SERF2, an RNA G-quadruplex Binding Protein, promotes stress granule formation

Bikash R. Sahoo,[1,2‡] Xiexiong Deng,[1,2‡] Ee Lin Wong,[1,2] Nathan Clark,[1,2] Harry Yang,[1,2] Vivekanandan Subramanian,[3] Bryan B. Guzman,[4] Sarah E. Harris,[5] Budheswar Dehury,[6] Emi Miyashita,[7] Hirohide Saito,[7] Daniel Dominguez,[4] 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]Bioinformatics Division, ICMR-Regional Medical Research Centre, Odisha-751023, India

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

[‡] equal contribution

Summary

Small EDRK-Rich Factor (SERF) is a member of a family of partially disordered proteins whose normal biological function is unclear apart from evidence they can speed disease-related amyloid formation. We show here that human SERF2 binds specifically to non-canonical RNA structures known as G-quadruplexes and plays an important role in stress granule formation. In cells, SERF2 colocalizes with stress granule marker proteins, and its depletion significantly affects stress granule size and abundance. In vitro, SERF2 undergoes liquid-liquid phase transitions in the presence of RNA G-quadruplexes. Structural analysis of the interactions between SERF2 and RNA G4 quadruplexes gives us a high-resolution view of the multivalent interactions and protein and RNA conformational changes that appear to represent some of the early steps in liquid-liquid phase transitions and stress granule formation.

INTRODUCTION

RNA G-quadruplex structures (rG4) are widely distributed throughout evolution, being particularly common in eukaryotic genomes. They have been detected in vivo,[1–3] and seem particularly abundant under stress conditions,[4–6] but their biological functions remain enigmatic.[7] These non-canonical four-stranded nucleic acid secondary structures are characterized by stacks of four-guanine planar structures called G-quartets.[1,8] G-quartets consist of guanine (G-G) bases that bond using Hoogsteen geometry and are stabilized by a centrally placed monovalent cation, characteristics that make G-quadruplexes very different from canonical Watson-Crick base-pairs.[9] rG4-specific antibodies and small-molecule probes have demonstrated rG4 localization to the nucleus, cytoplasm, mitochondria, and the endoplasmic reticulum of live cellls,[10–12] rG4 quadruplex forming sequences are enriched in untranslated regions of mRNA, ribosomal RNA, pre-mRNA, microRNA, long non-coding RNA, and the TERRA component of the telomerase complex.[13] Evidence links rG4 quadruplexes to a number of cellular processes including telomere maintenance, initiation of DNA replication, control of transcription and translation, and genetic and epigenetic instability.[14,15] The hexanucleotide repeat sequence G4C2, which is present in the *C9orf72* mRNA which is connected to the neurodegenerative disease ALS, forms stable RNAG4 quadruplexes.[13,16–19] rG4 quadruplexes are enriched in cell in stress conditions[4] bind disordered proteins and appear to be involved in the formation and organization of membrane-less compartments with liquid-like properties, such

as stress granules.[4,19,20] Despite the emerging importance of RNA G4 quadruplexes, relatively little is known regarding how their biological functions are regulated.

A number of proteins with disordered domains, including nucleolin,[21] helicases,[22,23] FUS,[24] hnRNPA1,[25] and TRF2,[26] have been shown to interact with rG4 quadruplexes and regulate quadruplex function. Helicases, for example, are thought to safeguard against G-quadruplex induced genome instability,[27] in contrast, nucleolin can act to enhance the stability of G4 quadruplex structures.[28] Understanding the mechanism and structural details of protein-rG4 interactions is crucial to understanding how proteins regulate the functions of these peculiar structures. Our structural understanding of these interactions is complicated by the dynamic nature of both the protein and rG4 quadruplex components that are involved in these interactions and the tendency of these complexes to undergo phase transition. Moreover, portions of proteins that bind rG4 quadruplexes are predominantly intrinsically disordered with low amino acid complexity.[7]

When cells are challenged with stresses that dissociate the translation complex, ribosome run-off occurs, and the resulting naked mRNA will recruit proteins to form a liquid-liquid phase compartment known as the stress granule.[29] Stress granules may be involved in RNA, or protein storage, though their exact cellular function is not yet clear.[30] Liquid-liquid phase compartments are relatively recently recognized, membrane-less compartments that are apparently involved in the coordination of various intracellular biological activities including signal transduction and transcription. Although we have a broad overview of the importance of multivalency and disorder in driving liquid-liquid phase transitions,[31] there is currently limited understanding of the structural basis of phase transitions or the structural changes that occur in macromolecules as they undergo phase transitions.[32]

SERF proteins were initially identified as a driver of amyloid formation, a process that has been linked to a number of age related diseases,[33–36] but their normal physiological functions remain obscure.[36] SERF proteins are remarkably small, averaging about ~70 amino acids in length, and they share a conserved, highly charged N-terminal domain.[36] Here, we present evidence that SERF2 facilitates stress granule formation through its ability to interact with rG4 quadruplexes, colocalizes with rG4 quadruplexes and stress granule marker proteins such as FUS and G3BP1. Multiple structural restraints were used to generate a high-resolution structural ensemble of SERF2 and SERF2-rG4 quadruplex complex that gives us a glimpse into the initial stages of liquid-liquid phase transition.

RESULTS

SERF2 is localized to stress granules

The SERF family proteins are rich in the amino acids E, D, R and K, from whence comes their name, Small EDRK-Rich Factor. SERF proteins were initially isolated for their ability to accelerate amyloid formation in a *C. elegans* model of Huntington's disease.[33] SERF1A has been reported to have a nuclear localization under normal growth conditions, but to be exported from nucleus into the cytosol upon stress treatment.[37] Under normal conditions we found that the related human SERF2 protein is also a least partially localized in the nucleus, but appears to be even more enriched in the nucleolus (Figure 1A). Following treatment with oxidative, osmotic, endoplasmic reticulum, mitochondrial and proteosome stresses (Figure 1B), but not heat stress (Figure S1A), we found that SERF2 forms prominent puncti in the cytoplasm. SERF2 colocalizes with several well-known stress granule marker proteins including G3BP1, FUS, TIA1, eIF2α and USP10 (Figures 1C, S1B) indicating SERF2 is likely a stress granule component.

SERF2 is involved in stress granule assembly

To test if SERF2 is important for stress granule formation, we knocked down SERF2 using RNAi and monitored the localization of a core stress granule marker protein, G3BP1 (Figure 2A). Sodium arsenite normally induces stress granule formation robustly.[38] Arsenate stressed SERF2 knockdown cells showed much more diffuse G3BP1 staining and significantly fewer and smaller stress granules than those formed in control knockdowns, in both U2OS cells and BJ fibroblast cells (Figures 2A, S2A). In a CRISPR-Cas9 knockout line of SERF2 in HEK293T cells, the number of stress granules after sodium arsenite treatment was also significantly reduced, as compared to those formed by arsenite treating wildtype HEK293T cells (Figure S2B). SERF2 depleted cells formed smaller stress granules (Figure 2B, C). We also observed a decrease of stress granule formation in live HeLa Kyoto cells bearing the stress granule marker EGFP-FUS (Figures 2D, S2C). Upon sodium arsenite, sorbitol or MG132 treatments which are all known to induce stress granules, SERF2 knockdown cells showed dramatically fewer EGFP-FUS cytosolic puncti as compared to the control knockdown cells (Figure S2C). These data strongly suggest SERF2 is involved in stress granule assembly or maintenance.

SERF2 binds RNA G4 quadruplexes

Stress granules are enriched in both RNA and ribonucleoproteins, and it has been reported that the SERF1a protein binds a 21-nucleotide long unstructured RNA species. We thus wondered if the related SERF2 protein can also interact with RNA. We used fluorescence polarization direct binding assays under physiological salt conditions and found that SERF2 failed to interact with polyA, polyC and polyU homo-polyribonucleotide sequences, or single or double stranded RNA sequences that contain hairpin structures (Figures 3A, S3A,B). However, we found that, polyG, (GGGA)4, (GGA)7, and (AGG)5 repeat RNA sequences did bind to SERF2 with sub-micromolar binding affinities (Figures 3A and S3). These repeat sequences are known to form RNA G-quadruplex structures.[39,40] To get a preliminary impression if SERF2 has some specificity in binding these non-canonical structures, we tested its interaction with two RNAs with similar sequences: (AGG)5, which folds into a rG4 quadruplex, and (ACG)5, that does not (Figure S3C).[40] SERF2 binds to (AGG)5 with a binding affinity (K_D) of 1.8 ± 0.5 μM, but shows no measurable binding to (ACG)5 (Figure S3C).

To search for SERF2 binding RNAs more unbiasedly, we utilized two independent high-throughput screening methods, RNA bind-n-seq (RBNS) and FOREST (folded RNA element profiling with structure). Both methods have been previously validated using known RNA binding proteins and structured RNAs including those that form G-quadruplexes.[41,42] RNA bind-n-seq, as illustrated in Figure 3B, is a high-throughput screening approach that uses a very diverse RNA pool that is generated by transcribing a randomized DNA library. The RNA species binding to SERF2 were fished out using three different conditions, binding reactions in a KCl buffer, which is known to promote G4 quadruplex folding, in a LiCl buffer which disfavors G4-quadruplex formation[43] and using an RNA pool synthesized using 7-deaza guanine in place of guanine, a substitution which eliminates G4 quadruplex formation[44] (Figure 3B). Features of RNAs that associated with SERF2 were identified by determining k-mer (k=6) enrichments.[41] In this case we assessed k-mers preferentially enriched in KCl (promoting G4 quadruplex formation) conditions vs those k-mers enriched in the 7-deaza RNA (G4 quadruplex breaker). The top 6-mers in this enrichment analysis are very rich in guanines (Figure 3C) and all are predicated to form G4 strong quadruplexes by the QGRS mapper tool.[45] Due to the massive RNA complexity of RBNS pools we were able to search for sequences predicted to form strong G-quadruplex structures as well as G-rich sequences that are not predicted to form G4 quadruplex structures. Strong G4 structures contained a repeated sequence motif $G_{(3-6)}N_{(0-7)}$ and the non G4 quadruplex forming sequences contained ≥ 8 guanines. We noted that strong G-

quadruplexes, were ~1.5-fold enriched for SERF2 binding in KCl while G-rich sequences that are not predicted to form G-quadruplexes were not enriched (Figure 3D). Furthermore, no G-quadruplex enrichment was observed in binding reactions in the presence of LiCl or with RNA pools made with 7-deaza, conditions known to disrupt rG4 structures (Figure 3D). These data support a model wherein SERF2 has a binding preference for RNA sequences predicted to fold into stable G4 structures.

To further probe if SERF2 has a binding preference for rG4 quadruplexes, SERF2 was used to screen an RNA structure library using the FOREST protocol.[42] We choose to use this approach for two reasons, first because it had been previously validated using known RNA G4 quadruplex forming sequences and G4 quadruplex binding proteins, and second the ~1800 human derived sequences in the FOREST library provide a diverse variety of folded RNA structures to screen for SERF2 binding.[42] The FOREST analysis identified a UG4U repeat sequence[46] as the top binding partner for SERF2 in this library and 8 out of the top 10 binding hits are G4 quadruplex forming sequences.[42] Overall, the binding intensities of SERF2 for the G4 quadruplex sub-library were statistically significantly higher (4.12×10^{-24}) than to the other 1800 sequences in the library (Figure 3E). The significance of the binding differences of SERF2 for the rG4 quadruplexes as compared to the rest of the library correlates to that observed previously for the anti rG4 quadruplex BG4 antibody (8.8×10^{-41}) and cold-induced binding protein (1.4×10^{-5}).[42]

We next measured the in vitro binding affinity of SERF2 to one of the top FOREST hit UGGGGU (UG4U) six-repeat sequence and to two additional well characterized rG4 quadruplex forming sequences, the TERRA repeat-containing RNA (UA GGG UUA GGG UUA GGG UUA GGG), which is found in human telomeres and the GGGGCC (G4C2)4 hexa-nucleotide repeat that occurs within the *C9orf72* gene, whose expansion is associated with Amyotrophic Lateral Sclerosis (ALS).[47,48] SERF2 bound to all three of these rG4 quadruplex forming sequences with low micromolar K_D values (G4C2: 0.9±0.1 μM; TERRA: 0.30±0.02 μM; and UG4U: 0.8±0.3 μM) using a fluorescence polarization assay (Figure 3F). Combined, the above data points towards a tendency of SERF2 to interact with rG4 quadruplex forming sequences.

RNA G-quadruplexes in stress granules are colocalized with SERF2

It has recently been shown that rG4 quadruplexes are more abundant under stress conditions, accumulate in stress granules and may be involved in their formation.[4,6] Therefore, we next checked if SERF2 and rG4s are co-localized in under stress conditions. We found that SERF2 co-localizes with the anti-G-quadruplex antibody, BG4, under normal conditions (Figure 1C), and we were able to reproduce previous findings that rG4 quadruplexes are found in stress granules.[4,6] This, together with our prior demonstration that SERF2 is also found in stress granules, and colocalizes with rG4 quadruplexes (Figure 1C) is suggestive of SERF2 interacting with G4 quadruplexes in vivo.

SERF2 forms liquid-liquid phase droplets with RNA G-quadruplexes

The nucleolus and stress granules are two types of membraneless compartments whose formation is mediated by liquid-liquid phase separation,[49–51] a phenomena that has recently been linked to a variety of cellular processes including RNA processing, chromatin organization, ribosome assembly, and a range of diseases.[52] Since rG4 quadruplexes have recently been shown to mediate liquid-liquid phase separation, we next investigated if SERF2 binding to rG4 quadruplexes can trigger phase separation. In isolation, SERF2 remains soluble in various salt (0-200 mM KCl) and crowding (0 or 20 % PEG) conditions (Figure S5A,B), even up to 1 mM in protein concentration. Phase separation was also not observed for SERF2 when are mixed with RNA or rG4s under non-crowding conditions (Figure S5A, S6A). However, in the presence of the crowding agent PEG8000, SERF2 phase separates when mixed with HeLa total RNA at various protein and RNA concentrations (Figure 4A, B). G4 quadruplexes have recently been identified to be able to induce proteins to undergo phase separation,[53] so we asked if rG4 quadruplexes can induce SERF2 condensation. We first confirmed rG4 quadruplexes (TERRA, G4C2, and UG4U) on their own do not form droplets (Figure S5C) in crowding conditions (10% PEG8000). However, SERF2 forms X-YD micron-sized droplets when mixed with all three rG4 quadruplexes tested (Figure 4C-E).

