UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Phase Separation as a Model of Nucleoprotein Organization

Permalink

https://escholarship.org/uc/item/93h7b022

Author

Jack, Amanda

Publication Date

2022

Supplemental Material

https://escholarship.org/uc/item/93h7b022#supplemental

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Phase Separation as a Model of Nucleoprotein Organization

By

Amanda Jack

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Biophysics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Ahmet Yildiz, Chair
Professor Kathy Collins
Professor Abby Dernburg
Professor Gary Karpen

Spring 2022

Abstract

Phase Separation as a Model of Nucleoprotein Organization

by

Amanda Jack

Doctor of Philosophy in Biophysics

University of California, Berkeley

Professor Ahmet Yildiz, Chair

Recent work has revealed the importance of liquid-liquid phase separation (LLPS) as a mechanism to organize cells through the formation of non-membrane-bound cellular compartments such as nucleoli, Cajal bodies, and stress granules. These compartments concentrate specific proteins and nucleic acids through weak, noncovalent interactions. I investigated phase separation as an organizing principle for both eukaryotic cells and viruses. In particular, I have studied the shelterin protein complex, which binds and protects human telomeres. Telomeres form unique nuclear compartments that prevent degradation and fusion of chromosome ends by recruiting shelterin proteins and regulating access of DNA damage repair factors. To understand how these dynamic components protect chromosome ends, I combine in vivo biophysical interrogation and in vitro reconstitution of human shelterin. I show that shelterin components form multicomponent liquid condensates with selective biomolecular partitioning on telomeric DNA. Tethering and anomalous diffusion prevent multiple telomeres from coalescing into a single condensate in mammalian cells. However, telomeres can coalesce when brought into contact via an optogenetic approach. TRF1 and TRF2 subunits of shelterin drive phase separation, and their N-terminal domains specify interactions with telomeric DNA in vitro. Telomeric condensates selectively recruit telomere-associated factors and regulate access of DNA damage repair factors. I propose that shelterin mediates phase separation of telomeric chromatin, which underlies the dynamic yet persistent nature of the end-protection mechanism.

My work also addressed how interactions between the nucleocapsid (N) protein and viral RNA in the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) drive LLPS. SARS-CoV-2 infection causes COVID-19, a pandemic that seriously threatens global health. SARS-CoV-2 propagates by packaging its RNA genome into membrane enclosures in host cells. Packaging of the viral genome into the nascent virion is mediated by the N protein, but the underlying mechanism is not fully understood. I found that the N protein forms biomolecular condensates with viral genomic RNA both in vitro and in mammalian cells. While the N protein forms spherical assemblies with homopolymeric RNA substrates that do not form base-pairing interactions, it forms asymmetric condensates with viral RNA strands. Cross-linking mass spectrometry identifies a region that mediates interactions between N proteins in condensates, and truncation of this region disrupts phase separation. I also identified small molecules that alter the formation of N protein condensates and inhibit the proliferation of SARS-CoV-2 in infected

cells. These results suggest that the N protein may promote biomolecular condensation to package the SARS-CoV-2 RNA genome into a viral particle.

Acknowledgments

My graduate studies were funded by the Molecular Biophysics Training Grant and the National Science Foundation Graduate Research Fellowship.

I could not have completed this work without a wide network of support, both professional and personal. First, I would like to thank my thesis advisor, Ahmet Yildiz. Throughout my time in the Yildiz lab, Ahmet has pulled me out of rabbit holes, pushed me to be a bolder scientist, and always made time for me. Thanks for the unwavering support and the baklava.

Next, I thank my mentor, Luke Ferro. Thanks for getting me off the ground, being my idea rat, and sharing Milo.

I also want to express my gratitude to my co-authors: Yoonji Kim, Amy R. Strom, Daniel S.W. Lee, Byron Williams, Jeffrey M. Schaub, Elizabeth H. Kellogg, Ilya J. Finkelstein, Clifford P. Brangwynne, Michael J. Trnka, Eddie Wehri, Amrut Nadgir, Xammy Nguyenla, Douglas Fox, Katelyn Costa, Sarah Stanley, and Julia Schaletzky. This research would not have been possible without each of them. I also thank Ruensern Tan, Jonathan Fernandes, John T. Canty, and the other members of the Yildiz laboratory; you're both indispensable and a lot of fun to be around.

Joshua Riback, Jorine Eeftens, Yi-Che Chang, David Sanders, Evangelos Gatzogiannis, Yavuz S. Dagdas, Shunsuke Shimobayashi, David Sanders, Sofi Quinodoz, Lindsay Becker, Titia de Lange, Huaiying Zhang, Thomas Graham, Kathy Li, Nevan Krogan, Amy Gladfelter, Volker Thiel, Don Cleveland, Denise Schichnes, the UC Berkeley Biological Imaging Facility, the UC Berkeley MacroLab, and the UC Berkeley Cell Culture Facility also contributed training, reagents, or helpful discussions. Thanks as well to the members of my thesis committee, Kathy Collins, Abby Dernburg, and Gary Karpen, for their feedback and support.

I would not have made it this far without board game marathons with Sara Volz, hiking with Meagan Esbin, ledge convos with Noah Rogers, concerts with Cameron Williams, the Spice Girls D&D crew, the UVC, Howard Shore, and finally my sister, Gracelyn Jack, and my mom, Elaine Jack – through video chats and mixtapes, amid exit ramps and brain worms, you've always had my back. Each of you has my gratitude (and possibly blame).

Introduction

Liquid-liquid phase separation (LLPS) has emerged as a mechanism to create membraneless cellular compartments or "condensates," such as nucleoli, Cajal bodies, and stress granules [1, 2]. These structures contain high local concentrations of proteins and nucleic acids that condense into liquid-like assemblies through multivalency and noncovalent interactions, selectively excluding non-interacting molecules [3, 4]. These compartments serve a wide variety of functions within the cell, including favoring or disfavoring reactions by modulating protein concentration, targeting molecules to different organelles, responding to stimuli, and exerting forces [5].

Recent research has focused on how interactions between proteins and nucleic acids drive phase separation. In particular, heterochromatin has been identified as a phase-separated compartment, suggesting new models for how chromosomes are organized [6, 7]. The heterochromatin protein HP1α forms liquid droplets in vitro in a manner dependent on its concentration and phosphorylation state, and it binds and compacts multi-kilobase DNA fragments [7]. This suggests that HP1α could help drive the packaging of chromosomal DNA and that the cell can regulate heterochromatin domain formation by modulating production or post-translational modification of the HP1α protein.

Modeling chromatin as a biological condensate could also help explain the mechanisms of transcriptional silencing. Thinking of heterochromatin as a more compact form of chromatin that sterically hinders RNA polymerase from accessing the DNA predicts that molecules should diffuse more slowly inside the heterochromatin compartment; however, this is not observed [6]. Instead, inert molecules diffuse quickly inside and outside the heterochromatin phase, with slower diffusion at the boundary between chromatin domains, as the phase separation model predicts [6]. Additionally, nucleosomes themselves have been shown to phase separate in vitro, and their condensation propensity depends both on the length of the nucleosomal array and the positioning of nucleosomes within it [8]. This previous research has broad implications in modeling chromatin structure as well as understanding how the cell controls nucleation and localization of heterochromatin compartments.

