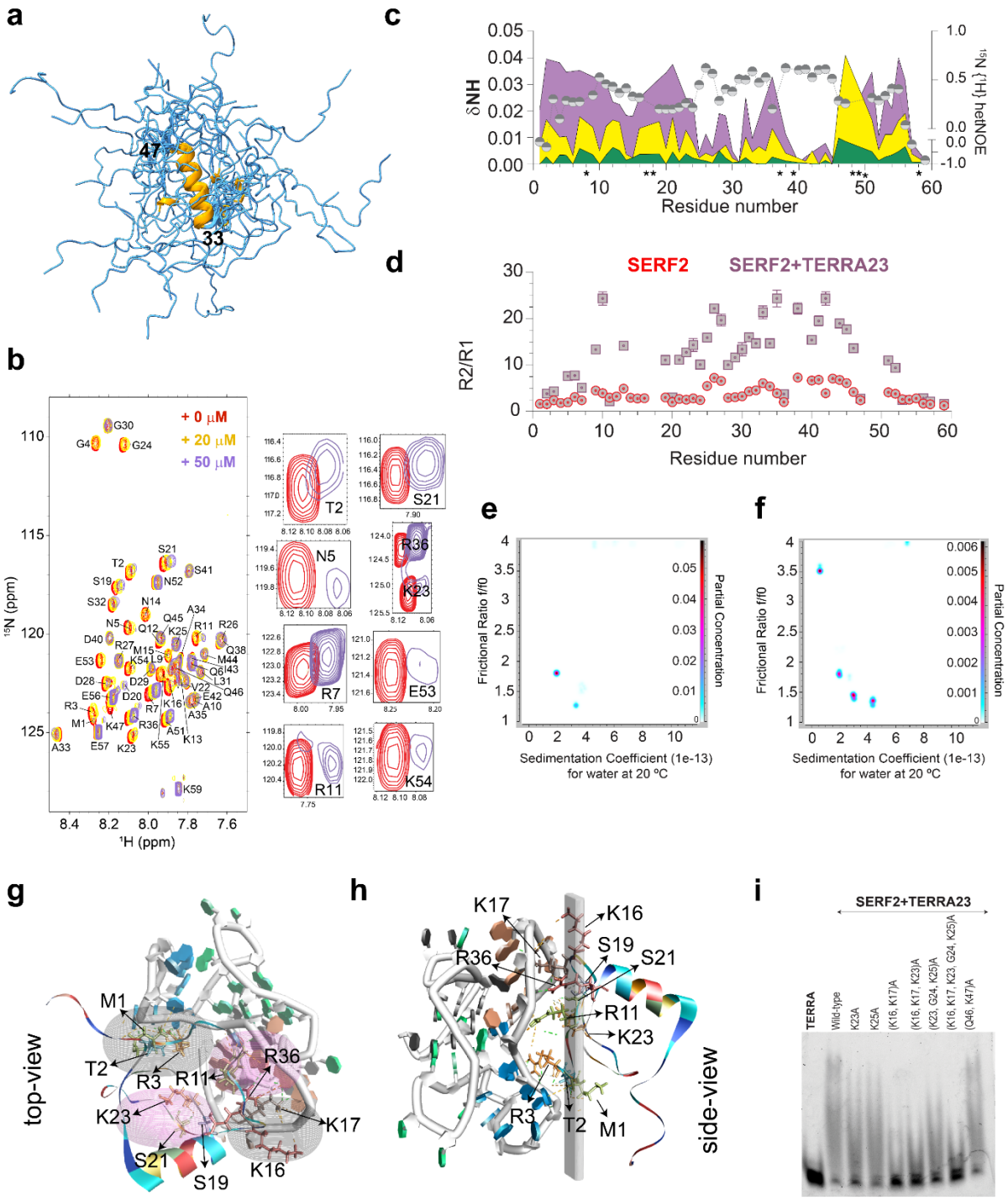
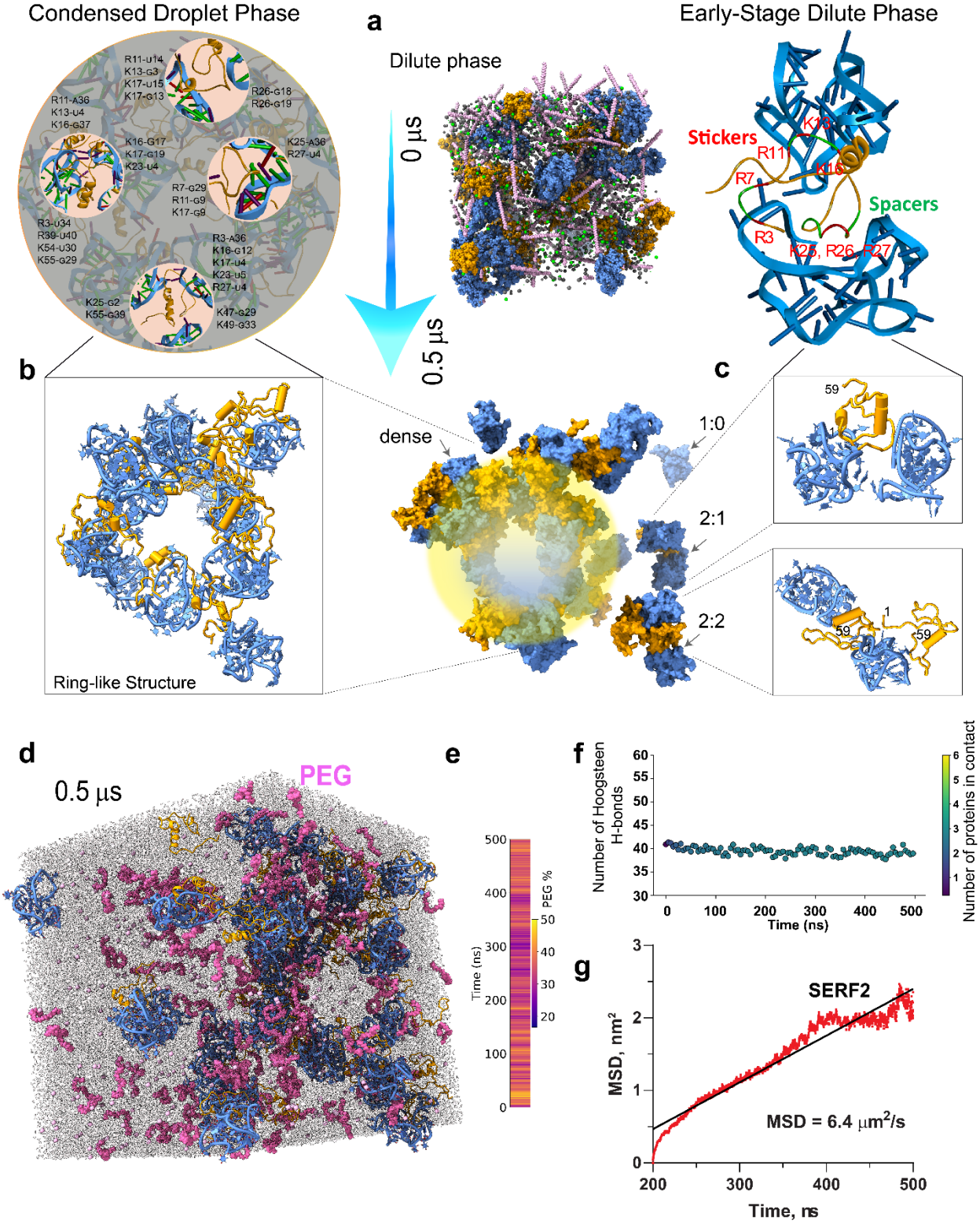