To better understand the phase regime behavior of SERF2-rG4 quadruplex condensate formation, we decided to focus on the TERRA rG4 quadruplex due to its previous extensive structural characterization, ability to be direct visualized within living cells, its structural similarity with many other parallel rG4 quadruplexes, and its tighter binding to SERF2 as compared to it binding to the G4C2 and UG4U repeat-sequences rG4 quadruplexes.[54] We found that SERF2-

TERRA droplets were formed at a variety of salt, PEG8000, protein, and RNA concentrations (Figures 4F, G, S6A). An excess of either RNA or SERF2 can disrupt the phase regime (Figure 4A, G) consistent with previous observations with other RNA-protein combinations.[31,55] Though SERF2 phase separates when mixed with the rG4 quadruplex forming TERRA sequence, it fails to when mixed with similar sized polyU and polyA RNAs which do not form rG4 quadruplexes (Figure S6B). We postulate that SERF2- binding to rG4 quadruplexes also drives phase separation in vivo, affecting the formation of stress granules, a well characterized cellular liquid-liquid phase compartment.

SERF2 and RNA G-quadruplex form slowly exchanging droplets

SERF proteins have been postulated to be cellular drivers of amyloid formation.[36] We demonstrate here that SERF2 also interacts with RNA in a way that can lead to liquid-liquid phase separation. These observations motivated us to look at the dynamics of the interactions of SERF2 with RNA in part because the transitions of liquid droplets to more solid aggregates have been linked to neurodegenerative diseases.[56] Fluorescence recovery after photobleaching (FRAP) experiments showed that SERF2 interacts in a rather dynamic fashion with bulk HeLa total RNA, showing recovery halftime of 13± 2 seconds (Figure 4H). In contrast, SERF2-rG4 droplets show considerably slower exchange dynamics for both the protein and rG4 quadruplex components, with halftime recovery ranging from 80 to 620 seconds (Figure 4C-E, bottom) indicating that SERF2 and rG4 quadruplex droplets are less fluid than those it forms with bulk RNA.

SERF2 and RNA droplets facilitates G3BP1 condensation

We show here that SERF2 is important for stress granule assembly and others have proposed G3BP1 to also be key in this process[49]. We thus asked if SERF-RNA interaction might affect G3BP1 condensation[57]. We found that SERF2 undergoes liquid-liquid phase separation in the presence of G3BP1 *in vitro* (Figure 4I) and that as SERF2 concentrations increase, one can observe remarkably large condensates composed of G3BP1 and total RNA (Figure 4J). These large condensates were not seen when total RNA is absent (Figure S6C). These results suggest that SERF2 can promote G3BP1-RNA condensation in vitro, seemingly reflecting the *in vivo* phenotype of SERF2's depletion which results in in smaller and fewer stress granule foci (Figure 2C).

Human SERF2 is partially disordered

To characterize the overall secondary structure of SERF2 we first performed CD spectroscopy. CD spectroscopy of SERF2 at 4 °C indicates a high proportion of helical content, but this helicity decreases substantially at physiological temperatures (Figure 5A). The narrow 15N/1H NMR chemical shift dispersion (~7.5 – 8.5 ppm) for SERF2 is additional evidence for structural disorder in SERF2, even at 4 °C (Figure S7). That the majority of its 15N/1H cross peaks are undetectable upon temperature upshift is suggestive of chemical exchange in the NMR measurement timescale. This could be due to the induction of different conformational states in the disordered regions in SERF2, a phenomenon that has been demonstrated earlier for other disordered proteins[58] which at low pH are detectable as demonstrated earlier[59] (Figure S7). We were able to assign the backbone (N, Cα, Cβ, NH, CO) NMR chemical shifts for 51 residues out of 58 non-proline residues and predicted the torsion angle restraints for the ϕ and ψ angles in SERF2 using TALOS-N program.[60] In addition, 835 NOEs restraints proved to be obtainable from 3D 15N-HSQC-NOESY measurements. In combination, these NMR restraints allowed us to build the SERF2 structural ensemble as shown in Figure 5B. The N-terminal domain (residues 1-36) of SERF2 is very dynamic, and present in multiple structures with a short helix spanning residue 9-13 (Figure 5B). SERF2 retains an average helical structure that spans residues 37-46, which is shown in green in Figure 5B which serves to anchor both the N- and C-terminal more dynamic structures.

A surprisingly large proportion of RNA interaction motifs are known to be disordered,[61] but it is not yet clear the role that this disorder plays in RNA-protein interactions. To expand our understanding of the mesoscopic phase behavior seen in SERF2 and RNA droplets we asked how SERF2 and rG4 interactions impact their structural dynamics. To map the protein binding sites in the model RNA G4 quadruplex TERRA, we first monitored the 15N/1H chemical shift perturbations that occur in SERF2 with increasing TERRA concentrations (Figure 5C). TERRA binding induces major chemical shift perturbations in both the N-terminal (3-21) and C-terminal 51-56 residues of SERF2 (Figure 5C, D). Saturation transfer difference NMR measurements revealed magnetization transfer from guanine imino protons to SERF2 (amide and methyl regions, red spectrum, Figure S8C, bottom) that are absent in TERRA sample alone (Figure S8A,B). Further, the saturation of the A33 amide proton in SERF2 (Figure S8D), which conveniently does not overlap with TERRA signals, shows magnetization transfer to TERRA guanine imino protons (Figure S8C, top). This observation suggests that the C-terminal residues in SERF2 are in proximity of TERRA binding and that several SERF2 residues are spatially

oriented close to the TERRA G-tetrad core. By studying TERRA interaction with individual domains in SERF2, we showed that the N-terminal domain (1-32), on its own, binds rG4 quadruplexes, though that binding is ~ 10-fold weaker as compared to the binding of full length SERF2. On the other hand, the C-terminal domain (31-59), in isolation, showed no binding saturation in a fluorescence polarization assay, implying that, on its own, the C-terminal domain interacts very weakly with SERF2 (Figure S9). Together these results suggest both N- and C-termini in SERF2 coordinate with one another in forming a SERF2-TERRA complex in solution but that the N-terminal domain is likely more important for rG4 recognition.

We next decided to test if SERF2 dynamics are associated with rG4 binding. By comparing the hetNOEs relaxation data that map the dynamic regions in SERF2 in the absence and presence of TERRA, we observed a strong correlation between the regions in SERF2 that are highly dynamic and those that are involved in TERRA binding (Figure 5D). The flexible, unstructured, or partially structured regions, which span residues 3-24 and 46-56, are characterized with an average hetNOEs < 3 showed the highest chemical shift perturbations and change in peak signal height upon addition of TERRA (Figure S10).

TERRA binding to SERF2 at 1:2 RNA to protein ratio displayed a considerable change in the peak signal height for several residues, which became stronger upon a further increase in TERRA concentration to a 1:1 ratio (Figure S10). The reduction of peak intensities in SERF2 are inversely correlated to spin-lattice or T2 relaxation rates and may arise from several factors such as conformational exchange upon TERRA binding, an increase in complex size leading to decreases in correlation time, and amide proton exchange with the bulk solvent. To determine whether TERRA binding has a significant effect on SERF2 dynamics, the backbone relaxation rate (R2/R1) of individual residues was calculated where a lower R2/R1 number represents a higher mobility. TERRA binding reduces the average tumbling in SERF2 explaining the overall increase in the R2/R1 values for most of the residues as compared to SERF2 alone (Figure 5E). These slow dynamics were observed for most of SERF2 except for a few N- (R11 and D20) and C-terminal (53-59) residues. The slowed dynamics and peak broadening are likely due to the increase in size that occurs upon protein-TERRA complex formation or multimerization. That the structurally disordered and dynamic region in SERF2 is directly involved in the interaction with a rG4 quadruplex could mean that disorder is a prerequisite for rG4 quadruplex recognition, or it could more simply reflect the gain in order that often accompanies rG4 binding.

Multivalent interactions drive liquid-liquid phase separation

Multivalent interactions are a prerequisite for in vitro phase separation.[62] Disordered proteins often are involved in multivalent interactions and thus are prone to form liquid-liquid phase separations, but the precise species that nucleate phase transitions is unclear. Since SERF2 is involved in phase separations with RNA, and since it well behaved biophysically, we thought that we might be in a good position to understand at a structural level the initial interactions that take place during phase transitions. Using a gel shift assay we observe an increase in the size of protein-TERRA complexes with increasing protein concentration (Figure 5F). Size-exclusion chromatography analysis also demonstrated the existence of at least three different-sized species that absorb at 260 nm, unbound TERRA, which eluted at ~15 kDa, a 1:1 protein-TERRA complex that eluted at ~30 kDa, and a multimeric complex of >30 kDa (Figure S11). As SERF2 has no absorbance at 260 nm, the detected peaks are only arising from TERRA.

To determine the protein-TERRA stoichiometry in these multimeric species, we performed analytical ultracentrifugation. In isolation, the major TERRA species have sedimentation coefficients of 1.95 and 1.80, suggestive of globular, folded structures (Figure 5G). However, in a SERF2:TERRA (2:1) mixture, three additional species are observed that sediment with higher coefficient values, two of these, which sediment at 3.00 and 4.38 appear to be globular, and one with frictional ratio 3.5 that corelate to an elongated/unfolded TERRA species (Figure 5H). The estimated sizes for the globular species are ~ 12.4, 17.4 and 27.1 kDa for sedimentation values 1.95, 3.00 and 4.38, respectively, corresponding to 1:0, 1:1 and 1:2 TERRA:SERF2 complexes. The estimated size for the unfolded TERRA species was ~5.8 kDa suggesting no complex formation with SERF2. This size-distribution analysis indicates that SERF2 forms multimeric species in complex with rG4 quadruplexes, suggesting their tendency to phase separate. The TERRA binding sites in SERF2 obtained from NMR measurements were used in combination with atomic simulation to build TERRA-SERF2 complex structure. By doing so we hoped to obtain some of the first ever glimpses into the complex structure and dynamics that accompanies liquid-liquid phase transitions. Using HADDOCK,[93] we first built the structures of 1:1 and 1:2 TERRA and SERF2 complexes that we had detected in our analytical ultracentrifugation experiments. We did this by parsing a set of ambiguous active site residues to HADDOCK program that were obtained from NMR titration and saturation transfer difference measurements. The docked structure of the 1:1 TERRA:SERF2 dimer obtained after a 1 μs atomistic simulation showed SERF2 interacts with TERRA primarily using its N-terminal

residues (Figure 6A). In the 2:1 TERRA:SERF2 complex, the C-terminal contacts are gained in addition to N-terminus contacts, suggesting that the multivalent interactions are necessary for oligomeric species (Figure 6A, C). These results are consistent with the observed increases in chemical shift perturbations for C-terminal residues seen upon increasing TERRA concentration (Figure 5D). We note that change in chemical shift perturbations could also arise from conformational alteration in the dynamics of SERF2. We obtained evidence that this is indeed occurring from our saturation transfer difference NMR experiments that showed that upon saturation of C-terminal residues, transfer magnetization occurs to TERRA guanine imino protons (Figure S8).

A hydrophilic core in the complex structure is formed by SERF2 residues that are located in regions that are dynamic and disordered in monomeric SERF2 (Figure 6B, C). This hydrophilic core constrains SERF2 dynamics via generating a planar contact surface that mounts on planar G-tetrads through a quadrupolar-like contact architecture (Figure 6C,D). Several N-terminally located charged residues (R3, R11, K16, K17 and K23) showed a high propensity to form hydrogen bonds with the G1, G3, G7, G8 and G9 residues located in the bottom two G-quartet that is formed by two intermolecular TERRA molecules (Figure 6C). The planar and quadrupolar-like contact surface between SERF2 and TERRA is formed by a highly charged surface facing opposite to a surface comprised of polar residues (Figure 6D). The contact surface form by the first three N-terminal SERF2 residues M1, T2 and R3 are facing opposite to the binding surface formed by two charged residues K16 and K17. Similarly, the charged residues R11and R36 in one contact surface face opposite to the polar surface formed by residues S19, S21 and K23 (Figures 6C).

The guanines in the tetrameric TERRA[63] are bonded in a nucleotide number ‘i’ to ‘i+6’ pattern to form 5 neatly stacked G-tetrads (Figure 6G). Interestingly, the simulation results reveal two G-quartets become disoriented upon SERF2 binding. The G9-G3 Hoogsteen bond is disrupted by R11 and R3 of SERF2, and on the other side S19-S21-K23 and M1-T2 interaction solvent exposed the buried U4 that stabilizes the TERRA tetrameric structure via H-bond networking with G2 and G3 (Figure 6C,E). K16-K17 forms hydrogen bonds with G7 and G8 and interrupts the G7-G1 and G7-G8 Hoogsteen base pairing (Figure 6C,E). The hydrophilic core in the 2:1 SERF2:TERRA complex is bigger than that present in the 1:1 complex due to the incorporation of several charged C-terminal residues including R36, K50, K54 and K55 that form hydrogen bonds with TERRA (Figure 6E). Hydrogen bond profiling as plotted versus atomic

simulation length showed that SERF2 has more tendency to interact with uracil nucleotides in the 1:1 complex which shift towards guanine nucleotides in the 2:1 complex (Figure 6F). SERF2 binding in the 2:1 complex increasingly distorts the TERRA's G4 quadruplex structural integrity. To see if we could obtain experimental evidence for this structural distortion, we next studied TERRA secondary structure in the presence of SERF2. The parallel structure content of TERRA as measured by CD (Figure S12) gradually decreases with increasing SERF2 concentration, consistent with SERF2 binding acting to distorting the parallel structure of the TERRA G4 quadruplex (Figure S12). These data correlate with the atomistic MD simulation results, though the MD simulation appears to overrepresent the structural distortion in the TERRA G-quadruplex than the CD indicates. The images shown in Figure 6 represent some of the high-resolution structural information concerning the interactions and conformational changes which may correlate to their early stages of liquid-liquid phase separation.[32]

DISCUSSION

The treatment of cells with a variety of stressors that act to halt translation generates naked RNA which is rapidly sequestered into a liquid-liquid compartments known as stress granules. Though the precise function of these granules is unclear, they may be involved in RNA or protein storage during stress.[64] Many of the proteins that are involved in stress granules contain elements of disorder and understanding how disordered proteins function and interact with other molecules is limited by a lack of experimentally determined conformational assembles for these rather difficult to handle proteins. We have a broad overview of factors affecting stress granule formation and liquid-liquid phase transitions including the importance of multivalency and disorder. However, a lack of detailed structural information leaves both the protein and RNA components of these membrane-less compartments often diagramed as vague spaghetti-like lines. Our replacing these vague lines with the structural and biophysical details diagrammed in figure 6 helps us understand the process of stress granule and liquid-liquid phase separation and the interactions that take place within these recently recognized compartments.