However, a better understanding of the interactions that drive heterochromatin condensate formation also suggests a more general model for nucleoprotein condensation. In this model, a nucleic acid scaffold can recruit a locally high concentration of a phase-separation-prone binding protein, and the stabilization of a liquid phase depends on the nucleic acid-protein interactions (i.e. the length of the scaffold and the concentration of the protein) and the protein-protein interactions (i.e. oligomerization as well as dynamic interactions modulated by charge, post-translational modifications, and amino acid motifs). In my dissertation, I will apply this model to both eukaryotic and viral nucleoprotein complexes.

Early research on chromatin phase separation focused on heterochromatin; however, in the first part of my dissertation, I will discuss evidence that telomeric chromatin is also a phase-separated structure. Telomeres are repetitive DNA sequences that form the ends of linear eukaryotic chromosomes. In dividing human cells, they are composed of 2-20 kilobases of double-stranded TTAGGG repeats followed by 50-500 basepairs of single-stranded 3' overhang [9]. Telomeres

contain nucleosomal arrays with similar post-transcriptional modifications as heterochromatin (although the nucleosomes are more mobile and packed more tightly) [10]. However, the most important telomere-associated proteins are the components of the six-component shelterin complex [9]. In humans, the homologous shelterin components TRF1 and TRF2 specifically bind to double-stranded telomeric tracts and recruit other subunits to telomeres. POT1/TPP1 binds to the single-stranded telomeric overhang, and TIN2 interconnects TRF1, TRF2, and TPP1. RAP1 binds to the hinge region of TRF2 [11, 12].

Telomeres serve two main functions in protecting the coding sequence. First, they help solve what is known as the end replication problem [13]. The cell's replication machinery cannot completely replicate the extreme terminus of the DNA during lagging strand DNA synthesis, meaning that chromosomes shorten during each cell division [9]. However, the presence of a noncoding telomeric cap ensures that genetic information is not lost [9]. Secondly, progressive shortening of the telomeric sequence leads to replicative senescence: After successive cell divisions, the telomere becomes critically short, which triggers activation of the DNA damage response (DDR) machinery and prevents further cell division [14, 15]. This process has been widely associated with aging [16], and malfunctions in this pathway, including erroneous telomere extension and chromosome fusions, can allow the cell to bypass senescence and continue dividing, leading to cancer [13, 17].

The structure of the telomere also presents a problem for the cell: Even when the telomere is not critically short, DDR proteins should theoretically recognize it as a damage site, either at the single-stranded overhang or at the junction between the double-stranded and single-stranded DNA [18]. However, while some DDR proteins can recognize and bind telomeric DNA in dividing cells, telomeres do not normally activate a DDR pathway [19]. Shelterin is the most important factor in protecting the chromosome end from recognition by the DDR machinery [18]. In particular, TRF2 is the key component in DDR suppression: knocking down TRF2 in live cells triggers the formation of telomere dysfunction-induced foci (TIFs) via the ataxia-telangiectasia mutated (ATM) and non-homologous end-joining (NHEJ) pathways [20]. TRF2 has been shown to drive the formation of t-loops, wherein the single-stranded telomere overhang invades and basepairs with the double-stranded region [21, 22]. It has been hypothesized that this restructuring of the overhang into a shelterin protective cap is the mechanism for DDR suppression [23, 24]. However, multi-kilobase telomeres (which should be long enough to allow for t-loop formation) still trigger replicative senescence [14, 25, 26]. More research is needed to understand how the interactions between the shelterin complex and telomeric DNA lead to telomere length regulation and end protection. I propose phase separation as a model for understanding how shelterin protects and restructures the telomere.

Phase separation is a compelling model for this system since the ability to stabilize a liquid phase should depend both on telomere length (the number of binding sites available to recruit shelterin) and shelterin concentration. Shortening telomeres or knocking the shelterin subunits out of the cell could disfavor condensate formation, leaving mid-length telomeres deprotected. Phase separation could also explain how the telomere can selectively recruit associated factors (for instance, telomerase, the enzyme that extends the telomere) while preventing DDR factor access [27]. In this work, I show that shelterin is capable of phase separation behavior in vitro. In addition, the ability of condensates to fuse is an important property of phase separating systems,

and we were able to observe telomeres coalescing on-demand in living cells, lending support to phase separation as a model for telomere structure. This relatively new way of thinking about how protein-DNA interactions drive chromatin organization may be important to further our understanding of chromosome end protection and telomere length maintenance.

In the second part of my dissertation, I apply the idea of nucleoprotein phase separation to the SARS-CoV-2 nucleocapsid (N) protein. N is a dimer that binds the virus' 30 kilobase single stranded RNA genome and carries out a variety of functions in the infected cell. N is the most highly expressed viral protein early in infection, and it accumulates at the replication transcription complex, where it enhances replication and transcription of both genomic and subgenomic RNA [28-32]. N is also responsible for restructuring the viral RNA into shell-shaped viral ribonucleoprotein complexes (vRNPs) which then bind to the viral membrane (M) protein on the surface of the ER-Golgi intermediate compartment (ERGIC) to trigger the budding of the vRNP complex [33, 34]. However, how N specifically packages a single copy of the genomic RNA per virion is not fully understood [30, 35, 36]. N (and its phosphorylation state) has been implicated in controlling the switches from subgenomic to genomic replication [32, 37, 38] and from transcription to viral packaging [31, 36]. In addition, the subgenomic RNA coding for N is the most abundant of the subgenomic RNAs during infection [39], and N protein interacts with its own RNA sequence [35], so changes in the concentration or length of the viral RNA could also drive these changes in N function. N presents a new target for inhibiting viral proliferation during infection; however, the mechanism of its interaction with genomic RNA is not fully understood.

Phase separation is an increasingly attractive model for viral organization in an infected cell, since viruses code for few proteins but express them at high levels [40]. RNA viruses that replicate in the cell cytoplasm (including SARS-CoV-2) typically concentrate their replication machinery within specialized compartments which resemble biological condensates [41-43]. These compartments have been shown to undergo fusion in vesicular stomatis virus [42], and studying reconstituted proteins in vitro generates predictions for their behavior in infected cells. For instance, the nucleocapsid protein of HIV-1 contains conserved domains that are required for ribonucleoprotein assembly and puncta formation in infected cells and also drive phase separation in vitro [43]. This indicates that phase separation could be a readout to study protein-protein and protein-nucleic acid interactions in a viral system. In addition, the measles virus separates the function of the N protein into two proteins: nucleoprotein, which binds genomic RNA, and phosphoprotein, which oligomerizes and chaperones nucleoprotein [41]. Different stoichiometric ratios of these proteins phase separate in vitro, but they cannot phase separate individually, suggesting that both these types of interactions – protein-protein and protein-RNA – are necessary to drive condensation. In vitro condensation can thus be used to observe changes in biomolecular interactions, and phase separation of viral proteins could be an additional druggable target to fight infection.