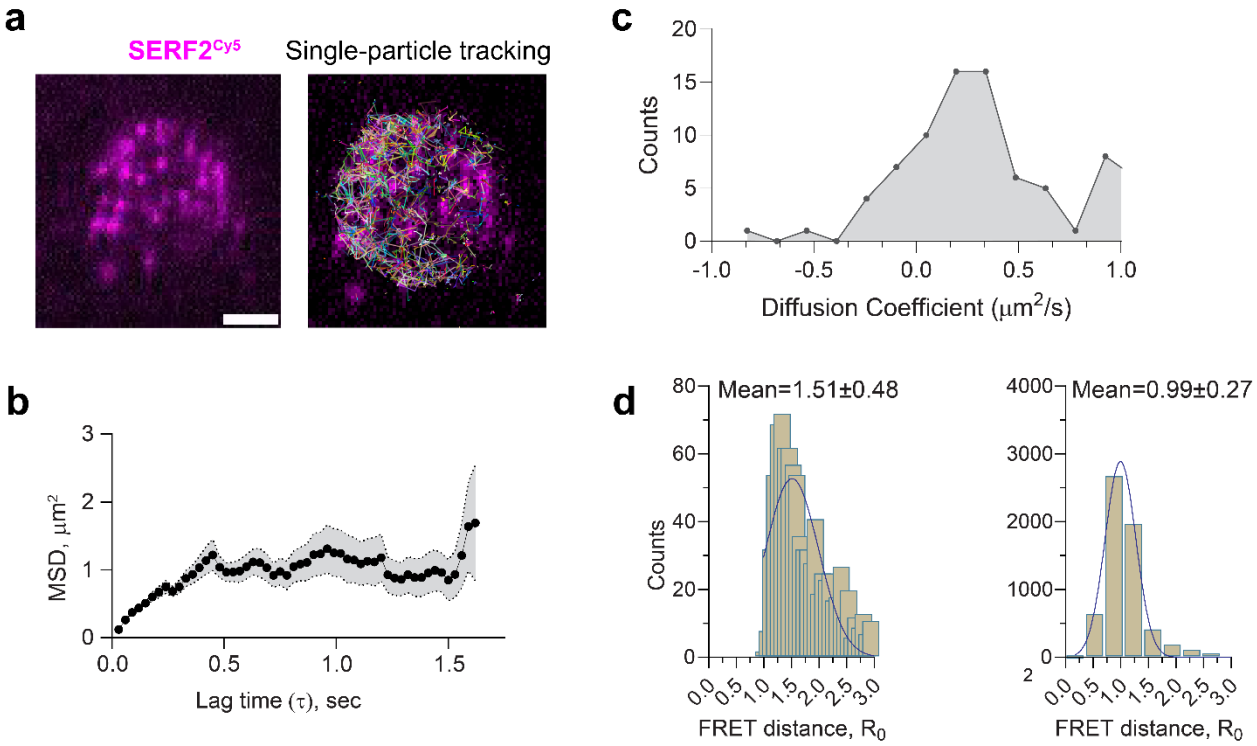