Here we show that SERF2, initially isolated by its ability to accelerate amyloid formation, is not just a component of stress granules but also appears to be important for their formation. A number of other proteins are thought to be important for the formation of stress granules, their overproduction also triggers stress granule formation in the absence of stress and their depletion inhibits the formation of these granules.[29,87] Transcriptomic[65,66] and proteomic[67–69] studies have sought to delineate a comprehensive picture of stress granule components, so it

may seem initially surprising that SERF related proteins do not occur either on the extensive lists of the protein components of stress granules or the much smaller lists of proteins involved in their formation. However, the high abundance of lysine and arginine in SERF has been shown to result in such extensive digestion by trypsin that becomes essentially invisible to the trypsin proteolysis-based mass spectrometry approaches[70] that have generally been used in helping to generate these lists.

Two high-throughput screening approaches supported by biophysical experiments show that SERF2 shows a strong tendency to bind to RNA G4 quadruplex structures. rG4 quadruplexes are non-canonical RNA structures containing Hoogsteen base pairs. As an emerging cellular component,[9–13] rG4 quadruplexes may function as translational repressors and transcriptional regulators, as well as being involved with mRNA processing, mRNA polyadenylation and splicing, telomere maintenance, and RNA translocation, [3,10,14,16,47,71–74] although their mechanism of action remains unclear. [36,75] In findings likely relevant to ours, it had recently been found that stress enhances the amount of rG4 quadruplexes present,[4] rG4 quadruplexes are abundant in stress granules, stress promotes rG4 quadruplex folding[4] and RNA quadruplexes may facilitate the formation of stress granules.[5,76]

A handful of RNA binding proteins, such as FMRP, nucleolin, CNBP, eIF4A, hnRNPA1, and DHX36, have been shown to bind rG4 quadruplexes, and modulate their folding.[20–26] SERF2 binds to known rG4 quadruplex structures with sub-micromolar binding affinities similar the affinities of other rG4 quadruplex binding proteins such as cold-inducible RNA-binding protein,[77] FUS,[78] and FMRP,[82] as well as G4 quadruplex binding small-molecules such as pyridostatin, NMM, and BRACO-19.[79] A common feature of reported rG4 quadruplex binding proteins is the engagement of an intrinsic disordered domain in the binding, however a lack of detailed structural and residue level dynamic information in these complex systems has limited our understanding of the role of disorder in these systems. Upon interaction of SERF2 with G4-quadruplexes, a planar G-tetrad RNA-protein interaction forms through a quadrupole-like interaction. The dynamics of SERF slow and it becomes more rigid and the rG4 quadruplex structure becomes distorted (Figure 6). The conformational adaptability of SERF2 in interacting with the TERRA G4 quadruplex structure shows the importance of its structural disorder in modulating the parallel topology of TERRA G4 quadruplex structure. Planar interaction of G4 quadruplexes with binding partners has been observed before, for instance, yeast Rap1 interacts with G4 quadruplexes using planar G-tetrads stabilizing the G4 quadruplex.[85] In

contrast the DHX36 helicase unfolds G4 quadruplexes[80] via forming a flat non-polar surface on the G-quartet in a manner somewhat similar to that of SERF2. Planar interactions of small molecules with G4 quadruplexes have been used in structure-based small-molecule design for both G4 stabilization[81] and destabilization.[83]

If rG4 quadruplexes thermodynamically stabilize the SERF2 protein they may affect the levels of SERF2 in vivo and thus its ability to affect stress granule formation.[84] As G-quadruplex folding elevates the local charge density, it may be possible that rG4 quadruplexes in stress granules not only attract charged misfolded proteins, but also limit the rate of stress granule dissociation. Notably, unlike droplets that have undergone a liquid-to-gel transition or aggregates which are characterized with extremely slow or no measurable exchange,[86] droplets that form between SERF2 and rG4 quadruplexes are still reversible, consistent with the dissipation of stress granules seen upon removal of stress conditions.[87]

Given SERF's action in speeding protein aggregation and our observations that SERF2 is important for stress granule formation and can engage in liquid-liquid phase separation, it is tempting to consider possible links between these processes. Persistent stress introduced by the overexpressing amyloidogenic proteins can lead to long-term stress granule formation.[88] Within the high concentration environment of the stress granule, fibrillization of amyloid may be accelerated by SERF2. Alternatively, by binding G4 quadruplexes, which have previously been shown to function as very potent anti-aggregation agents,[89] SERF2 may affect the maturation of liquid-liquid phase separations into pathological solid-like aggregates. An appropriate balance of the anti-aggregation behavior of G4 quadruplexes and the pro-aggregation properties of SERF2 may be necessary for maintaining liquid-liquid phase separation droplets in a dynamic and reversible state and preventing solidification reactions that have an irreversible effect on protein structure and function.

LIMITATIONS OF THE STUDY

Though the current work provided structural insights into SERF2 binding to a model rG4 quadruplex TERRA and how it distorts the TERRA rG4 structure, future work will be necessary to probe if it interacts with in vivo rG4 quadruplexes binding partners, in stress granules or elsewhere, in a similar or different manner. To enable these studies, it would be necessary to determine which of the ~100,000 rG4 quadruplexes that are predicted to exist in vivo that SERF2 interacts with. Attempts to use enhanced UV-crosslinking immunoprecipitation to isolate

SERF2 binding RNAs were unsuccessful, possibly due to lack of aromatic residues in SERF2 protein, which causes a low frequency of UV crosslinking. In addition, a commercially available antibody against SERF2 could not recognize RNA-bound SERF2. This is not surprising, as structural studies make it evident that SERF2 utilizes both its N- and C-terminus residues in binding RNA, possibly interfering with antibody recognition. Epitope tagging was attempted to provide another way of pulling down these in vivo SERF2-RNA complexes, but the epitopes tested interfered with the function of this tiny protein. Transient overexpression using tagged SERF2 not only altered the localization of these fusion proteins in cell, but also elevated non-specific nucleic acid interaction. A better understanding of these structure-function relationships will aid us in understanding rG4 quadruplex linked biological functions such as gene regulation and stress granule formation.

STAR METHODS

Cell Culture, Treatment, and Transfection

All cells were cultured in Dulbecco's Modified Eagle Medium (Fisher Scientific, 11-995-073) supplemented with 10% heat inactivated fetal bovine serum (Sigma, F4135) and 1X Penicillin-Streptomycin-Glutamine (Fisher Scientific, 10-378-016). Stress treatments were performed by treating cells with vehicle alone or different stress inducers. Cells were treated with a final concentration of 500 μM sodium arsenite to induce oxidative stress. For ER stress, cells were treated with a final concentration of 2 mM Dithiothreitol for 30 minutes in culture medium at 37°C. To achieve proteasome inhibition, cells were treated with 10 μM of MG132 (Sigma, 474790) diluted in culture medium and incubated for 30 minutes at 37°C. For mitochondrial stress, cells were treated with 75 mM NaN_3 for 30 minutes at 37°C. For heat shock cells were incubated in 43°C for 1 hour. For plasmid transfection, cells were transiently transfected using Lipofectamine™ LTX and PLUS™ reagent (Fisher Scientific, 15338030) according to the manufacturer's instructions. For knockdown, 13 μM of control (Horizon Discovery, D-001206-13-20) or SERF2 (Horizon Discovery, M-016317-01-0010) siRNAs were transfected using RNAiMAX reagent (Thermo Fisher, 13778150). After 48 h of incubation, transfected cells were harvested for western blot analysis, and RT-qPCR.

Immunofluorescence

Cells grown on coverslips were treated using the indicated conditions and then fixed with 4% paraformaldehyde (Electron Microscopy Sciences, 157-8-100) in 1X DPBS (Fisher Scientific, 14-190-144) at room temperature for 10 minutes. Fixed cells were washed three

times with 1X DPBS and permeabilized with 0.1 % Triton X-100 at room temperature for 30 minutes. Permeabilized cells were then blocked at room temperature for 1 hour using UltraCruz™ blocking reagent (Santa Cruz, sc-516214), followed by incubation with primary antibody solutions for 1 hour at room temperature. Primary antibody solutions were prepared in UltraCruz™ blocking reagent using the following dilution factors: 1:200 rabbit anti-SERF2 (Proteintech, 11691-1-AP), 1:1000 mouse anti-G3BP1 (BD Biosciences, 611127), 1:1000 mouse anti-Fibrillarin (Boster Bio, M03178-3), 1:100 mouse anti-USP10 (Santa Cruz, sc-365828), 1:200 mouse anti-FUS (Thermo Fisher, 50-554-337), 1:500 mouse anti-BG4 (Absolute antibody, Ab00174-1.1), 1:100 mouse anti-TIA1 (Santa Cruz, sc-398372), and 1:100 mouse anti-eIF2α (Santa Cruz, sc-133132). Cells were then washed three times with 1X DPBS and incubated in secondary antibody solutions in dark at room temperature for 1 hour. Secondary antibody solutions were prepared in blocking reagent using 1:1000 dilution, goat-anti-rabbit secondary Alexa Fluor Plus 488 (Thermo Scientific, A32731) and goat-anti-mouse secondary Alexa Fluor 647 (Thermo Scientific, A21235). After three washes with 1X DPBS, cells were stained with 0.25 µg/mL DAPI (Thermo Scientific, D1306) in 1X DPBS for 3 minutes. Cells were further washed three times with 1X DPBS, air dried, and mounted with ProLongGlod™ antifade reagent (Cell Signaling, 9071S) on glass slides. After 24 hours mounting, slides were sealed with nail polish and imaged using Leica SP8 confocal or Thunder™ microscope.

Immunoblotting

Cell lysate from ~ 2×10^5 cells were loaded in 4-15% Mini-PROTEAN™ TGX Stain-Free™ protein gels and resolved in Tris/Glycine/SDS running buffer (25 mM Tris, 192 mM Glycine, 0.1% w/v SDS, pH 8.3). Proteins were transferred to nitrocellulose membranes using Trans-Blot Turbo™ transfer system (Bio-Rad) and RTA™ transfer kit (Bio-Rad, Cat #170-4270). Membranes were then blocked with 5% non-fat milk in 1×TBST (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% v/v Tween20) at room temperature for 1 hour. The blocked membranes were incubated in primary antibody solutions at 4 °C overnight. The primary antibody solutions were prepared with 5% BSA in 1×TBST. The dilution factors are as follows: rabbit anti-SERF2 (Proteintech, 11691-1-AP) 1:1000 and mouse anti-GAPDH (Invitrogen, MA5-15738) 1:1000. After overnight incubation, membranes were washed three times with 1×TBST. Each wash was 5 minutes at room temperature. The washed membranes were then incubated with secondary antibody solutions at room temperature for 1 hour. IRDye™ 800CW goat anti-rabbit (Licor, 925-32211) and IRDye™ 680RD goat anti-mouse (Licor, 926-68070) were used as the secondary

antibodies. The membranes were washed three times and imaged with ChemiDoc™ Touch imaging system for chemiluminescence or Li-COR Odyssey DLx™ scanner.

Oligonucleotide synthesis

High-performance liquid chromatography (HPLC) purified unlabeled and 6-FAM fluorescent labeled RNA nucleotides were either purchased from Integrated DNA Technology (IDT) or synthesized at Slovenia NMR Center using previously described methods.[90,91] Homo-ribopolynucleotides (polyA, polyU, polyC, and polyG) were purchased from Sigma. All oligonucleotides were suspended in nuclease-free water or buffer prepared using nuclease-free water. All oligonucleotides were desalted using a 3 kD cutoff filter (Amicon ®Ultra 0.5 mL) by 10-time buffer exchange and their concentration was measured using the extinction coefficients obtained using the IDT OligoAnalyzer™ tool. G4 sequences were folded by cooling samples prepared in KCl (20mM sodium phosphate (NaPi), pH7.4 and 100 mM KCl) or LiCl buffers (20 mM Tris-HCl, pH7.4 and 100 mM LiCl) using a thermocycler with 1°C/min. Folded G4 quadruplexes were stored at 4 °C for immediate use or stored at -20°C for future use. All chemicals and reagents used in this study were commercially purchased with >98% purity and used without further purification.

Purification of recombinant proteins

The plasmids containing codon-optimized human SERF2 gene were synthesized commercially by GenScript, and subcloned to a pET28a-SUMO vector as reported elsewhere.[35] The expression and purification of SERF2 protein was like that for yeast SERF as previously we reported.[35] Briefly, the plasmids were transformed into competent BL21(DE3) *Escherichia coli* cells, incubated overnight in 10 mL of cell culture medium, and then transferred to freshly prepared 1 L of PMEM medium containing 50 mg/L Kanamycin. The cells were grown at 37°C, under shaking, until the OD_{600} reached 1.0. They were then transferred to a 20°C shaker for 1 hour and then protein expression was induced by adding 0.1 mM IPTG and cells were incubated overnight. Isotope labeled ^{15}N and $^{15}N/^{13}C$ SERF2 proteins for NMR studies, were produced by growing cells in M-9 minimal media supplemented with 100% ^{15}N NH_4Cl for the ^{15}N labelling or ^{15}N NH_4Cl and D-Glucose-$13C_6$ for the $^{15}N/^{13}C$ labeling. Cells were subsequently harvested and lysed by sonication in ice-cold lysis buffer (40 mM Tris-HCl pH 8.0, 10 mM NaPi, 400 mM NaCl, 20 mM imidazole, 10% glycerol, 1 tablet of cOmplete protease inhibitor (Sigma, 5056489001), and 1.25 µg/mL DNase I (Sigma, 10104159001)). The cell lysate was centrifuged at 36,000 g for 30 minutes and the supernatant was passed through a HisTrap

column (Cytiva, 17-5248-02). Lysis buffer containing 0.5 M imidazole was used to elute the His-SUMO tagged SERF2 proteins. The elution was then supplemented with 5 mM beta-mercaptoethanol (final concentration) and incubated overnight at 4°C with 10 µL of homemade SUMO protease 6His-ULP1. The SUMO cleaved digestion mixture was dialyzed in 40 mM Tris-HCl pH8.0 and 300 mM NaCl overnight at 4°C, using a 3.5 kD cutoff dialysis membrane (Repligen, 132724). The dialyzed proteins were run through a 5 mL HisTrap column to remove the cleaved His-ULP1 and His-SUMO. The flow-through SERF2 protein was further purified by an ion exchange HiTrap SP column (Cytiva, 17-5161-01) using buffer A (50 mM sodium phosphate and 125 mM NaCl pH6.0) and buffer B (50 mM sodium phosphate and 1 M NaCl, pH 6.0). A final purification of the ion-exchange purified SERF2 protein was conducted using a size-exclusion chromatography column Hiload75 (Cytiva, 28989333) in 20 mM NaPi pH7.5 and 150 mM NaCl or 40 mM HEPES pH 7.5 and 100 mM NaCl. The protein samples used in the biophysical and biochemical studies were prepared in indicated buffers as needed via buffer exchange using a 3kD cutoff filter (EMD Millipore, UFC503024). The SERF2 protein concentration was determined using the Pierce TM BCA assay calibrated with SERF2 A51W mutant serving as a standard. A similar expression and purification protocol, as described above, was used to produce the SERF2 (T2C) mutant, His-tagged and GST-tagged SERF2 for the high throughput screening assay.