Interestingly, the changes that are observed in the cell during the SARS-CoV-2 infection cycle – namely, changes in the length and abundance of the RNA scaffold as well as changes in N phosphorylation state – would be predicted to affect condensation propensity as well, suggesting that modulating phase separation might drive the switch from replication to packaging [44, 45] and suggesting that the stoichiometric dependence of condensate formation may favor packaging

of a single RNA genome [35, 36]. Further evidence of viral phase separation is found in N's interactions with the host's immune response machinery: N protein is recruited to stress granules, which are also phase separated structures, and may inhibit their formation [31, 40, 46-48]. Thinking about N as a phase separating protein could help us understand its interactions with native proteins as well as explain N's ability to suppress the innate immune response [49-51], allowing us to create novel or more effective modes of treatment.

In my work with the SARS-CoV-2 system, I show that N requires an RNA scaffold to form condensates and that the material properties of these condensates are strongly affected by the structure of the RNA. I also test the effects of sequence motifs, post-translational modifications, and small molecules on phase separation in order to predict how these changes in the N protein or to the cellular environment will affect N puncta formation as well as its function in the infection cycle. Notably, the small molecules that nonspecifically modulated phase separation in vitro also inhibited viral replication in infected cells. These results suggest that we can use simplified, reconstituted systems and visual measurements of in vitro phase separation to predict the behavior of a protein within more complex cellular systems.

Part 1: Compartmentalization of Telomeres through DNA-scaffolded Phase Separation

Preface

The nucleus contains the biological software of the cell – the genome – which is organized into individual chromosomes. Eukaryotic chromosomes end with telomeres, nucleoprotein structures containing repetitive DNA, which protect the genome over successive cell divisions [52-54]. Unlike germline cells in which the average telomere length is set, the telomeres in somatic cells shorten over time [55-57]. This mechanism has been viewed as a tumor-suppressing pathway, as the gradual shortening of telomeres leads to replicative senescence or cell death [52].

In humans, telomeres consist of 2-20 kilobases of double-stranded TTAGGG (dsTEL) repeats followed by 50-500 bases of single-stranded telomeric (ssTEL) overhang [9]. Telomeres associate with the six-protein complex shelterin [58], which prevents degradation, chromosome end-to-end fusions, and unwanted DDR [52, 53]. The shelterin protein subunits suppress a wide variety of DDR pathways at telomeres by masking the chromosome ends from being improperly recognized as DNA break sites [10, 59]. In particular, TRF2 inhibits the ATM pathway and NHEJ of telomeres [20]. Moreover, TRF1 prevents replication fork stalling [23, 60], POT1/TPP1 specifically suppresses the ataxia telangiectasia and Rad3-related (ATR) pathway, and TIN2 suppresses ATM, ATR, and NHEJ pathways [9].

End-protection by telomeres is mechanistically attributed to the formation of t-loops, wherein TRF2 enables the ssTEL overhang to invade dsTEL tracts and form lasso-like structures [21, 22]. The t-loop model provides an explanation for how shelterin sequesters the chromosome ends from the ATM and NHEJ pathways [21, 22]. However, this model does not adequately explain

how cells enter senescence while their telomeres still contain kilobases of telomeric repeats [14, 25, 26] since t-loops have been observed for telomeric DNA as short as one kilobase [61] [62]. Shelterin is also hypothesized to protect telomere ends through the three-dimensional compaction of telomeric chromatin [63], but decompaction of telomeres upon shelterin knockdown has not been observed by others [64, 65]. The network of interactions between shelterin components and telomeric DNA could also function as a selectivity barrier to regulate the preferential binding of shelterin and prevention of DNA damage response signaling [63], but interactions between shelterin and telomeric DNA are too dynamic to serve as a steric barrier. Thus, it remains unclear what physical picture best describes telomere organization and function.

LLPS has recently been implicated in controlling chromatin structure [66] and in heterochromatin domain formation [6, 7], raising the possibility that telomeric DNA may also form condensates with associated shelterin components. Consistent with this hypothesis, TRF1 and TRF2 display many of the characteristics common in phase separating systems, including intrinsically disordered regions (IDRs), a dimerization domain, and a DNA binding domain [9, 20]. However, this liquid phase model has not been tested.

We combined intracellular biophysical interrogation and in vitro reconstitution to reveal that shelterin components and telomeric DNA organize into liquid-like condensates. Using an optogenetic approach to bring two telomeres together, we found that telomeres are capable of undergoing coalescence, forming a single larger telomeric body. In living cells, we showed that telomeres exhibit quantitative signatures of multicomponent LLPS, but their hindered diffusivity results in extremely few coalescence events. I reconstituted the human shelterin complex and found that the interactions between shelterin and telomeric DNA promote the formation of liquid condensates. TRF1 and TRF2 drove phase separation of the shelterin complex, and these liquid droplets selectively recruited telomere-associated factors in vitro. I proposed that LLPS of shelterin components builds the telomere compartment and could protect chromosome ends by selectively recruiting telomere-associated factors while limiting access of DDR factors.

Results

Telomeres in living cells are liquid-like

We first investigated whether telomeres exhibit liquid-like features in human cells. We expressed TRF1 (miRFP-TRF1) and TRF2 (mGFP-TRF2 and miRFP-TRF2) and confirmed that TRF1 and TRF2 formed distinct puncta in nuclei of U2OS cells (Figure 2A-B). As previously reported [67] fluorescence recovery after photobleaching (FRAP) assays showed that TRF1 and TRF2 rapidly exchange between telomeres and the nucleoplasm (Figure 1A), which is typical for phase separating systems [68, 69]. If telomeres are liquid-like, they should coalesce and round up due to surface tension. Consistent with previous studies [70-72], telomeres exhibited subdiffusive motion and typically did not encounter one another (Figure 1B, Movie 1), likely because they were in a viscoelastic environment and tethered to chromosomes [73, 74]. Based on mean-squared displacement (MSD) analysis, we estimated that it would take ~5 days for a telomere to reach its nearest neighbor (2.4 ± 1.2 μm, mean ± SD) and as long as ~200 days to reach the average pairwise distance between telomeres (6.8 ± 3.2 μm) via diffusion (Figure 2C-D). Consistently, we were able to detect only one potential coalescence event after imaging 60 cells

for 1 h (Figure 2E), demonstrating that telomeres do not frequently merge with one another and remain distinct within living cells due to their suppressed diffusivity.

Due to the infrequency of telomere coalescence, the liquid phase model could not be tested through passive microscopic examination of telomeres in living cells. To controllably pull two or more telomeres into contact, we developed an optogenetic approach based on the Corelet system (Figure 1C) [75]. The synthetic Corelet droplets were made by triggering interaction of a phase-separation prone protein (in this case, FUS_N) with a multivalent (24-mer Ferritin) core through light-triggered heterodimerization between sspB, attached to FUS_N, and iLID, attached to the core. We tethered the droplets to telomeres with FUS_N-miRFP-TRF1, which bound the telomeric DNA and interacted with the droplet through homotypic FUS_N interactions. With light activation, two closely-positioned telomeres could be induced to nucleate FUS_N droplets, which fused to create one FUS_N droplet stably interacting with two telomeres (Figure 1D, Movie 2). Following removal of the blue light stimulus ('deactivation'), the FUS_N droplet shrank, surface tension pulled telomeres inwards, and the telomeres ultimately coalesced into a single spot in three dimensions (Figure 1F). In 46% (11 out of 24) of the attempts, we observed these droplet-guided telomere coalescence events (Figures 1E-G and 2F), which remained single spots for at least 8 minutes following the dissolution of the FUS_N droplets. In several instances, the telomeres detached from the FUS_N droplet before contacting each other and relaxed back to their original or more distal positions (Figure 2F-G, Movie 2), indicating that the local viscoelastic constraints on telomeres tend to maintain their relative separation (Shin et al., 2018).