a
47
33
b
+ 0 μM
+ 20 μM
+ 50 μM
¹⁵N (ppm)
¹H (ppm)
c
δNH
¹⁵N {¹H} hetNOE
Residue number
d
SERF2
SERF2+TERRA23
R2/R1
Residue number
e
f
Frictional Ratio f/f0
Partial Concentration
Sedimentation Coefficient (1e-13) for water at 20 °C
g
top-view
h
side-view
i
SERF2+TERRA23
TERRA
Wild-type
K23A
K25A
(K16, K17)A
(K16, K17, K23)A
(K23, G24, K25)A
(K16, K17, K23, G24, K25)A
(Q46, K47)A

Fig. 4


Condensed Droplet Phase
a
Early-Stage Dilute Phase
Dilute phase
0 µs
0.5 µs
Stickers
Spacers
b
c
Ring-like Structure
dense
1:0
2:1
2:2
d
0.5 µs
PEG
e
Time (ns)
PEG %
f
Number of Hoogsteen H-bonds
Time (ns)
Number of proteins in contact
g
SERF2
MSD, nm²
MSD = 6.4 µm²/s
Time, ns

1673 **Fig. 5**



1674

Fig. 6

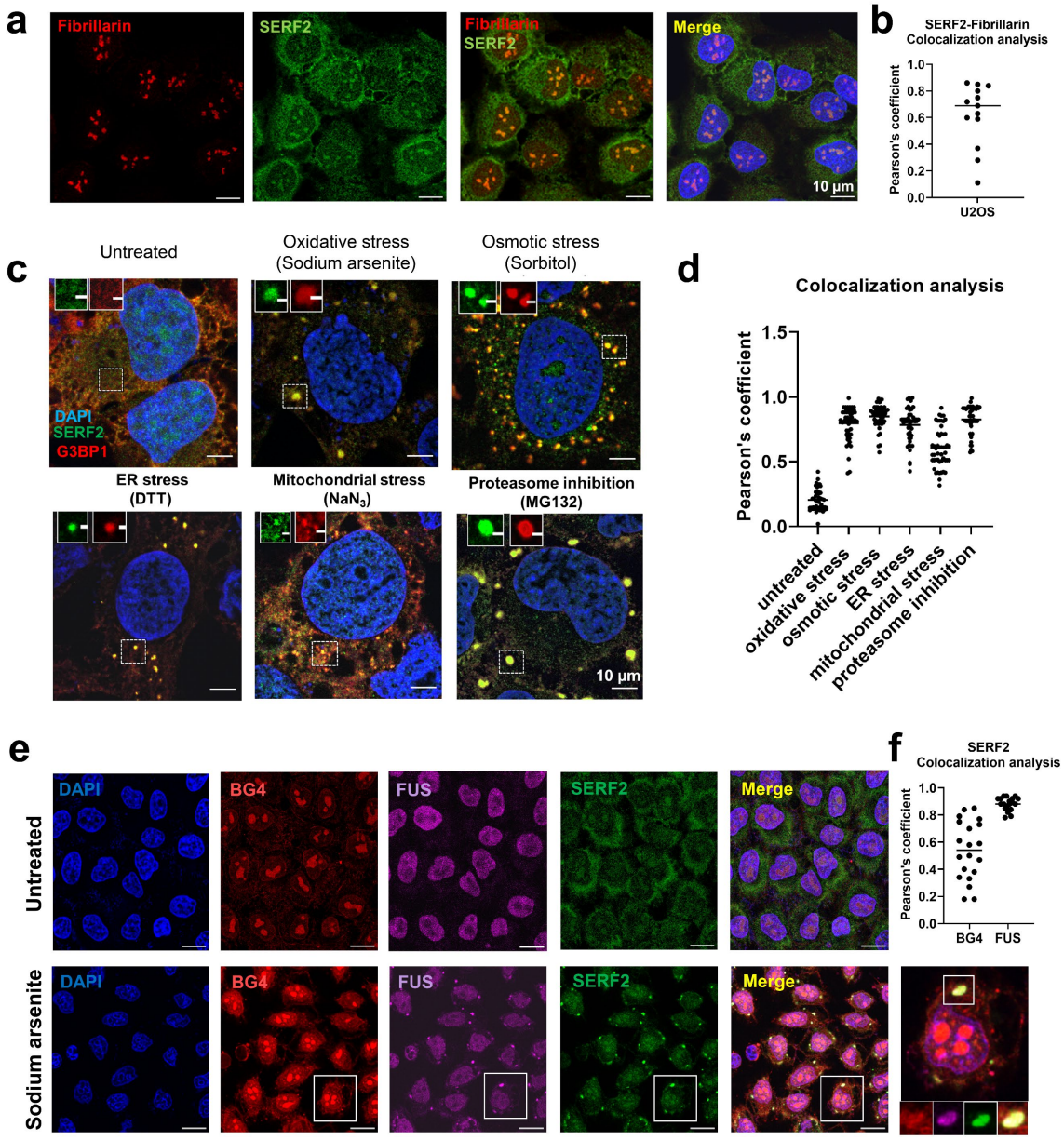
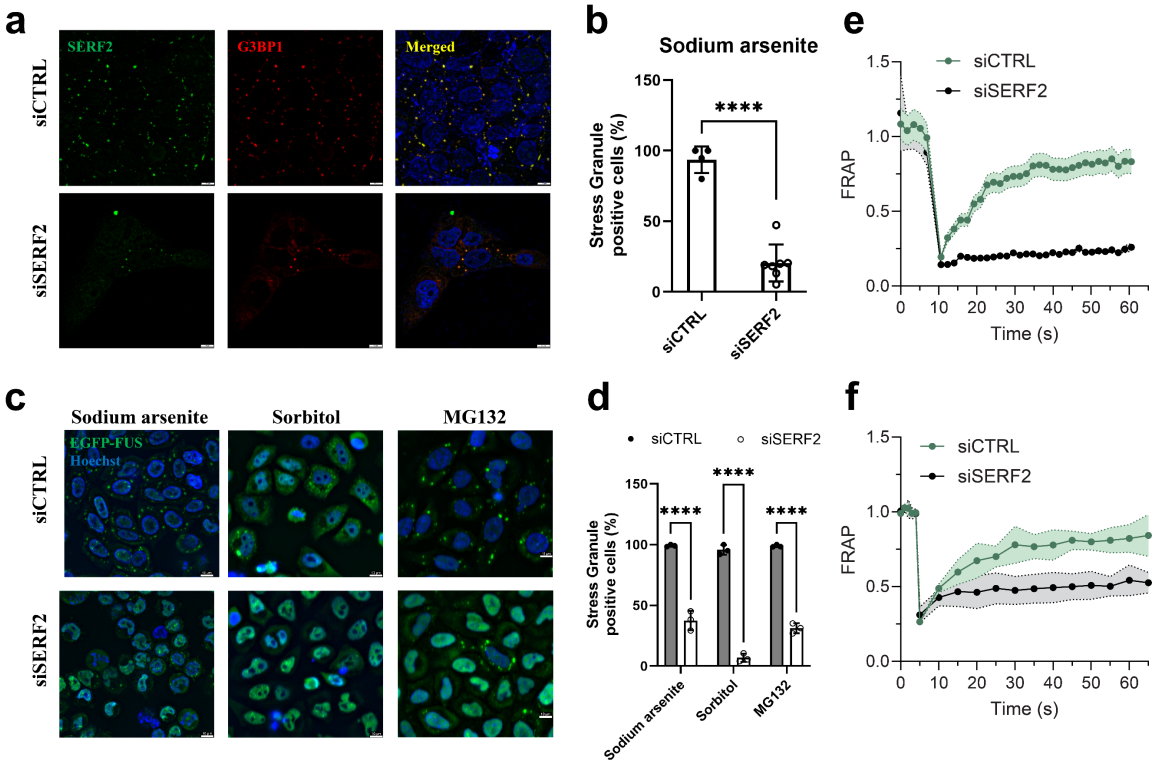
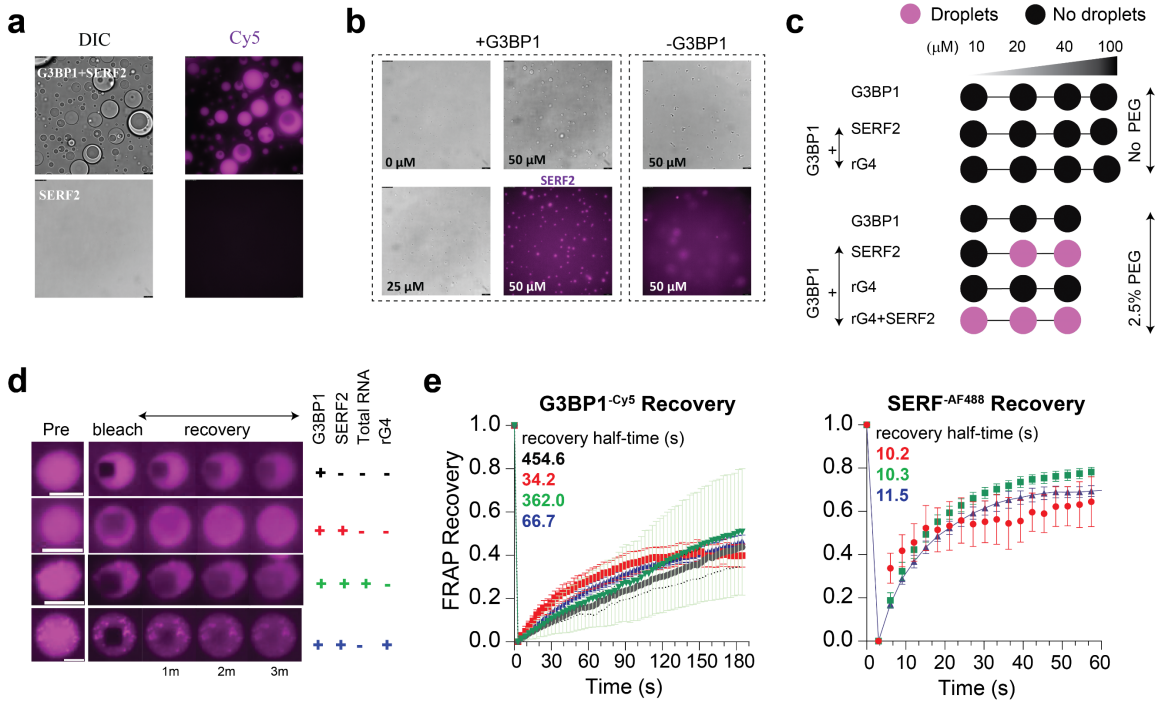