Protein labeling

Cy-5 labeling was done by incubating 200 µM of SERF2 (T2C) with a 10-molar excess of Cy5 maleimide mono-reactive dye (Cytiva, PA25031) in 20 mM Tris-HCl pH7.4 and 100 mM KCl buffer overnight, at 25 °C, under continuous shaking at 300 rpm. The free excess label was removed by passing the sample through a PD-10 desalting column in a dark room. The samples were concentrated using a 3 kD Amicon Ultra-15 Centrifugal Filter Unit (EMD Millipore, UFC800324) and any residual free dyes were removed by resuspending the proteins in a working buffer after centrifugation at 8,500 rpm (8 times) for 15 minutes using an Amicon ®Ultra 0.5 mL 3 kD cutoff filter.

High-throughput screening assays

RNA bind-n-seq assay and analysis

For the RNA bind-n-seq assay (RBNS), a single-strand DNA library containing a randomized 40 nt region was obtained from IDT, gel-purified and the RNA library was prepared following a previously described method[41] using a T7 promoter and *in vitro* transcription. A

second pool of RNA pool was made by replacing guanines with 7dG (Trilink) to eliminate RNA G4 quadruplex folding while preserving the sequence. Residual DNA in the RNA pool was removed with DNAseI (Promega) followed by a phenol-chloroform extraction. RNA was resolved in a 6% TBE-Urea gel, the band of the expected size was excised, and gel-purified as previously described.[41]

RNA bind-n-seq method was modified from a previous study.[41] Briefly, 60 µL of recombinant GST-SBP-SERF2 at different concentrations (250 nM and 50 nM) in binding buffer (25 mM tris-HCl pH 7.5, 150 mM KCl or LiCl, 3 mM $MgCl_2$, 500ug/mL Ultrapure BSA and SUPERase-In RNase Inhibitor) was equilibrated with 60 µL of pre-washed (binding buffer) magnetic beads (Dynabeads MyOne Streptavidin T1, Invitrogen) for 30 minutes at 4°C. RNA pools were heated in the presence of 150 mM KCl or LiCl for 5 minutes at 100°C and allowed to cool down to room temperature for at least 10 minutes. 60 µL of 3 RNA pools (KCl, LiCl and 7dG (in LiCl)) were then mixed with GST-SBP-SERF2 and further incubated for 1 hour at 4°C. Final concentration of protein in the binding reaction was 250nM and 50nM and RNAs was 1uM final. Protein-RNA complexes were washed with wash buffer (25 mM tris-HCl pH 7.5, 150 mM KCl or LiCl and SUPERase-In RNase Inhibitor (Invitrogen)). Protein-RNA complexes were magnetically isolated and RNA eluted with elution buffer (4 mM biotin and 25 mM tris-HCl pH 7.5) for 30 minutes at 37°C. Elution was performed twice, eluates combined, and RNA was purified by phenol chloroform method.[41]Before reverse transcription, RNAs were heated for 5 minutes at 100°C in the presence of 150 mM LiCl (to facilitate reverse transcription through G-quadruplexes). Following reverse transcription, samples were prepared for sequencing as described elsewhere.[41]

We performed sequence enrichment analysis as previously described.[41,94] Briefly, we analyzed k-mer (*k*=6) enrichments to determine 'R' values defined as the frequency of given k-mer in the protein-associated pool divided by the frequency of that *k*mer in the input pool. To specifically determine enrichment of G quadruplexes, we searched for strong G4 patterns $(G_{(3-6)}N_{(0-7)})_4$. As a control we removed all sequences that matched the G4 pattern but still had greater than 8 Gs in the randomized region. G4 pattern analysis was performed on the randomized region plus the adapters (as described in Reference[41] for RNA structure analysis) as they are part of the RNAs presented to the protein in the binding reaction. The enrichment of each pattern was the frequency of the pattern in the protein-bound sample divided by frequency of the input pool.

FOREST assay

Library-1 from the previously published paper[1] that include 1800 pre-miRNA and 10 RNA G4 quadruplex sequences was used for FOREST screening. A detailed method for oligonucleotide template pool, DNA barcode microarray design, in vitro transcription, RNA fluorophore labeling, hybridization, and microarray scanning is provided in the supporting information. Briefly, the templates used here were synthesized by oligonucleotide library synthesis and the template size was limited to 170-nucleotides (OLS, Agilent Technologies). The in vitro transcribed RNA structure library was labeled with Cy5 at the 3' end to detect and quantify RNA probes on a microarray. The RNA structure library was next prepared in K^+ folding buffer (10 mM Tris-HCl pH 7.5, 100 mM KCl), heated at 95 °C and cooled to 4 °C at a rate of −6 °C/s on a ProFlex Thermal Cycler (Thermo Fisher Scientific) to allow G4 folding.

His-tagged SERF2 recombinant protein was used for the FOREST binding assay. For this purpose, a target protein (100 pmol of SERF2), 20 µL of TALON magnetic beads (Clontech) and 1 µg of the refolded RNA structure library were mixed in 1 mL of protein-binding buffer (10 mM Tris-HCl pH 7.5, 100 mM KCl, 10% glycerol, and 0.1 µg/µL BSA). A mixture containing no protein was also prepared as a control. The mixture was incubated on a rotator at 4 °C for 30 min and washed three times with the protein-binding buffer. Then, 200 µL of elution buffer was added to the magnetic beads, and the mixture was heated at 95 °C for 3-min. The RNA was collected from the supernatant by removing the magnetic beads. The RNA structure library in the mixture was extracted with phenol and chloroform with ethanol precipitation for purification. The enriched RNA sample was hybridized for microarray analysis as detailed in the supporting methods. To determine the protein-binding intensities of each RNA probe, we subtracted the fluorescence intensities of the negative control sample (samples without protein) from those of the enriched protein samples. To account for any undesired interactions with the barcode region, we calculate the average fluorescence intensity of each RNA structure by averaging the intensities of the RNA probes that have the same RNA structure but different RNA barcodes.

NMR experiments

NMR samples were prepared in NaPi buffer containing 92%/8% H_2O/D_2O. Isotopically unlabeled protein and oligonucleotide samples were used to collect 1D proton NMR spectrum with a recycle delay of 1 second. Multidimensional NMR data were collected using uniformly isotope-labeled 15N or 15N/13C SERF2 samples. A series of 2D and 3D NMR experiments,

that included 15N-HSQC,13C-HSQC, HNCO, HNCA, HNCOCA, HNCACO, CBCACONH, HNCACB, 15N-HSQC-NOESY, and 15NHSQC-TOCSY, were conducted for backbone assignments of 1 mM proteins dissolved in 20 mM d_3-NaAc (sodium acetate) (pH 5.5) and 8% D_2O at 4 °C. Unless indicated, all NMR measurements were done for protein or protein-RNA mixture samples suspended in 20 mM NaPi and 100 mM KCl (pH 7.4) containing 8% D_2O at 4 °C. 2D 15N/1H TROSY experiments were recorded for 100 μM 15N labeled SERF2 titrated with an increasing concentration of rG4 for binding site analysis. The effect of pH and temperature on SERF2 chemical exchange was studied by acquiring the 2D spectrum at 4 and 37 °C in buffers containing either 20 mM NaPi and 100 mM KCl (pH7.4) or 20 mM d_3-NaAc (pH 5.5). 15N relaxation NMR measurements, that included heteroNOE, T1 and T2, were done using 500 μM 15N SERF2 mixed with and without 250 μM of TERRA. The relaxation delays used for the T1 experiments were 20, 50, 90, 130, 200, 320, 450, 580, 750, 900, 1200, and 2005 ms and T2 relaxation delays used were 16.96, 50.88, 84.80, 135.68, 169.60, 220.48, 305.28, 373.12, 457.92, 542.72, 678.40, and 1356.80 ms. Saturation transfer difference (STD) NMR spectra were recorded with 512 scans, 4 s saturation time, and on-resonance excitation at different chemical shift regions, and off-resonance excitation at -40.0 ppm. 2D TROSY, 3D, and STD NMR data were collected on a Bruker 800 MHz spectrometer equipped with a triple resonance cryoprobe. 15N relaxation data were collected on a Bruker 600 MHz equipped with a triple resonance cryoprobe. The NMR data was processed using Bruker's Topspin 4.1.4 and spectra assignment and analysis was done using an NMRFAM-Sparky 1.47.

Fluorescence polarization and anisotropy assay

Fluorescently labeled 6-FAM RNA probes were prepared in nuclease-free water containing 20 mM NaPi (pH 7.4) and 100 mM KCl. Fluorescence polarization assays were done using 20 nM RNA 6-FAM RNA probe mixed with increasing concentrations of SERF2, ranging from 0.009 μM to 20 μM which were dissolved in 20 mM NaPi and 100 mM KCl (pH7.4). The sample mixture was incubated for 30 minutes at room temperature. Fluorescence polarization data were recorded on a TECAN Infinite M1000 microplate reader at 25 °C with the excitation and emission wavelengths set at 470 and 530 nm, respectively. Fluorescence anisotropy measurements were done by titrating SERF2 to 200 nM of Cy3 (excitation/emission, 550/570 nm) or FAM (excitation/emission, 493/517 nm) labelled oligonucleotides in a 1 mL quartz cuvette (Hellma,101-QS) using Cary Eclipse spectrofluorometer (Agilent) at 25 °C. Slit bandwidths were set at 5 nm and 10 nm for excitation and emission, respectively. The binding

constant (K_D) was calculated from the change in polarization or anisotropy values in GraphPad Prism 9.5.1 using non-linear regression for curve fitting with a one-site specific binding model.

Circular dichroism spectroscopy

The secondary structure of G4 quadruplexes (15 μM) and SERF2 (50 μM) suspended in 20 mM NaPi and 100 mM KCl (pH7.4) was studied by collecting CD spectra using a JASCO J-1500 spectropolarimeter. For folding analysis, the CD spectra of SERF2 were collected at different temperatures (4, 25 and 37 °C). The buffer CD spectrum was subtracted from the average CD spectrum obtained from 8 scans.

Liquid-liquid phase transition assay

16-well Culture-Well chamber slips (Grace Bio-Labs) or 384-well plates (Cellvis, P384-1.5H-N) pre-treated with 5% (w/v) Pluronic™ F-127 (Sigma, P2443) overnight were used to study in vitro phase separation. Well chambers were washed three times with NaPi buffer and air dried. 50 or 20 μl of reaction sample mixtures were incubated for 30 minutes at room temperature in various conditions (varying protein, total RNA and rG4 quadruplex concentration) in a buffer with or without 10% of PEG8000 (Sigma, P5413). For fluorescence imaging, 1/200th fluorescence labelled (6-FAM or Cy-5) protein/rG4 samples are mixed to unlabeled sample mixtures Sample imaging and fluorescence recovery after photobleaching measurements were done using a Nikon Ti2-E motorized, inverted microscope. This microscope is controlled by NIS Elements software containing a SOLA 365 LED light source and used a 100X oil immersion objective. Recovery half-life analysis was done using GraphPad Prism and image processing was done using Fiji ImageJ.

Size-distribution analysis

Gel shift assay

20 μM of TERRA quadruplexes dissolved in NaPi buffer were mixed with increasing concentration of SERF2 (0.5 to 2x molar excess) and incubated for 30 minutes at room temperature. Gel shift mobility assay was performed by loading 5 μM TERRA sample mixture containing SERF2 and 20% glycerol to a 4-20% TBE gel (Invitrogen, EC6225).

Size-exclusion chromatography

20 μM of TERRA quadruplexes dissolved in NaPi buffer was mixed with 40 μM of SERF2 and incubated for 30 minutes at room temperature. The sample mixture was next injected to a Superdex 200 Increase 10/300 GL size-exclusion chromatography column (Cytiva, 28-9909-44).

Analytical Ultra Centrifugation (AUC)

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

Structure calculation and MD simulation

SERF2 backbone and nOE assignments were done using NMRFAM-Sparky and the dihedral angles were predicted utilizing the backbone chemical shifts by TALOS-N program.[60] A total of 835 NOE distance constraints were used for the multiple-state ensemble calculation in CYANA 3.98.15. One hundred conformers were calculated using 10,000 torsion-angle dynamics steps. The 20 conformers with the lowest target function values were used to represent the calculated SERF2 structure. To gauge the intrinsic structural dynamics of *SERF2-TERRA* systems, we performed all-atom MD simulations, that were defined in a structure-based balanced forcefield, i.e., CHARMM36m using GROMACS 2022.4.[92] The SERF2-TERRA complex structure was built using the HADDOCK program[93] by parsing the NMR distance restraints obtained from chemical shift perturbations and saturation transfer NMR data analysis. A set of ambiguous active site SERF2 residues (T3, N5, R7, R11, Q12, K16, S19, S21, K23, A33, Q46, K47, A51, N52, K55 and E56) and TERRA guanines (G3, G5 and G9) were provided to the HADDOCK program to allow it to build the structure of the complex. The energetically best cluster complex structure was used for MD simulation analysis by solvating in TIP3P water model in a orthorhombic water boxes. These model systems were electro-neutralized using 100 mM KCl and 20 mM NaPi, pH 7.4 followed by energy minimization using the steepest descent algorithms in less than 5000 steps to remove steric clashes. The energy-minimized systems were then subjected to two-step equilibrations using NVT and NPT ensemble for 50 ns by keeping the temperature to 37 °C and pressure to 1 bar using the V-rescale thermostat and the Berendsen barostat. Equilibrated systems are next subjected to a production MD of 0.5 μs. All the bonds involving hydrogen atoms were restrained using the SHAKE algorithm. The atomic coordinates of each system were saved at every 100 ps resulting 5000 snapshots for each system for post dynamics analysis.