To rule out the possibility that coalescence of telomeres is driven by linking FUS_N to TRF1, we linked iLID to TRF1 and FUS_N to mCherry-sspB (Figure 2H, Movie 3). In this case, iLID-miRFP-TRF1 only became a seed when FUS_N-mCherry-sspB was bound upon light-activation, and FUS_N-mCherry-sspB was released from the telomere after deactivation. We observed droplet-guided telomere coalescence events in 10 out of 13 attempts (77%) (Figure 2I-J), demonstrating that observed liquid-like telomere coalescence was driven by the endogenous telomere protein interactions. We observed these merger events in both U2OS (Figure 1D-F) and telomerase-positive hTERT-RPE1 cells (Figure 2F), indicating that telomere coalescence is not due to alternative lengthening of telomeres (ALT)-associated PML bodies (APBs) in U2OS cells [76-79]. Taken together, these data suggest that inducing contact of two telomeres causes their coalescence, which is consistent with telomeres behaving as liquid-like condensates.

Telomeric DNA is an oligomerizing scaffold to promote TRF1 and TRF2-mediated condensation

To examine whether components of the shelterin complex drive LLPS of telomeres, I purified human shelterin complex proteins and tested if they could create a biomolecular condensate with telomeric DNA in vitro. I first characterized whether TRF1 and TRF2 phase separated under physiological salt concentration (150 mM NaCl). While TRF2 did not form liquid droplets in the absence of DNA, the addition of short telomeric DNA with multiple TRF2 binding sites (8 dsTEL and 3 ssTEL repeats; 8ds3ss) initiated the formation of TRF2 droplets (Figures 3A and 4A-C, Table 1). I also observed that TRF2 did not form droplets with nontelomeric DNA at the same length as 8ds3ss, but still formed droplets with a 3-kb long nontelomeric DNA, suggesting that both telomeric DNA sequence and the length of the DNA backbone contribute to TRF2 phase separation.

In the presence of telomeric DNA (2.5 μM 8ds3ss), the volume of the droplets increased linearly with TRF2 concentration (Figures 3B and 4D). The minimum TRF2 concentration that triggered phase separation (c_{sat}) was 1.8 ± 0.2 μM (Figure 3B, see Methods), which was comparable to the estimated nuclear concentration of TRF2 (~1 μM) and lower than the local concentration of TRF2 at telomeres (>100 μM) [63, 80]. A telomeric substrate that could not recruit more than one TRF2 (2ds0ss) nucleated small droplets only at the highest TRF2 concentration tested (61 μM, Figure 4E), whereas increasing the length of the dsTEL tracts to 39 repeats substantially reduced the TRF2 concentration required for triggering phase separation (Figures 3C and 4E, Table 1). These results indicated that the multivalency of the DNA scaffold increases the local concentration of TRF2 dimers and drives phase separation. Interestingly, TRF2 exhibited first an increasing and then decreasing tendency to phase separate as a function of telomeric DNA concentration, and the addition of excess DNA abolished TRF2 condensation (Figure 3D-E). Such reentrant phase behavior has been reported for nucleoprotein assemblies [81-83] and indicates that the stoichiometry of TRF2 and dsTEL repeats is critical for their pairwise interactions to result in phase separation.

TRF1 also formed liquid droplets in vitro but under markedly different conditions than TRF2. I found that TRF1 forms liquid droplets in the absence of DNA (Figure 3F). The addition of DNA, either 8ds3ss or nontelomeric DNA, played an inhibitory role in TRF1 phase separation (Figures 3F and 4F), with TRF1 droplets no longer present at increased 8ds3ss concentrations (Figure 3E-H). The c_{sat} of TRF1 with or without DNA (19 ± 5 μM and 18 ± 2 μM, respectively, Figure 3G) was higher than that of TRF2 with DNA, indicating that TRF1 has a lower propensity to phase separate.

I also found that TRF1 and TRF2 droplets readily dissolved at high salt (0.5 M NaCl, Figure 4G) and coalesced when they encountered each other (Figure 3I), indicating that they exhibit liquid-like material properties. The average fusion time of TRF1 droplets (21 ± 2 s after contact, ±SE) was comparable to that of TRF2 droplets (27 ± 4 s) in the presence of DNA (Figure 4H, Movie 4). TRF1 droplets fused an order of magnitude faster in the absence of DNA (Figure 4H), suggesting that strong interactions between TRF1 with telomeric DNA increase the viscosity of these droplets.

My in vitro findings highlight how telomeric DNA can strongly impact the phase behavior of shelterin components. Because TRF1 phase separates in the absence of DNA (Figure 3F), we used the Corelet system to test whether TRF1-TRF1 interactions could create liquid condensates independent of telomeres in living cells (Figure 3J). We marked telomeres in U2OS cells by expressing miRFP-TRF2, and fused sspB to either $TRF1^{WT}$ or the TRF1 dimerization mutant ($TRF1^{A75P}$) [63, 84] to synthetically oligomerize up to 24 TRF1 molecules on the Ferritin core upon local light activation. When the Ferritin core was recruited to a single telomere, enrichment of $TRF1^{WT}$ or $TRF1^{A75P}$ at that telomere was slightly increased (Figure 3K). Interestingly, *de novo* TRF1 puncta were not observed to form when we locally activated a region away from a telomere (Figure 3K), except under very high expression conditions (Figure 4I-J). This is consistent with the concept that high concentration and valency are required for shelterin-mediated phase separation. We also obtained similar results with TRF2 in both U2OS and

hTERT-RPE1 cell lines (Figures 3K and 4I-J), demonstrating that TRF1 and TRF2 condensation is dependent on multivalent interactions with the telomeric DNA in living cells.

Differences in TRF1 and TRF2 phase separation are driven by their N-terminal domains. TRF1 and TRF2 are homologous proteins with flexible N-terminal charged domains (acidic in TRF1; basic in TRF2), structured TRFH dimerization domains [85], flexible hinge domains, and C-terminal DNA-binding Myb domains (Figure 5A) [18]. To investigate which domains were primarily responsible for phase separation, I generated fragments and domain swapping mutants of TRF1 and TRF2 (Figures 5A and 6A-B). Deleting the N-terminal acidic domain of TRF1 ($TRF1^{\Delta A}$) triggered the formation of irregularly shaped, solid-like condensates across conditions, underscoring its role in the solubility of TRF1 (Figure 5B). Interestingly, replacing the acidic domain of TRF1 with the basic domain of TRF2 ($TRF1^{Basic}$) resulted in a reentrant response of phase separation to DNA concentration, similar to $TRF2^{WT}$ (Figure 5C). In addition, swapping the acidic domain of TRF1 into TRF2 ($TRF2^{Acidic}$) inhibited phase separation of TRF2 in the presence of DNA (Figures 5B-C and 6D), similar to $TRF1^{WT}$. These results suggest that the TRF2 basic domain promotes phase separation through its interactions with telomeric DNA [86, 87], whereas the TRF1 acidic domain weakens electrostatic interactions with the DNA backbone and reduces phase separation in the presence of DNA. I also observed that removing the basic domain from TRF2 ($TRF2^{\Delta B}$) did not strongly change its phase behavior (Figure 5B), presumably because this construct is still highly positively charged and the TRFH domain of TRF2 may be sufficient for favorable interactions with telomeric DNA [88] in the absence of the basic domain. Differences in phase separation were not due to reduced DNA binding, as these mutants maintained a high affinity to bind telomeric DNA (Figure 6B).