a
Fibrillarin
SERF2
Fibrillarin
SERF2
Merge
10 µm
b
SERF2-Fibrillarin
Colocalization analysis
Pearson's coefficient
1.0
0.8
0.6
0.4
0.2
0.0
U2OS
c
Untreated
Oxidative stress
(Sodium arsenite)
Osmotic stress
(Sorbitol)
DAPI
SERF2
G3BP1
ER stress
(DTT)
Mitochondrial stress
(NaN3)
Proteasome inhibition
(MG132)
10 µm
d
Colocalization analysis
Pearson's coefficient
1.5
1.0
0.5
0.0
untreated
oxidative stress
osmotic stress
ER stress
mitochondrial stress
proteasome inhibition
e
Untreated
Sodium arsenite
DAPI
BG4
FUS
SERF2
Merge
f
SERF2
Colocalization analysis
Pearson's coefficient
1.0
0.8
0.6
0.4
0.2
0.0
BG4
FUS

Fig. 7


a
SERF2
G3BP1
Merged
siCTRL
siSERF2
b
Sodium arsenite
Stress Granule positive cells (%)
****
siCTRL
siSERF2
e
siCTRL
siSERF2
FRAP
Time (s)
c
Sodium arsenite
Sorbitol
MG132
EGFP-FUS
Hoechst
siCTRL
siSERF2
d
siCTRL
siSERF2
Stress Granule positive cells (%)
****
Sodium arsenite
Sorbitol
MG132
f
siCTRL
siSERF2
FRAP
Time (s)

1733 **Fig. 8**
1734



1735
1736