REFERENCES

1. Lipps, H.J., and Rhodes, D. (2009). G-quadruplex structures: in vivo evidence and function. Trends Cell Biol *19*, 414-422 10.1016/j.tcb.2009.05.002.
2. Summers, P.A., Lewis, B.W., Gonzalez-Garcia, J., Porreca, R.M., M Lim, A.H., Cadinu, P., Martin-Pintado, N., Mann, D.J., Edel, J.B., Baptiste Vannier, J., et al. Visualising G-quadruplex DNA dynamics in live cells by fluorescence lifetime imaging microscopy. Nat Commun *12*, 162. 10.1038/s41467-020-20414-7.
3. Yang, X., Cheema, J., Zhang, Y., Deng, H., Duncan, S., Umar, M.I., Zhao, J., Liu, Q., Cao, X., Kwok, C.K., et al. (2020). RNA G-quadruplex structures exist and function in vivo in plants. Genome Biol *21*, 226. 10.1186/s13059-020-02142-9.
4. Kharel, P., Fay, M., Manasova, E. V., Anderson, P.J., Kurkin, A. V., Guo, J.U., and Ivanov, P. (2023). Stress promotes RNA G-quadruplex folding in human cells. Nat Commun *14*, 205. 10.1038/s41467-023-35811-x.
5. Byrd, A.K., Zybailov, B.L., Maddukuri, L., Gao, J., Marecki, J.C., Jaiswal, M., Bell, M.R., Griffin, W.C., Reed, M.R., Chib, S., et al. (2016). Evidence that G-quadruplex DNA accumulates in the cytoplasm and participates in stress granule assembly in response to oxidative stress. Journal of Biological Chemistry *291*, 18041–18057. 10.1074/JBC.M116.718478.
6. Yang, X., Yu, H., Duncan, S., Zhang, Y., Cheema, J., Liu, H., Benjamin Miller, J., Zhang, J., Kwok, C.K., Zhang, H., et al. (2022). RNA G-quadruplex structure contributes to cold adaptation in plants. Nat Commun *13*, 6224. 10.1038/s41467-022-34040-y.
7. Fay, M.M., Lyons, S.M., and Ivanov, P. (2017). RNA G-Quadruplexes in Biology: Principles and Molecular Mechanisms. J Mol Biol *429*, 2127-2147. 10.1016/j.jmb.2017.05.017.
8. Spiegel, J., Adhikari, S., and Balasubramanian, S. (2020). The Structure and Function of DNA G-Quadruplexes. Trends Chem *2*, 123–136. 10.1016/j.trechm.2019.07.002.
9. Lane, A.N., Chaires, J.B., Gray, R.D., and Trent, J.O. (2008). Stability and kinetics of G-quadruplex structures. Nucleic Acids Res *36*, 5482-5515. 10.1093/nar/gkn517.
10. Chen, X.C., Chen, S. bin, Dai, J., Yuan, J.H., Ou, T.M., Huang, Z.S., and Tan, J.H. (2018). Tracking the Dynamic Folding and Unfolding of RNA G-Quadruplexes in Live Cells. Angewandte Chemie - International Edition *57*, 4702–4706. 10.1002/anie.201801999.
11. Laguerre, A., Hukezalie, K., Winckler, P., Katranji, F., Chanteloup, G., Pirrotta, M., Perrier-Cornet, J.M., Wong, J.M.Y., and Monchaud, D. (2015). Visualization of RNA-Quadruplexes in Live Cells. J Am Chem Soc *137*, 8521–8525. 10.1021/jacs.5b03413.
12. Biffi, G., Di Antonio, M., Tannahill, D., and Balasubramanian, S. (2013). Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells. Nature Chemistry *6*, 75–80. 10.1038/nchem.1805.
13. Varshney, D., Spiegel, J., Zyner, K., Tannahill, D., and Balasubramanian, S. (2020). The regulation and functions of DNA and RNA G-quadruplexes. Nat Rev Mol Cell Biol *21*, 459–474. 10.1038/s41580-020-0236-x.
14. Lyu, K., Chow, E.Y.C., Mou, X., Chan, T.F., and Kwok, C.K. (2021). RNA G-quadruplexes (rG4s): Genomics and biological functions. Nucleic Acids Res *49*, 5426–5450. 10.1093/nar/gkab187.
15. Dumas, L., Herviou, P., Dassi, E., Cammas, A., and Millevoi, S. (2021). G-Quadruplexes in RNA Biology: Recent Advances and Future Directions. Trends Biochem Sci *46*, 270-283. 10.1016/j.tibs.2020.11.001.
16. Kharel, P., Becker, G., Tsvetkov, V., and Ivanov, P. (2020). Properties and biological impact of RNA G-quadruplexes: From order to turmoil and back. Nucleic Acids Res *48*, 12534–12555. 10.1093/nar/gkaa1126.

17. Simone, R., Fratta, P., Neidle, S., Parkinson, G.N., and Isaacs, A.M. (2015). G-quadruplexes: Emerging roles in neurodegenerative diseases and the non-coding transcriptome. FEBS Lett *589*, 1653-1668. 10.1016/j.febslet.2015.05.003.
18. Cammas, A., and Millevoi, S. (2017). RNA G-quadruplexes: emerging mechanisms in disease. Nucleic Acids Res *45*, 1584-1595. 10.1093/nar/gkw1280.
19. Fay, M.M., Anderson, P.J., and Ivanov, P. (2017). ALS/FTD-Associated C9ORF72 Repeat RNA Promotes Phase Transitions In Vitro and in Cells. Cell Rep *21*, 3573–3584. 10.1016/j.celrep.2017.11.093.
20. Asamitsu, S., Yabuki, Y., Matsuo, K., Kawasaki, M., Hirose, Y., Kashiwazaki, G., Chandran, A., Bando, T., Wang, D.O., Sugiyama, H., et al. (2023). RNA G-quadruplex organizes stress granule assembly through DNAPTP6 in neurons. Sci Adv *9*, eade2035. 10.1126/sciadv.ade2035.
21. Santos, T., Salgado, G.F., Cabrita, E.J., and Cruz, C. (2022). Nucleolin: a binding partner of G-quadruplex structures. Trends Cell Biol *32*, 561-564. 10.1016/j.tcb.2022.03.003.
22. Tippana, R., Chen, M.C., Demeshkina, N.A., Ferré-D'Amaré, A.R., and Myong, S. (2019). RNA G-quadruplex is resolved by repetitive and ATP-dependent mechanism of DHX36. Nat Commun *10*, 1855. 10.1038/s41467-019-09802-w.
23. Sauer, M., Juranek, S.A., Marks, J., De Magis, A., Kazemier, H.G., Hilbig, D., Benhalevy, D., Wang, X., Hafner, M., and Paeschke, K. (2019). DHX36 prevents the accumulation of translationally inactive mRNAs with G4-structures in untranslated regions. Nat Commun *10*, 2421. 10.1038/s41467-019-10432-5.
24. Ishiguro, A., Lu, J., Ozawa, D., Nagai, Y., and Ishihama, A. (2021). ALS-linked FUS mutations dysregulate G-quadruplex-dependent liquid–liquid phase separation and liquid-to-solid transition. Journal of Biological Chemistry *297*, 101284. 10.1016/j.jbc.2021.101284.
25. Flynn, R.L., Centore, R.C., O'Sullivan, R.J., Rai, R., Tse, A., Songyang, Z., Chang, S., Karlseder, J., and Zou, L. (2011). TERRA and hnRNPA1 orchestrate an RPA-to-POT1 switch on telomeric single-stranded DNA. Nature *471*, 532–538. 10.1038/nature09772.
26. Biffi, G., Tannahill, D., and Balasubramanian, S. (2012). An intramolecular G-quadruplex structure is required for binding of telomeric repeat-containing RNA to the telomeric protein TRF2. J Am Chem Soc *134*, 11974–11976. 10.1021/ja305734x.
27. Mendoza, O., Bourdoncle, A., Boulé, J.B., Brosh, R.M., and Mergny, J.L. (2016). G-quadruplexes and helicases. Nucleic Acids Res *44*, 1989–2006. 10.1093/nar/gkw079.
28. Tosoni, E., Frasson, I., Scalabrin, M., Perrone, R., Butovskaya, E., Nadai, M., Palù, G., Fabris, D., and Richter, S.N. (2015). Nucleolin stabilizes G-quadruplex structures folded by the LTR promoter and silences HIV-1 viral transcription. Nucleic Acids Res *43*, 8884-8897. 10.1093/nar/gkv897.
29. Buchan, J.R., and Parker, R. (2009). Eukaryotic stress granules: the ins and outs of translation. Mol Cell *36*, 932-941. 10.1016/j.molcel.2009.11.020.
30. Hofmann, S., Kedersha, N., Anderson, P., and Ivanov, P. (2021). Molecular mechanisms of stress granule assembly and disassembly. Biochim Biophys Acta Mol Cell Res *1868*, 118876. 10.1016/j.bbamcr.2020.118876.
31. Alberti, S., Gladfelter, A., and Mittag, T. (2019). Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. Cell *176*, 419-434. 10.1016/j.cell.2018.12.035.
32. Galvanetto, N., Ivanović, M.T., Chowdhury, A., Sottini, A., Nüesch, M.F., Nettels, D., Best, R.B., and Schuler, B. (2023). Extreme dynamics in a biomolecular condensate. Nature *619*, 876-883. 10.1038/s41586-023-06329-5.
33. van Ham, T.J., Holmberg, M.A., van der Goot, A.T., Teuling, E., Garcia-Arencibia, M., Kim, H. eui, Du, D., Thijssen, K.L., Wiersma, M., Burggraaff, R., et al. (2010). Identification of MOAG-4/SERF as a regulator of age-related proteotoxicity. Cell *142*, 601–612. 10.1016/j.cell.2010.07.020.

34. Falsone, S.F., Meyer, N.H., Schrank, E., Leitinger, G., Pham, C.L.L., Fodero-Tavoletti, M.T., Holmberg, M., Dulle, M., Scicluna, B., Gesslbauer, B., et al. (2012). SERF Protein Is a Direct Modifier of Amyloid Fiber Assembly. Cell Rep *2*, 358–371. 10.1016/j.celrep.2012.06.012.
35. Meinen, B.A., Gadkari, V. v, Stull, F., Ruotolo, B.T., and Bardwell, J.C.A. (2019). SERF engages in a fuzzy complex that accelerates primary nucleation of amyloid proteins. *116*, 23040–23049. 10.6084/m9.figshare.c.4684733.v1.
36. Sahoo, B.R., and Bardwell, J.C.A. (2022). SERF, a family of tiny highly conserved, highly charged proteins with enigmatic functions. FEBS Journal *290*, 4150-4162. 10.1111/febs.16555.
37. Meyer, N.H., Dellago, H., Tam-Amersdorfer, C., Merle, D.A., Parlato, R., Gesslbauer, B., Almer, J., Gschwandtner, M., Leon, A., Franzmann, T.M., et al. (2020). Structural Fuzziness of the RNA-Organizing Protein SERF Determines a Toxic Gain-of-interaction. J Mol Biol *432*, 930–951. 10.1016/j.jmb.2019.11.014.
38. Bernstam, L., and Nriagu, J. (2000). Molecular aspects of arsenic stress. J Toxicol Environ Health B Crit Rev *3*, 293-322. 10.1080/109374000436355.
39. Palumbo, S.L., Memmott, R.M., Uribe, D.J., Krotova-Khan, Y., Hurley, L.H., and Ebbinghaus, S.W. (2008). A novel G-quadruplex-forming GGA repeat region in the c-myb promoter is a critical regulator of promoter activity. Nucleic Acids Res *36*, 1755-1769. 10.1093/nar/gkm1069.
40. Turner, M., Danino, Y.M., Barshai, M., Yacovzada, N.S., Cohen, Y., Olender, T., Rotkopf, R., Monchaud, D., Hornstein, E., and Orenstein, Y. (2022). rG4detector, a novel RNA G-quadruplex predictor, uncovers their impact on stress granule formation. Nucleic Acids Res *50*, 11426-11441. 10.1093/nar/gkac950.
41. Dominguez, D., Freese, P., Alexis, M.S., Su, A., Hochman, M., Palden, T., Bazile, C., Lambert, N.J., Van Nostrand, E.L., Pratt, G.A., et al. (2018). Sequence, Structure, and Context Preferences of Human RNA Binding Proteins. Mol Cell *70*, 854-867. 10.1016/J.MOLCEL.2018.05.001.
42. Komatsu, K.R., Taya, T., Matsumoto, S., Miyashita, E., Kashida, S., and Saito, H. (2020). RNA structure-wide discovery of functional interactions with multiplexed RNA motif library. Nat Commun *11*, 6275. 10.1038/s41467-020-19699-5.
43. You, J., Li, H., Lu, X.M., Li, W., Wang, P.Y., Dou, S.X., and Xi, X.G. (2017). Effects of monovalent cations on folding kinetics of G-quadruplexes. Biosci Rep *37*, BSR20170771. 10.1042/BSR20170771.
44. Han, H., Hurley, L.H., and Salazar, M. (1999). A DNA polymerase stop assay for G-quadruplex-interactive compounds. Nucleic Acids Res *27*, 537-542. 10.1093/nar/27.2.537.
45. Kikin, O., D'Antonio, L., and Bagga, P.S. (2006). QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. Nucleic Acids Res *34*, W676–W682. 10.1093/NAR/GKL253.
46. Cheong, C., and Moore, P.B. (1992). Solution Structure of an Unusually Stable RNA Tetraplex Containing G-and U-Quartet Structures. Biochemistry *31*, 8406-8414. 10.1021/bi00151a003.
47. Bettin, N., Oss Pegorar, C., and Cusanelli, E. (2019). The Emerging Roles of TERRA in Telomere Maintenance and Genome Stability. Cells *8*, 246. 10.3390/cells8030246.
48. Zhang, K., Donnelly, C.J., Haeusler, A.R., Grima, J.C., Machamer, J.B., Steinwald, P., Daley, E.L., Miller, S.J., Cunningham, K.M., Vidensky, S., et al. (2015). The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. Nature *525*, 56-61. 10.1038/nature14973.
49. Protter, D.S.W., and Parker, R. (2016). Principles and Properties of Stress Granules. Trends Cell Biol *26*, 668–679. 10.1016/J.TCB.2016.05.004.
50. Sanders, D.W., Kedersha, N., Lee, D.S.W., Strom, A.R., Drake, V., Riback, J.A., Bracha, D., Eeftens, J.M., Iwanicki, A., Wang, A., et al. (2020). Competing Protein-RNA Interaction Networks Control Multiphase Intracellular Organization. Cell *181*, 306-324.e28. 10.1016/j.cell.2020.03.050.
51. Feric, M., Vaidya, N., Harmon, T.S., Mitrea, D.M., Zhu, L., Richardson, T.M., Kriwacki, R.W., Pappu, R. V., and Brangwynne, C.P. (2016). Coexisting Liquid Phases Underlie Nucleolar Subcompartments. Cell *165*, 1686–1697. 10.1016/J.CELL.2016.04.047.