I also tested the possible roles of IDRs and dimerization domains in TRF1 and TRF2 phase separation. Deletion of the hinge (ΔHinge) or both the hinge and N-terminal domains (ΔIDR) reduced the solubility and inhibited LLPS of TRF1 and TRF2 (Figure 5C,E). While Hinge-Myb of TRF2 was unable to drive LLPS in the presence or absence of DNA, Hinge-Myb of TRF1 formed droplets only at very high concentrations (>100 μM) (Figure 6C). Artificial dimerization of Hinge-Myb triggered phase separation in TRF1 with telomeric DNA, but additionally required the N-terminal basic domain in TRF2 (GSTSub) for phase separation (Figures 5A and 6C,F). Taken together, dimerization and IDRs of TRF1 and TRF2 are essential for phase separation, and the N-terminal domain regulates interactions with telomeric DNA in liquid droplets.

The shelterin complex phase separates in vitro

To examine the role of interactions among TRF1, TRF2, and the rest of the shelterin components in driving telomeric phase separation, I co-expressed human shelterin components in insect cells (Figure 7A). Four component complexes containing TRF1, TIN2, TPP1, and POT1 (4comp1); TRF2, TIN2, TPP1, and POT1 (4comp2); and the five-component complex that contains both TRF1 and TRF2 (5comp) eluted as a stable complex from gel filtration (Figures 7B-C and 8A). Unlike TRF1 and TRF2, neither POT1 nor co-purified TPP1 and TIN2 formed liquid condensates (Figure 8C), showing that these proteins do not phase separate on their own. Although POT1 does not specifically bind to TRF1 or TRF2 [89, 90], it partitioned into TRF1 and TRF2 droplets (Figure 8D-E), indicating that interactions within the condensate can recruit POT1 in the absence of TPP1-TIN2 and ssTEL tracts.

I then characterized the phase behavior of the shelterin complexes with and without telomeric DNA (Figures 7D and 8F). Similar to TRF1 condensates, 4comp1 efficiently formed liquid droplets in the absence of telomeric DNA, and its phase separation was inhibited by the addition of telomeric DNA. Similar to TRF2 condensates, 4comp2 only phase-separated in the presence of DNA and formed droplets across a range of telomeric DNA concentrations (Figures 7D-E and 8F). 5comp, containing equimolar TRF1 and TRF2, phase-separated across a broader range of telomeric DNA (8ds3ss) concentrations (Figure 7D) and in the absence of DNA (Figure 8F). Therefore, TRF1 and TRF2 are synergistic in driving phase separation of shelterin, possibly due to complex coacervation between their charged domains.

Next, I investigated the material state of the shelterin condensates in vitro. 5comp droplets could adhere to one another and change shape, but they could not complete fusion into a spherical droplet within 200 s (Figure 7F, Movie 4). Therefore, shelterin droplets appear to be more viscoelastic than TRF1 or TRF2 droplets and exhibit gel-like properties in the presence of telomeric DNA. Because shelterin components form subcomplexes with fewer subunits [18], and TRF1 and TRF2 are more abundant than POT1-TPP1 at telomeres [80], I asked whether changing the stoichiometry of shelterin subunits affect these condensates. The equimolar mixtures of separately purified TRF1, TRF2, TPP1-TIN2, and POT1 also formed liquid droplets with all components present and responded to changes in DNA concentration similar to co-purified 5comp (Figure 8G-H). The addition of ~3-fold excess TRF1 or TRF2 substantially reduced fusion times of shelterin droplets (Figures 7F-G and 8I-K). These results suggest that the stoichiometry of shelterin components could serve to regulate the viscoelasticity of telomeres in vivo.

Telomeres exhibit quantitative signatures of multicomponent liquids

To determine how altered phase behavior of TRF1 and TRF2 mutants might affect phase separation of shelterin complexes in vitro, I assembled shelterin complexes using N-terminal swap or deletion mutants of TRF1 and TRF2 (Figure 10A). Replacing $TRF1^{WT}$ with $TRF1^{\Delta A}$ or $TRF1^{Basic}$ resulted in phase separation of 4comp1 over a wider range of DNA concentrations (Figure 9A-B). Replacing $TRF2^{WT}$ with $TRF2^{\Delta B}$ in 4comp2 and 5comp did not inhibit phase separation (Figures 9A-B and 10B), but adding $TRF2^{Acidic}$ reduced the size and number of droplets of the complex (Figure 9A-B).

We next tested whether altered phase separation of TRF2 affects end-protection of telomeres in living cells, by expressing $TRF2^{WT}$, $TRF2^{\Delta B}$, or $TRF2^{Acidic}$ upon knockdown of endogenous TRF2 in hTERT-RPE1 cells (Figures 9C-D and 10C-F). I quantified the number of DNA damage foci formed by the localization of p53 binding protein 1 (53BP1), a downstream signaling protein recruited to these DDR foci (Figure 9C) [91]. As previously reported [92], knockdown of endogenous TRF2 led to a significant increase in the percent of nuclei with greater than ten 53BP1 foci (23.6% compared to 6.7% of untreated cells, Figures 9D and 10C-F). Consistent with previous studies [20, 93], both $TRF2^{WT}$ and $TRF2^{\Delta B}$ rescued telomere end-protection, with few cells exhibiting greater than ten 53BP1 foci (7.3% and 8.8%, respectively; Figure 9D). We found that $TRF2^{Acidic}$ also rescued end-protection (4.5%, Figure 9D), consistent with phase separation of this mutant with the rest of the shelterin complex in vitro.

To further probe whether telomere compartmentalization requires interactions between multiple components in living cells, we quantified the relative importance of homotypic vs heterotypic interactions in telomere formation in U2OS cells. We measured the nucleoplasmic, or dilute phase concentration (c_{dil}) of miRFP-TRF2 at increasing expression levels (Figure 9E-F). If homotypic interactions of miRFP-TRF2 dominate its phase separation, c_{dil} would remain constant as miRFP-TRF2 concentration is increased [4]. However, we observed that c_{dil} continues to increase with miRFP-TRF2 overexpression (Figure 9E-F). Our results suggest that telomeres are thermodynamically stabilized by heterotypic interactions, which is consistent with the necessity of telomeric DNA for shelterin condensation in live cells.

Phase separation of shelterin is modulated by telomere-associated factors

In mammalian cells, telomeres also associate with the sixth component of shelterin, RAP1, and nucleosomes, and I sought to examine their impact on telomeric phase separation in vitro. I found that RAP1 fully inhibits phase separation of TRF2 when mixed at an equimolar concentration (Figure 11A), presumably because RAP1 binding to the TRF2 hinge domain prevents this region to contribute to phase separation [83]. RAP1 also reduced the total volume of shelterin droplets without DNA when mixed at equal concentrations, and it moderately reduced the volume of 4comp2 droplets in the presence of DNA (Figures 11B and 12A-B). However, the addition of RAP1 only had a minor effect in the presence of telomeric DNA when all six shelterin subunits were present (Figure 11B), suggesting that the telomeric DNA scaffold and the interactions of the other shelterin subunits counteract RAP1's inhibitory effect on phase separation.