52. Wang, B., Zhang, L., Dai, T., Qin, Z., Lu, H., Zhang, L., and Zhou, F. (2021). Liquid–liquid phase separation in human health and diseases. Signal Transduct Target Ther *6*, 290. 10.1038/s41392-021-00678-1.
53. Zhang, Y., Yang, M., Duncan, S., Yang, X., Abdelhamid, M.A.S., Huang, L., Zhang, H., Benfey, P.N., Waller, Z.A.E., and Ding, Y. (2019). G-quadruplex structures trigger RNA phase separation. Nucleic Acids Res *47*, 11746–11754. 10.1093/NAR/GKZ978.
54. Xu, Y., Suzuki, Y., Ito, K., and Komiyama, M. (2010). Telomeric repeat-containing RNA structure in living cells. Proc Natl Acad Sci U S A *107*, 14579–14584. 10.1073/pnas.1001177107.
55. Maharana, S., Wang, J., Papadopoulos, D.K., Richter, D., Pozniakovsky, A., Poser, I., Bickle, M., Rizk, S., Guillén-Boixet, J., Franzmann, T.M., et al. (2018). RNA buffers the phase separation behavior of prion-like RNA binding proteins. Science *360*, 918-921. 10.1126/science.aar7366.
56. Posey, A.E., Holehouse, A.S., and Pappu, R. V. (2018). Phase Separation of Intrinsically Disordered Proteins. In Methods in Enzymology *611*, 1-30. 10.1016/bs.mie.2018.09.035.
57. Guillén-Boixet, J., Kopach, A., Holehouse, A.S., Wittmann, S., Jahnel, M., Schlüßler, R., Kim, K., Trussina, I.R.E.A., Wang, J., Mateju, D., et al. (2020). RNA-induced conformational switching and clustering of G3BP drive stress granule assembly by condensation. Cell *181*, 346-361.e17. 10.1016/j.cell.2020.03.049.
58. Yamaguchi, T., Matsuzaki, K., and Hoshino, M. (2011). Transient formation of intermediate conformational states of amyloid-β peptide revealed by heteronuclear magnetic resonance spectroscopy. FEBS Lett *585*, 1097–1102. 10.1016/j.febslet.2011.03.014.
59. Whitehead, R.D., Teschke, C.M., and Alexandrescu, A.T. (2019). NMR Mapping of Disordered Segments from a Viral Scaffolding Protein Enclosed in a 23 MDa Procapsid. Biophys J *117*, 1387-1392. 10.1016/j.bpj.2019.08.038.
60. Shen, Y., and Bax, A. (2015). Protein structural information derived from NMR chemical shift with the neural network program TALOS-N. Methods in Molecular Biology *1260*, 17-32. 10.1007/978-1-4939-2239-0_2.
61. Zeke, A., Schád, É., Horváth, T., Abukhairan, R., Szabó, B., and Tantos, A. (2022). Deep structural insights into RNA-binding disordered protein regions. Wiley Interdiscip Rev RNA *13*, e1714. 10.1002/wrna.1714.
62. Dignon, G.L., Best, R.B., and Mittal, J. (2020). Biomolecular phase separation: From molecular driving forces to macroscopic properties. Annu Rev Phys Chem *71*, 53-75. 10.1146/annurev-physchem-071819-113553.
63. Martadinata, H., and Phan, A.T. (2013). Structure of human telomeric RNA (TERRA): Stacking of two G-quadruplex blocks in K+ solution. Biochemistry *52*, 2176-2183. 10.1021/bi301606u.
64. Ivanov, P., Kedersha, N., and Anderson, P. (2019). Stress granules and processing bodies in translational control. Cold Spring Harb Perspect Biol *11*, a032813. 10.1101/cshperspect.a032813.
65. Khong, A., Matheny, T., Jain, S., Mitchell, S.F., Wheeler, J.R., and Parker, R. (2017). The stress granule transcriptome reveals principles of mrna accumulation in stress granules. Mol Cell *68*, 808-820.e5. 10.1016/j.molcel.2017.10.015.
66. Namkoong, S., Ho, A., Woo, Y.M., Kwak, H., and Lee, J.H. (2018). Systematic Characterization of Stress-Induced RNA Granulation. Mol Cell *70*, 175-187.e8. 10.1016/j.molcel.2018.02.025.
67. Jain, S., Wheeler, J.R., Walters, R.W., Agrawal, A., Barsic, A., and Parker, R. (2016). ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. Cell *164*, 487-498. 10.1016/j.cell.2015.12.038.
68. Youn, J.Y., Dunham, W.H., Hong, S.J., Knight, J.D.R., Bashkurov, M., Chen, G.I., Bagci, H., Rathod, B., MacLeod, G., Eng, S.W.M., et al. (2018). High-Density Proximity Mapping Reveals the Subcellular Organization of mRNA-Associated Granules and Bodies. Mol Cell *69*, 517-532.e11. 10.1016/j.molcel.2017.12.020.

69. Markmiller, S., Soltanieh, S., Server, K.L., Mak, R., Jin, W., Fang, M.Y., Luo, E.C., Krach, F., Yang, D., Sen, A., et al. (2018). Context-Dependent and Disease-Specific Diversity in Protein Interactions within Stress Granules. Cell *172*, 590-604.e13. 10.1016/j.cell.2017.12.032.
70. Huynh, M.L., Russell, P., and Walsh, B. (2009). Tryptic digestion of in-gel proteins for mass spectrometry analysis. Methods Mol Biol *519*, 507-513. 10.1007/978-1-59745-281-6_34.
71. Song, J., Perreault, J.-P., Topisirovic, I., and Richard, S. (2016). RNA G-quadruplexes and their potential regulatory roles in translation. Translation *4*, e1244031. 10.1080/21690731.2016.1244031.
72. Shao, X., Zhang, W., Umar, M.I., Wong, H.Y., Seng, Z., Xie, Y., Zhang, Y., Yang, L., Kwok, C.K., and Deng, X. (2020). RNA G-quadruplex structures mediate gene regulation in bacteria. mBio *11*, e02926-19. 10.1128/mBio.02926-19.
73. Zhang, R., Liu, Y., Zhang, X., Xiao, K., Hou, Y., Liu, H., and Sun, X. (2021). Detecting and profiling endogenous rna g-quadruplexes in the human transcriptome. Int J Mol Sci *22*, 8012. 10.3390/ijms22158012.
74. Bugaut, A., and Balasubramanian, S. (2012). 5'-UTR RNA G-quadruplexes: Translation regulation and targeting. Nucleic Acids Res *40*, 4727-4741. 10.1093/nar/gks068.
75. Colombo, J., Provazzi, P.J.S., Calmon, M.F., Pires, L.C., Rodrigues, N.C., Petl, P., Fossey, M.A., de Souza, F.P., Canduri, F., and Rahal, P. (2013). Expression, purification, and molecular analysis of the human ZNF706 protein. Biol Proced Online *15*, 10. 10.1186/1480-9222-15-10.
76. Iida, K., Suzuki, N., Sasaki, A., Ishida, S., and Arai, T. (2022). Development of a novel light-up probe for detection of G-quadruplexes in stress granules. Sci Rep *12*, 12892. 10.1038/s41598-022-17230-y.
77. Huang, Z.L., Dai, J., Luo, W.H., Wang, X.G., Tan, J.H., Chen, S. Bin, and Huang, Z.S. (2018). Identification of G-Quadruplex-Binding Protein from the Exploration of RGG Motif/G-Quadruplex Interactions. J Am Chem Soc *140*, 17945–17955. 10.1021/JACS.8B09329/ASSET/IMAGES/LARGE/JA-2018-09329J_0007.JPEG.
78. Yagi, R., Miyazaki, T., and Oyoshi, T. (2018). G-quadruplex binding ability of TLS/FUS depends on the β-spiral structure of the RGG domain. Nucleic Acids Res *46*, 5894-5901. 10.1093/nar/gky391.
79. Awadasseid, A., Ma, X., Wu, Y., and Zhang, W. (2021). G-quadruplex stabilization via small-molecules as a potential anti-cancer strategy. Biomedicine and Pharmacotherapy *139*, 111550. 10.1016/j.biopha.2021.111550.
80. Chen, M.C., tippana, ramreddy, Demeshkina, N.A., Murat, P., Balasubramanian, S., Myong, S., and Ferré-D, A. (2018). Structural basis of G-quadruplex unfolding by the DEAH/RHA helicase DHX36. *558*, 65-469. 10.1038/s41586-018-0209-9.
81. Kosiol, N., Juranek, S., Brossart, P., Heine, A., and Paeschke, K. (2021). G-quadruplexes: a promising target for cancer therapy. Molecular Cancer *20*, 40. 10.1186/S12943-021-01328-4.
82. Zhang, Y., Gaetano, C.M., Williams, K.R., Bassell, G.J., and Mihailescu, M.R. (2014). FMRP interacts with G-quadruplex structures in the 3 '-UTR of its dendritic target shank1 mRNA. RNA Biol *11*, 1364-1374. 10.1080/15476286.2014.996464.
83. Mitteaux, J., Lejault, P., Wojciechowski, F., Joubert, A., Boudon, J., Desbois, N., Gros, C.P., Hudson, R.H.E., Boulé, J.B., Granzhan, A., et al. (2021). Identifying G-Quadruplex-DNA-Disrupting Small Molecules. J Am Chem Soc *143*, 12567-12577. 10.1021/jacs.1c04426.
84. Glauninger, H., Wong Hickernell, C.J., Bard, J.A.M., and Drummond, D.A. (2022). Stressful steps: Progress and challenges in understanding stress-induced mRNA condensation and accumulation in stress granules. Mol Cell *82*, 2544-2556. 10.1016/j.molcel.2022.05.014.
85. Traczyk, A., Liew, C.W., Gill, D.J., and Rhodes, D. (2020). Structural basis of G-quadruplex DNA recognition by the yeast telomeric protein Rap1. Nucleic Acids Res *48*, 4562–4571. 10.1093/nar/gkaa171.
86. Babinchak, W.M., and Surewicz, W.K. (2020). Liquid–liquid phase separation and its mechanistic role in pathological protein aggregation. J Mol Biol *432*, 1910-1925. 10.1016/j.jmb.2020.03.004.

87. Marcelo, A., Koppenol, R., de Almeida, L.P., Matos, C.A., and Nóbrega, C. (2021). Stress granules, RNA-binding proteins and polyglutamine diseases: too much aggregation? Cell Death Dis *12*, 592. 10.1038/s41419-021-03873-8.
88. Wolozin, B., and Ivanov, P. (2019). Stress granules and neurodegeneration. Nat Rev Neurosci *20*, 649-666. 10.1038/s41583-019-0222-5.
89. Begeman, A., Son, A., Litberg, T.J., Wroblewski, T.H., Gehring, T., Cabral, V.H., Bourne, J., Xuan, Z., and Horowitz, S. (2020). G-Quadruplexes act as sequence-dependent protein chaperones. EMBO Rep *21*, e49735. 10.15252/EMBR.201949735.
90. Brčić, J., and Plavec, J. (2017). ALS and FTD linked GGGGCC-repeat containing DNA oligonucleotide folds into two distinct G-quadruplexes. Biochim Biophys Acta Gen Subj *1861*, 1237–1245. 10.1016/j.bbagen.2016.11.018.
91. Kocman, V., and Plavec, J. (2017). Tetrahelical structural family adopted by AGCGA-rich regulatory DNA regions. Nat Commun *8*, 15355. 10.1038/ncomms15355.
92. Abraham, M.J., Murtola, T., Schulz, R., Páll, S., Smith, J.C., Hess, B., and Lindah, E. (2015). Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX *1–2*, 19–25. 10.1016/j.softx.2015.06.001.
93. De Vries, S.J., Van Dijk, M., and Bonvin, A.M.J.J. (2010). The HADDOCK web server for data-driven biomolecular docking. Nat Protoc *5*, 883-897. 10.1038/nprot.2010.32.
94. Goering, R., Hudish, L.I., Guzman, B.B., Raj, N., Bassell, G.J., Russ, H.A., Dominguez, D., and Taliaferro, J.M. (2020). FMRP promotes RNA localization to neuronal projections through interactions between its RGG domain and G-quadruplex RNA sequences. Elife *9*, e52621. 10.7554/eLife.52621.

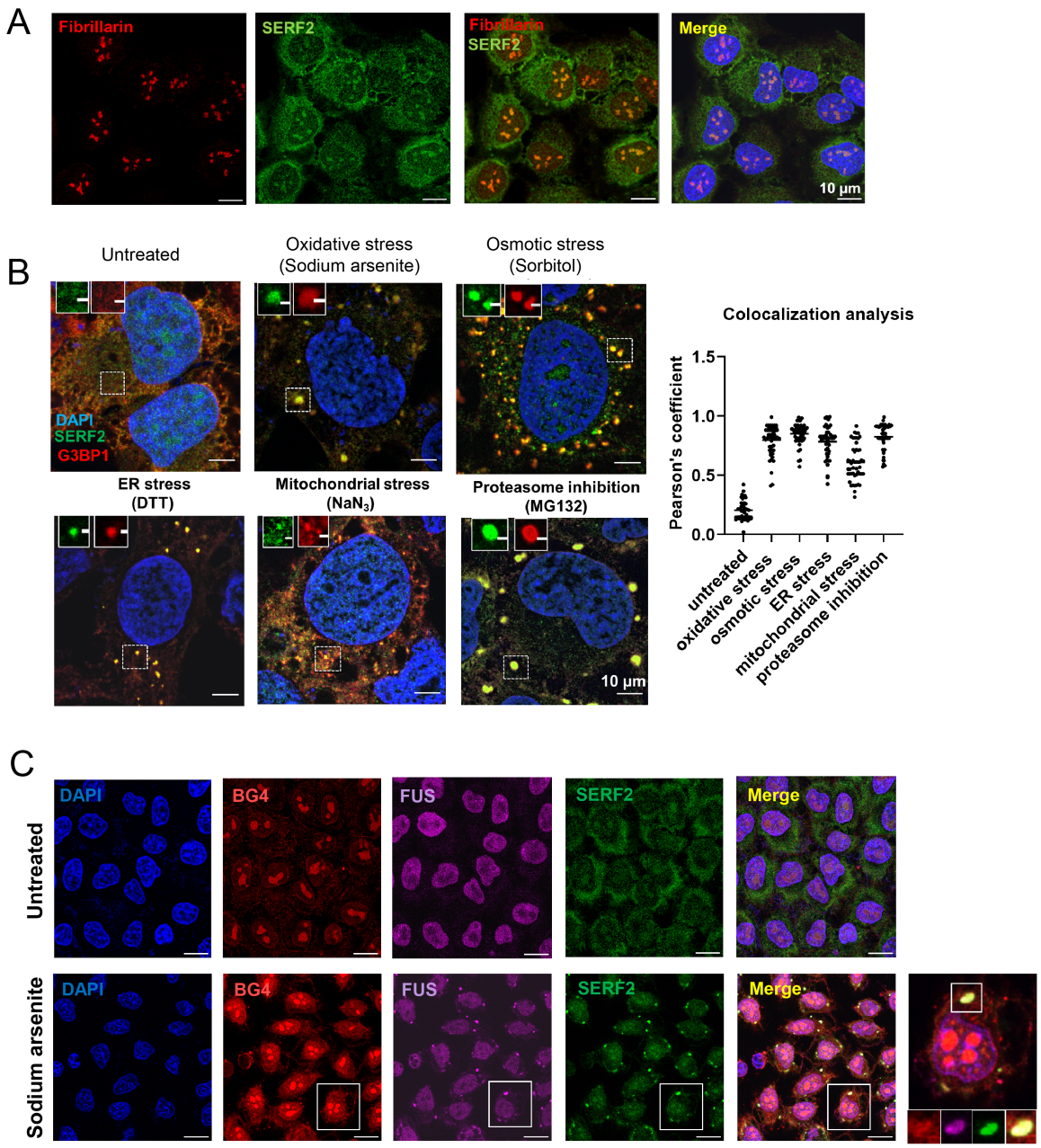

Figure 1: SERF2 colocalizes with stress granule upon various stress conditions. (**A**) Immunofluorescence images show endogenous SERF2 is prevalently distributed in the nucleolus in fixed U2OS cells as evidenced from fibrillarin staining. (**B**) SERF2 shows the formation of cytoplasmic foci that are co-localized with G3BP1 protein, a stress granule marker in different stress conditions suggesting its involvement in stress-granule formation. The plot on the right shows the quantification of stress granules retrieved from (B) under various stress conditions containing both SERF2 and G3BP1. **(C)** Immunofluorescence images of fixed U2OS cells showing oxidative stress induced stress granules contain SERF2, FUS and G4 quadruplexes as detected by specific antibodies indicated in green, purple, and red, respectively. Scale bar is 10 µm.

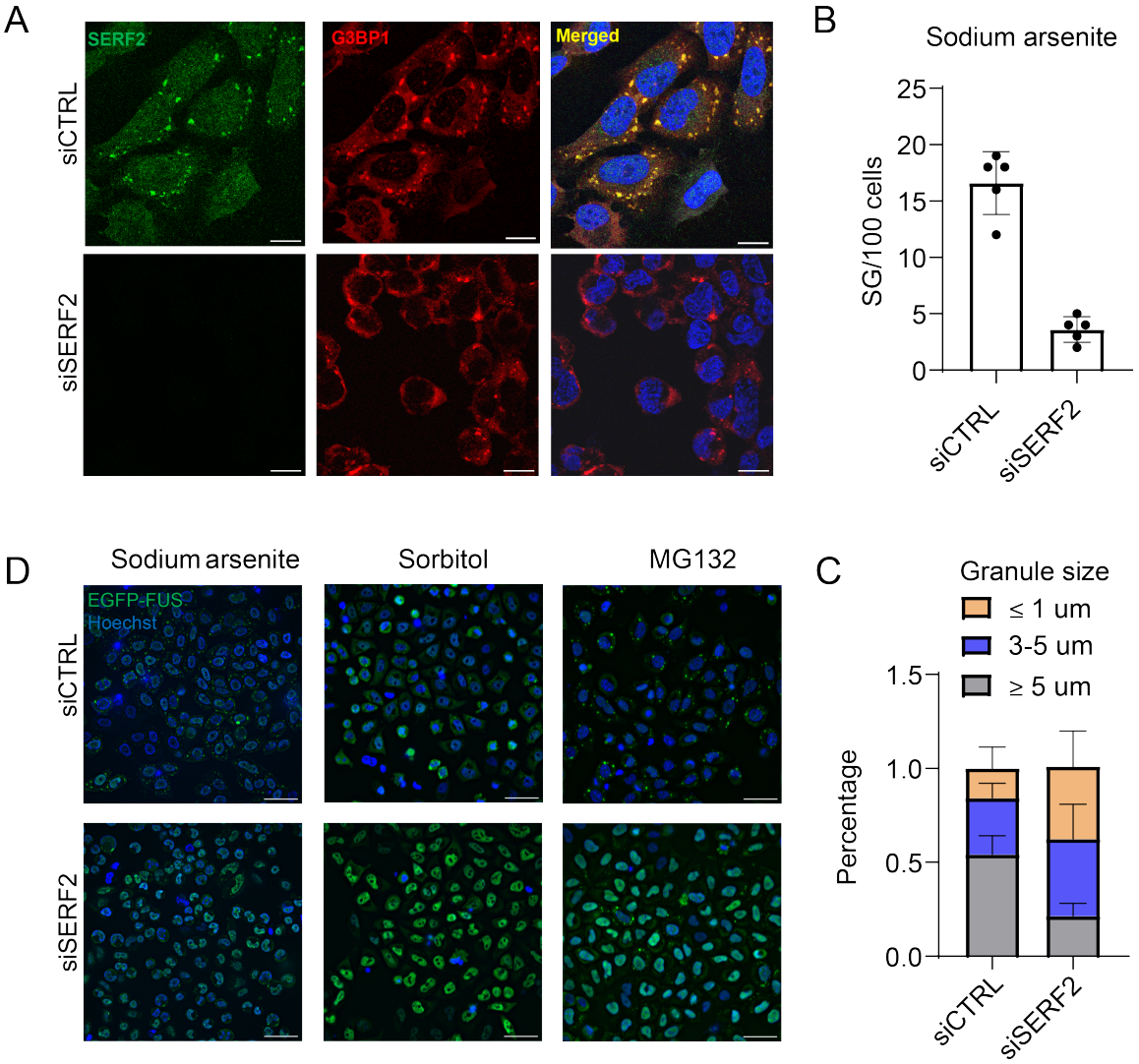

Figure 2: SERF2 regulates stress granule formation. (**A**) Immunofluorescence of SERF2 and G3BP1 in U2OS cells treated with 0.5 mM sodium arsenite for 1 hour. Scale bar is 10 µm. **(B)** Quantification of stress granule number under sodium arsenite treatment. Error bars are calculated from stress granules distributed in five independent regions of interest with 100 cells. **(C)** Quantification of stress granule sizes in sodium arsenite treated (0.5mM for 1 hour) U2OS cells with control knockdown (siCTRL) or SERF2 knockdown (siSERF2). (**D**) Live-cell imaging of EGFP-FUS HeLa Kyoto cells with siCTRL or siSERF2, treated with different stresses (Sodium arsenite, 0.5 mM; Sorbitol, 0.4 M; MG132, 10 µM) for 1 hour. Scale bar is 50 µm.

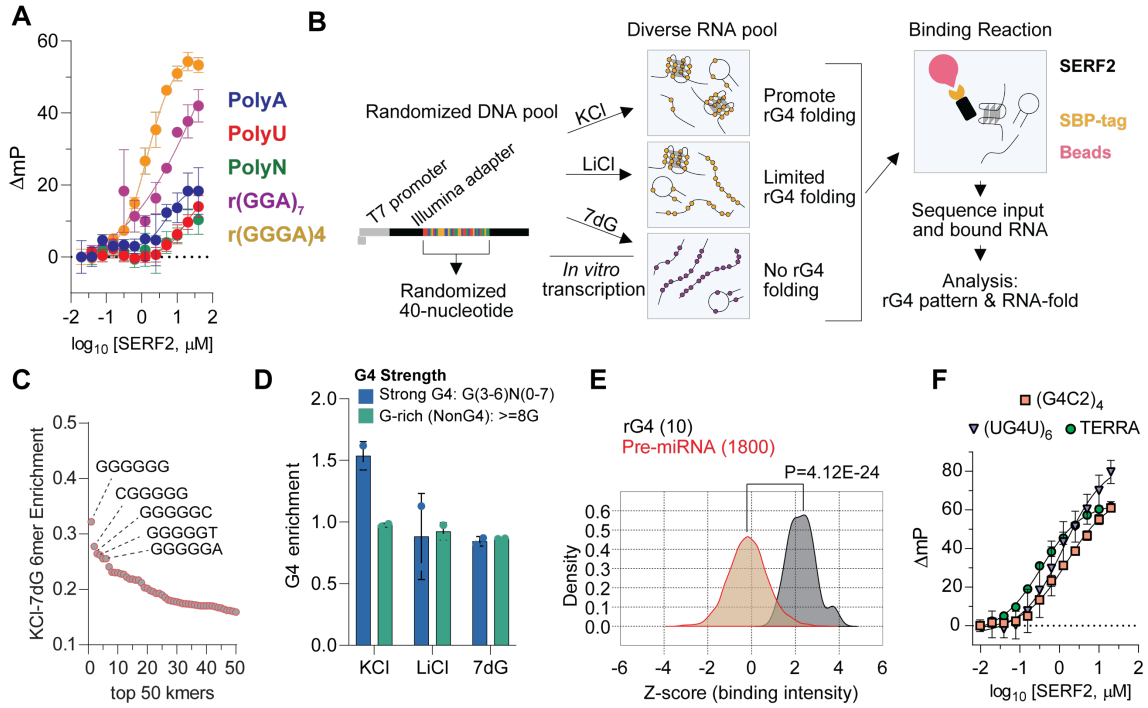

Figure 3. High-throughput screening of SERF2 binding substrates. **(A)** Fluorescence polarization (FP) assay to measure the binding specificity of SERF2 with random ribopolynucleotides and G4 quadruplex sequences. **(B)** A schematic representation of RNA bind-n-seq (RBNS) experiments with varying RNA pools. A randomized DNA oligo was transcribed to RNA and folded in KCl or LiCl. An additional RNA pool was made by replacing guanines (G) with 7-deaza (7dG) to limit rG4 quadruplex folding while preserving the sequence. These pools were mixed with GST-SERF2, at the indicated concentration, and bound RNA was isolated and sequenced to ~10-20+e6 reads. **(C)** RNA bind-n-seq analysis for enrichment of 6-mers in KCl versus RNA made with 7dG in sample mixture containing 50 nM SERF2. The guanine-rich 6mers in the top 5 kmers are indicated. **(D)** Enrichment of rG4 quadruplex patterns, in each of the indicated conditions, with varying G4 quadruplex strengths in 50 nM SERF2 RNA bind-n-seq. Sequences containing 3 or more guanines in the G-tetrad are referred to as strong G4 quadruplexes and sequences with >=8 guanines but lacking a defined G4 forming motif are referred to as non-G4 quadruplexes. **(E)** Average binding intensities for SERF2 obtained from FOREST analysis containing 1800 human pre-miRNA and 10 rG4 quadruplex sequences. The p-value was determined by the two-tailed Brunner–Munzel test. **(F)** The binding affinity between SERF2 and three different rG4 quadruplexes as indicated prepared in 20 mM NaPi and 100 mM KCl pH7.4. 20 nM of 6-FAM labeled G4 quadruplexes, or polynucleotides was mixed with varied concentration of protein, as indicated, for 1 hour at room temperature and the change in fluorescence polarization was measured at 25°C. The standard deviations are calculated from three independent replicates.