We also purified mono-nucleosomes wrapped with Widom positioning DNA that contains either a telomeric or a nontelomeric overhang. I observed that mono-nucleosomes do not form liquid droplets on their own (Figure 12C) but are sequestered strongly into 5comp droplets in the absence of additional DNA (Figure 11C). The telomeric nucleosomes stimulated phase separation of 5comp, while less droplet formation occurred in the nontelomeric nucleosomes or buffer-only conditions (Figures 11D and 12D). These results indicate that heterotypic interactions between shelterin and telomeric DNA drive phase separation, even in the presence of other abundant factors like nucleosomes that localize to telomeres.

Phase-separated shelterin selectively recruits associated factors.

To investigate selective permeability of shelterin droplets in an in vitro system that could mimic protection of the ssTEL overhang, I settled 5comp droplets onto surfaces coated with 8ds3ss (Figure 13A) and flowed fluorescently labeled queries into the chamber. When TRF1, telomeric noncoding RNA (TERRA) [94], or telomeric DNA were introduced into the chamber, they strongly partitioned into the droplets within a few minutes (Figures 13B-C and 14A), indicating that these settled droplets can accumulate favorable biomolecules. Consistently, TERRA partitioned less strongly into shelterin droplets containing $TRF2^{\Delta B}$ or $TRF2^{Acidic}$ (Figure 14B), likely due to the loss of the interactions between TERRA and TRF2's basic domain [95].

I then tested access of replication protein A (RPA), which activates the ATR pathway by binding to the ssTEL overhang [96-99], and the Mre11-Rad50-Nbs1 (MRN) complex, which activates the ATM pathway at DNA double-strand breaks (DSBs) [100, 101]. Both GFP-RPA and Alexa488-MRN were distributed uniformly inside and outside the droplet after 10 min, rather

than being enriched inside the droplet (Figure 13B-D, Movie 5). Downstream DDR factors PARP1, which is involved in the homologous recombination (HR) pathway [93], and the Ku70-Ku80 complex (Ku), which binds DSBs and is part of the non-homologous end joining (NHEJ) pathway [18], also exhibited near-uniform partitioning inside and outside shelterin droplets (Figures 13D and 14C-E). Furthermore, GFP-RPA diffused into the droplets more slowly than telomere-associated factors and GFP (Figure 14A), suggesting that telomeric condensates could act as a diffusion barrier to biomolecules with a large Stokes radius [102, 103]. The addition of excess TRF1 or TRF2 did not speed up the diffusion of RPA (Figure 14F), suggesting that slow diffusion of RPA is not due to the high viscosity of 5comp droplets. However, RPA partitioned more strongly into 5comp shelterin containing TRF2^{Acidic} (Figure 14G), likely due to weakening of phase separation under these conditions. These results suggest that LLPS of shelterin selectively recruits and enriches telomere-associated factors independent of their size (Figure 13E).

Discussion

In this study, we used in vivo and in vitro biophysical interrogation to demonstrate that telomeres represent a phase-separated liquid-like compartment. This compartment is formed through protein-protein and protein-DNA interactions, which give rise to the unique physicochemical properties of telomeres. I propose that the repetitive nature of telomeric DNA serves as a "super-scaffold" [104], effectively oligomerizing phase separation-prone proteins to drive the formation of a liquid compartment that protects the chromosome terminus (Figure 13E). In addition, my findings elucidate important organizing principles that likely underlie the formation of other genomic compartments.

Due to their constrained diffusion in mammalian cells, telomeres do not coalesce into a single condensate as would be expected in an equilibrium system. We speculate that random merger events would promote telomeric DNA end-to-end fusions and genome instability, and therefore the cell maintains telomeres as multiple distinct condensates. However, telomere clustering has been reported in TERT positive human cells [105] and ALT cells show fewer telomere puncta than the number of chromosome ends [106]. It remains to be determined whether end-to-end fusion of telomeres is due to higher mobility or interactions with other phase separating condensates (such as APBs) in these cells. We repurposed the Corelet system to bring telomeres together on-demand, and showed that telomeres coalesce upon contact. This optogenetic method can be used to bring other chromatin loci together and could thus be a powerful approach to study the role of genomic compartmentalization in gene regulation and cellular function.

Previous examples of intracellular phase separation have primarily focused on the role of homotypic IDR-IDR interactions that behave as single component systems and exhibit a fixed c_{sat} [107, 108]. However, many phase separating systems utilize both homotypic and heterotypic interactions to form a complex network of multicomponent interactions [5, 109]. Here, we show that the telomere is a multicomponent compartment whose formation relies on heterotypic interactions of the shelterin components and the scaffolding telomeric DNA [4]. Consistently, synthetic oligomerization of shelterin components cannot form *de novo* condensates away from telomeres in living cells, except at exceedingly high concentration and valence.

Using an in vitro reconstitution approach, I found that the TRF1 and TRF2 subunits of human shelterin form liquid droplets, in agreement with an emerging study on in vitro phase separation of TRF2 [83]. Both IDRs and dimerization domains are required for TRF1 and TRF2 phase separation, and the differentially charged N-terminal domains are responsible for their distinct properties of condensation in the presence and absence of telomeric DNA. Consistent with our in vivo results, TRF1 and TRF2 together drive phase separation when in complex with other shelterin components and telomeric DNA.

Collectively, my results are consistent with a model of telomeres as condensed liquid compartments, in which shelterin components drive local condensation around the valence-amplifying super-scaffold of telomeric DNA. A balance between the length of telomeric DNA and the stoichiometry of the component factors affect the formation and composition of telomere condensation (Figure 13E). In accordance with this model, telomere function in vivo is controlled by protein levels of TRF1 and TRF2 [110-112] and the length of telomeric repeats [10, 113]. An imbalance of these factors may destabilize the structure and deprotect the telomere ends. Partial knockdown of TRF1 and TRF2 triggers several DDR pathways even though TRF1 and TRF2 are still abundant at these telomeres [114, 115]. Additionally, as telomeres shorten in aging tissues, they fail to recruit sufficient shelterin to suppress DDR signals [116]. Consistent with these observations, I have shown that increasing the length of the telomeric DNA triggers phase separation at a lower concentration of TRF2 in vitro. Phase separation could also explain the mechanism of action of a dominant-negative allele of TRF2 within the context of the multicomponent network. This mutation is capable of dimerization with endogenous TRF2 but lacks DNA binding and the N-terminal domain, which may alter the interaction valence of the shelterin complex, leading to loss of compartmentalization and end-protection [117].

I found that in vitro shelterin droplets are more enriched with telomere-associated factors than the DDR proteins. These results suggest that through selective permeability the telomeric condensate could potentially recruit specific complexes like the DNA replication machinery and telomerase to the telomeric DNA, while limiting access of DDR factors during large-scale rearrangements of telomeric DNA. Selective partitioning of POT1, but not RPA, into shelterin droplets may also explain how POT1 can outcompete RPA binding to the displacement loop (D-loop) and the ssTEL overhang even though RPA is more abundant and has similar affinity for the ssTEL [99, 118].

Figures and Tables Related to Part 1

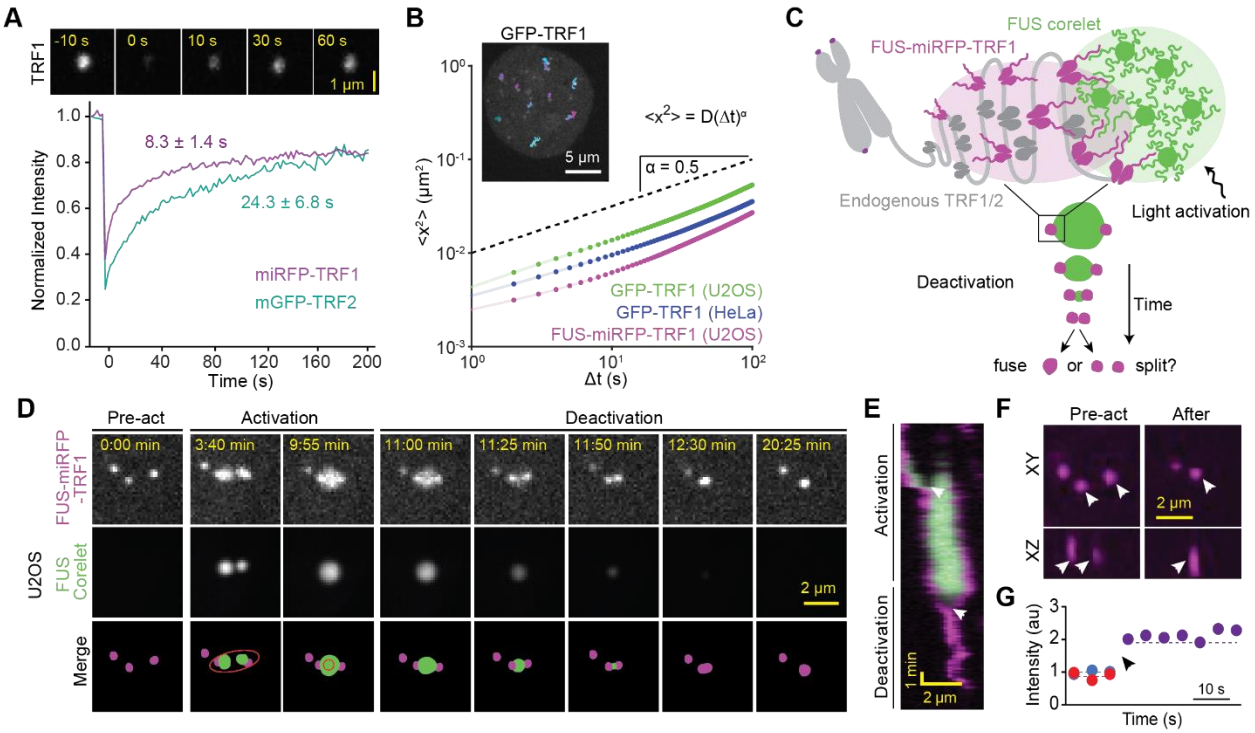

Figure 1. Telomeres in living cells exhibit liquid-like behavior. A. (Top) FRAP of miRFP-TRF1 at a telomere in a U2OS cell. (Bottom) Recovery of mGFP-TRF2 or miRFP-TRF1 fluorescence at telomeres in U2OS cells (±SD, n = 9 and 11 telomeres respectively, from one biological replicate). **B.** (Inset) Trajectories of individual telomeres are colored separately by trajectory duration in a HeLa cell expressing GFP-TRF1. MSD analysis of these trajectories revealed subdiffusive motion with exponent $\alpha = 0.54 \pm 0.01$ and diffusion coefficient $D = 2.8 \pm 0.1 \times 10^{-3} \mu m^2 s^{-\alpha}$ (±SE). The slope of the dashed line serves as a reference for $\alpha = 0.5$. **C.** Schematic of the optogenetically-induced telomere coalescence experiment: FUS_N-miRFP-TRF1 serves as a seed at telomeres to recruit FUS_N Corelet droplets upon local light activation. After two of these droplets merge, light is deactivated to pull telomeres together as the FUS_N droplet shrinks. **D.** Pre-activation, activation, and deactivation of FUS_N-miRFP-TRF1 and FUS_N Corelets in U2OS cells. The ellipse in the schematic merged images shows the local activation pattern. **E.** Kymograph shows that the two telomeres coalesce and remain as a single spot after deactivation. White arrowheads indicate the merging of FUS_N Corelet droplets and telomeres. **F.** XY and XZ views of the telomeres before and after activation. White arrowheads mark two telomeres that merge. **G.** The average intensities of the two telomeres add up (dashed lines) as they coalesce (black arrowhead) into a single spot.