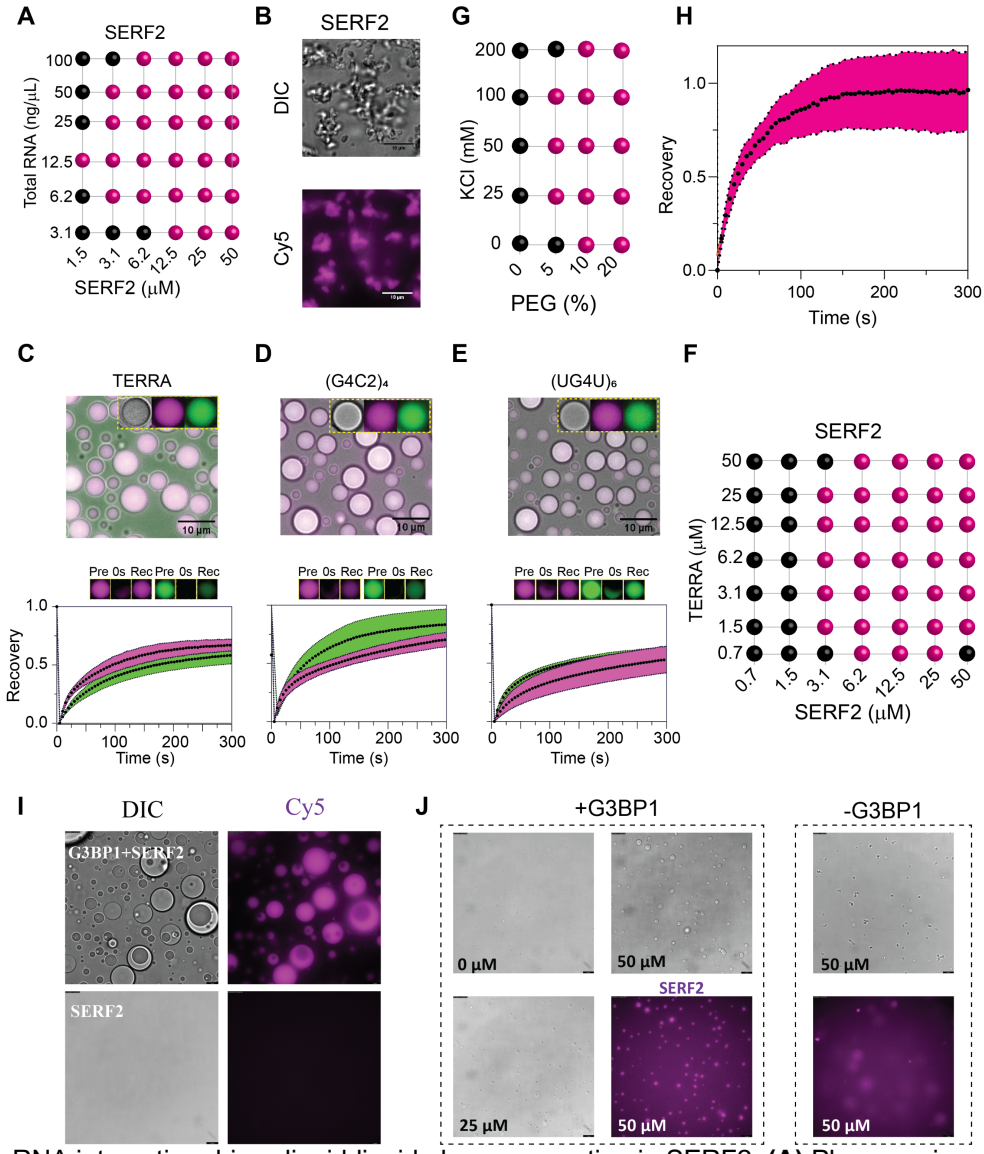

Figure 4. RNA interaction drives liquid-liquid phase separation in SERF2. **(A)** Phase regimes illustrating LLPS in SERF2 as a function of total RNA concentration extracted from HEK293T cells. **(B)** Fluorescence imaging shows gel-like structures in 50 μM SERF2 (left) mixed with 200 ng/μL of total RNA containing 10% (w/v) PEG8000 incubated for 30 minutes at room temperature. (**C-E**) 50 μM of SERF2 dissolved in 20 mM NaPi (pH 7.4), 100 mM KCl readily undergoes LLPS formation when mixed with equimolar concentration of rG4 quadruplexes that include TERRA, $(G4C2)_4$, and $(UG4U)_6$. The sample mixture contains 1/200th Cy-5 label protein (purple signals) and 6-FAM rG4 quadruplex (green signals) as indicated in the figure inset. Two-component FRAP analysis are done to measure the protein recovery (purple plot) and rG4 quadruplex recovery (green plot) rates in the SERF2-rG4s droplets. The pre-bleached, after-bleached (0 s) and recovered droplets (300 s) are shown on the top of each FRAP plot. The FRAP data were fitted in GraphPad Prism using non-linear regression one-phase association model to obtain the recovery halftime (t1/2). **(F,G)** Schematics showing SERF2 and TERRA sample mixture phase separation in 10% PEG8000 at different protein to RNA concentrations (F), and at different salt and PEG8000 concentrations **(G)**. The black and purple spheres in all phase regime represents no formation or formation of droplets, respectively. **(H)** Dynamics and recovery of Cy5-labelled proteins in SERF2-total RNA droplets obtained by FRAP analysis suggest the gel-like structures are dynamic and reversible. Standard errors are calculated by analyzing 8 isolated droplets subjected to FRAP. **(I)** DIC and Fluorescence images showing co-phase separation of SERF2 (purple) and G3BP1 with HeLa total RNA. (J) SERF2 facilitates G3BP1-RNA condensation. 0, 25 or 50 μM of SERF2 protein was added to 12.5 ng/μl of HeLa total RNA with or without 25 μM of G3BP1.

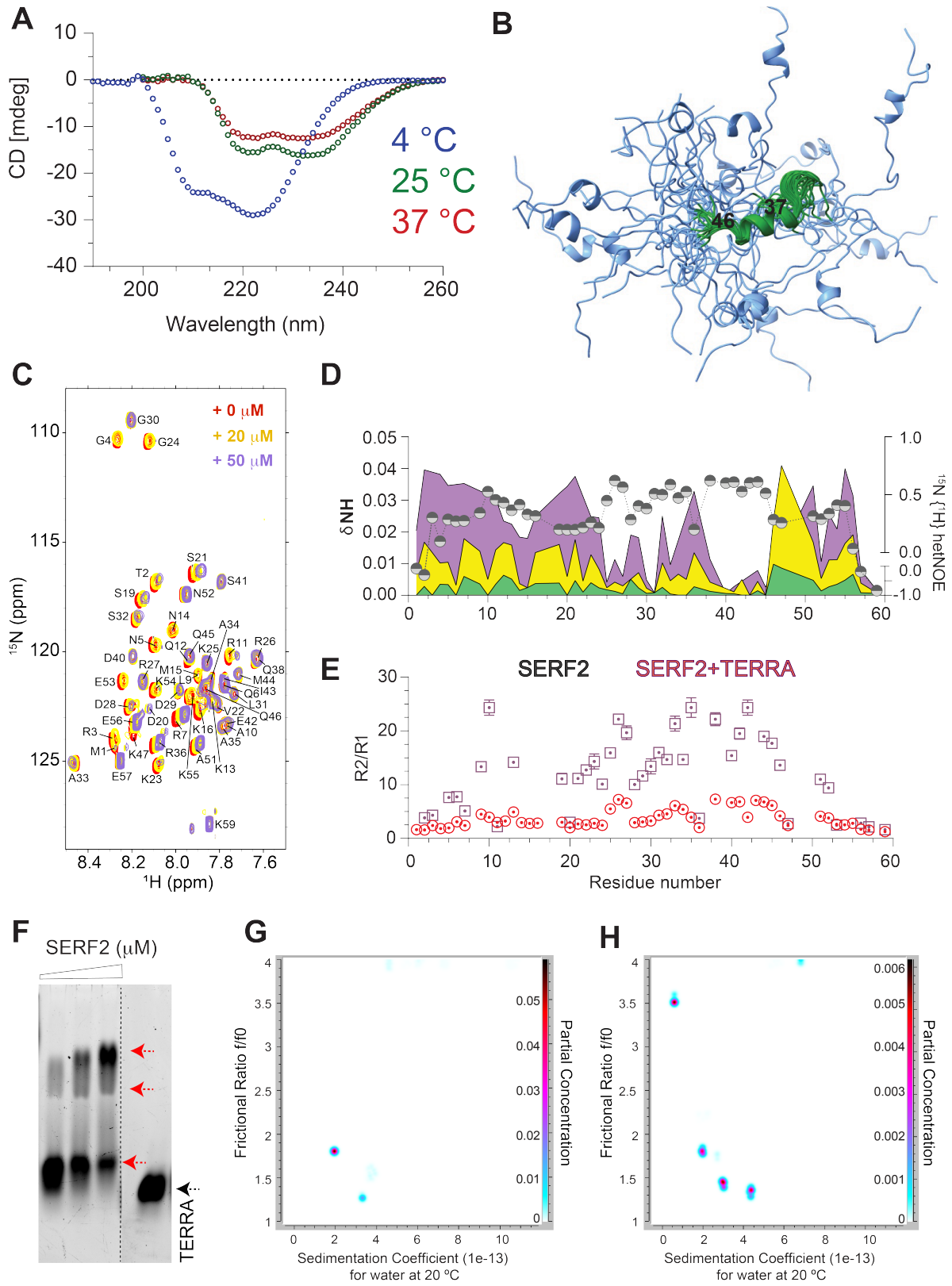

Figure 5. The disordered and dynamic SERF2 domain binds TERRA G-quadruplex. **(A)** Secondary structure analysis of human SERF2 by CD spectroscopy at the indicated temperature in 20 mM NaPi

(pH7.4), 100 mM KCl. **(B)** 20 best NMR ensemble models of SERF2 calculated using Cyana. The average converged structure spanning residues 37-46 in SERF2 is shown in green. **(C)** 15N/1H amide resonance assignment (red spectrum) of human SERF2 mixed with 20 μM (yellow) and 50 μM (purple) TERRA. **(D)** The 15N/1H chemical shift perturbations (CSPs) are calculated from (**C)** and CSPs are plotted for each assigned residue in SERF2 with increasing TERRA concentrations as indicated in colors. The heteroNOE values of SERF2 in the absence of TERRA were plotted with the CSPs on the right y-axes to highlight the correlation between SERF2 dynamics and TERRA interaction. **(E)** 15N relaxation rates R2/R1 demonstrating a significant change in 200 μM SERF2 dynamics interacting with 100 μM TERRA as a function of residues. Protein and RNA samples are dissolved in buffer containing 8% D_20 and spectra are recorded at 4 °C on a Bruker 600 MHz spectrometer. **(F)** EMSA of 5 μM TERRA rG4 quadruplex (black arrow) after 30 minutes incubation with increasing SERF2 (2.5, 5 and 10 μM) concentration in 20 mM NaPi (pH7.4), 100 mM KCl buffer. Red arrows indicate an increase in multimeric species as a function of increasing protein concentration. **(G,H)** 2D analysis plots derived from AUC for 4.7 μM TERRA mixed without (**G**) or with 2-molar excess SERF2 (**H**). The partial concentration shown in color on right y-axes represents the abundance of individual species in the sample solution.

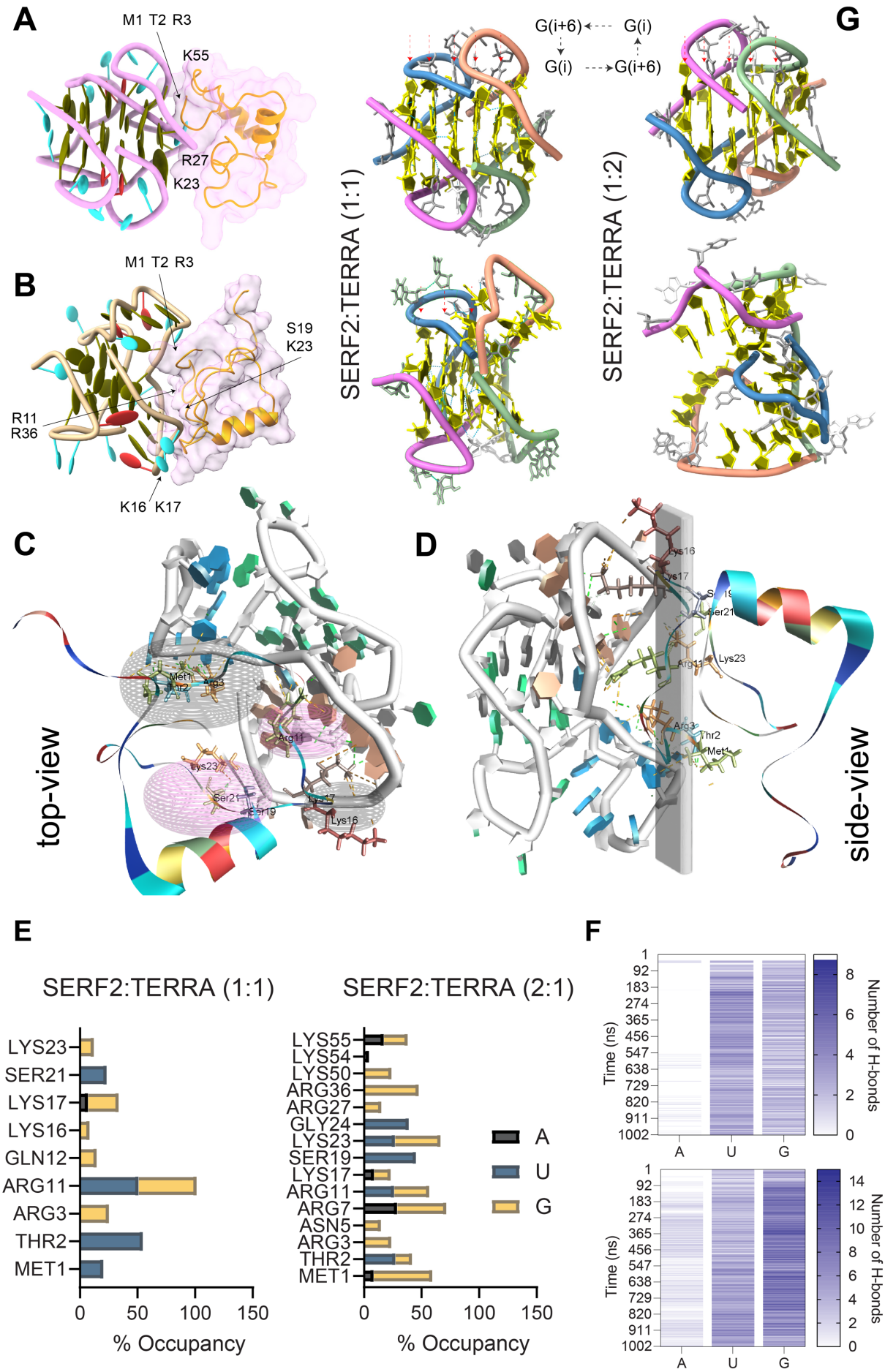
A
M1 T2 R3
K55
R27
K23
B
M1 T2 R3
S19
K23
R11
R36
K16 K17
SERF2:TERRA (1:1)
G(i+6)
G(i)
G(i)
G(i+6)
SERF2:TERRA (1:2)
G
C
top-view
D
side-view
E
SERF2:TERRA (1:1)
LYS23
SER21
LYS17
LYS16
GLN12
ARG11
ARG3
THR2
MET1
0 50 100 150
% Occupancy
SERF2:TERRA (2:1)
LYS55
LYS54
LYS50
ARG36
ARG27
GLY24
LYS23
SER19
LYS17
ARG11
ARG7
ASN5
ARG3
THR2
MET1
0 50 100 150
% Occupancy
A
U
G
F
Time (ns)
1 92 183 274 365 456 547 638 729 820 911 1002
A U G
Number of H-bonds
0 2 4 6 8
Time (ns)
1 92 183 274 365 456 547 638 729 820 911 1002
A U G
Number of H-bonds
0 2 4 6 8 10 12 14

Figure 6. High-resolution structural insights into SERF2-TERRA complex. **(A)** Cartoon structure of SERF2-TERRA complex built using HADDOCK by parsing ambiguous NMR restraints. **(B)** Average MD structure of SERF2-TERRA (1:1) complex obtained from 1 μs all-atom simulation. The labelled residues shown are involved in hydrogen-bond interactions. The tetrameric TERRA structure is shown in pink and protein in orange. **(C-D)** Cartoon shows the top- and side-view of a quadrupole-like (ellipses, **C**) and planar (vertical slab, **D**) interaction in SERF2-TERRA complex. Residues generating the quadrupole-like interaction and TERRA structure distortion are labelled and hydrogen bonds are indicated with dashed lines. **(E)** Hydrogen bond occupancy (%) of SERF2 residues in 1:1 and 1:2 complex with individual TERRA nucleotides A/U/G estimated from 1 μs MD trajectory using VMD program. **(F)** Existence and disappearance of total number of hydrogen-bonds in SERF2-TERRA 1:1 (top) and 2:1 (bottom) complex as a function of nucleotides A/U/G. The scale represents the total number of hydrogen-bonds between SERF2 and TERRA at a give time frame. **(G)** 3D structure of tetrameric TERRA (PDB ID: 2M18) before and after 1 μM MD simulation shows G-tetrad distortion in SERF2:TERRA 1:1 (center) and 2:1 (right) complex. G-tetrads are indicated by red arrows and each TERRA unit in the tetrameric structure are represented with different colors. A G-quartet in TERRA is formed by guanines in i and i+6 as shown on the top.