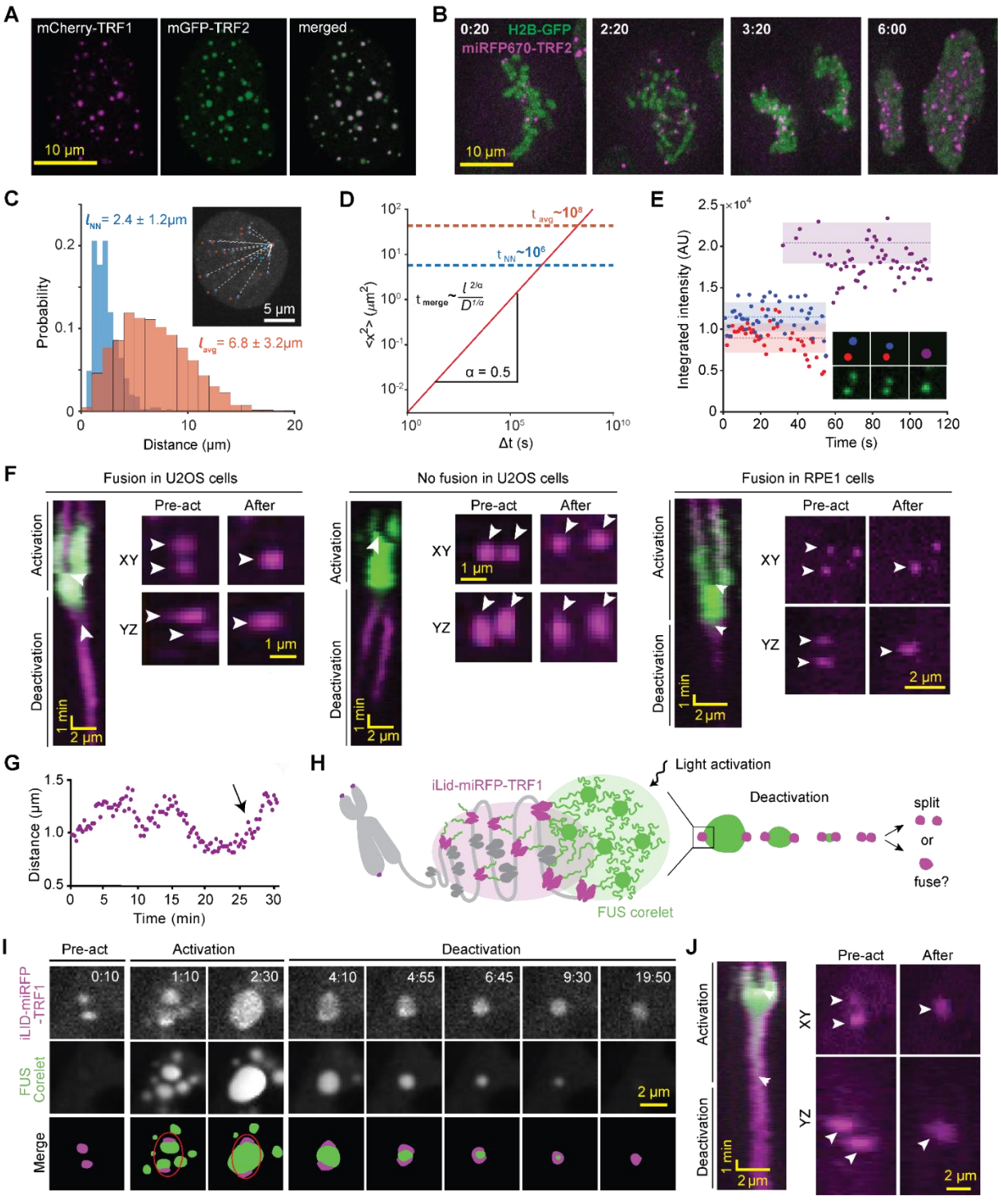

Figure 2. Characterization of telomere behavior and TRF1 and TRF2 constructs in live cells. A. U2OS cells expressing mCherry-TRF1 (left), mGFP-TRF2 (center), and a merged image showing co-localization of the two constructs (right). **B.** Panels from a 6-hour time-lapse movie, confirming miRFP-TRF2 localization at telomeres during metaphase. **C.** The histogram of the minimum pairwise distance (blue) and the average pairwise distances (orange) between

telomeres. **D.** Linear extrapolation of the timescale predicted for telomeres to randomly contact their nearest neighbor (blue) and average other telomere (orange) using the MSD of TRF1. **E.** The integrated intensities of red and blue pseudo-colored telomeres add up as they coalesce into one purple telomere. **F.** Kymographs and before and after images of successful telomere mergers (left) or no merger event (middle) with FUS_N-miRFP-TRF1 'seed' construct in U2OS cells. The right panels show successful telomere coalescence of FUS_N-miRFP-TRF1 'seed' in hTERT-RPE1 cells, suggesting the fusion is not dependent on APBs, which are not present in this cell line. **G.** In cases of unsuccessful telomere coalescence, the telomeres relax back to their original distance away from each other. **H.** Unlike FUS_N-miRFP-TRF1, iLID-miRFP-TRF1 becomes a seed upon light activation. TRF1 seed recruits FUS_N Corelet to telomeres. **I.** Pre-activation, activation, and deactivation of iLID-miRFP-TRF1 and FUS_N Corelet. The ellipse in merged images shows the local activation pattern. **J.** (Left) Kymograph shows that the two telomeres merge and remain as a single spot after deactivation. White arrowheads indicate the merging of droplets and telomeres. (Right) XY and YZ views of the telomeres before and after the coalescence event. White arrowheads mark two telomeres that coalesce upon activation.

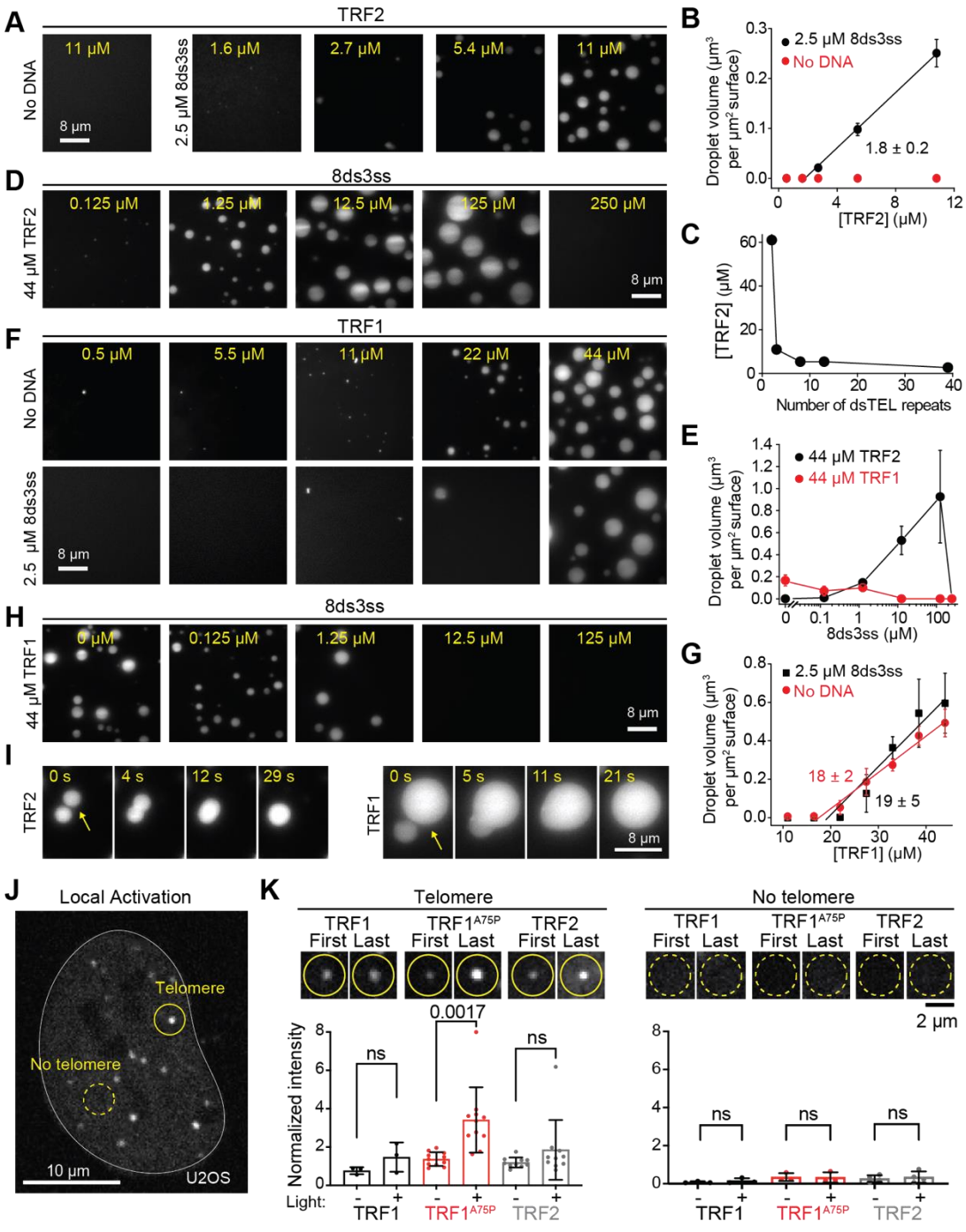

Figure 3: Telomeric DNA acts as an oligomerizing scaffold to promote TRF1 and TRF2-mediated condensation. A. Example images of Cy3-TRF2 in the presence and absence of 8ds3ss telomeric DNA. **B.** The total volume of TRF2 condensates settled per micron squared area on the coverslip in the presence or absence of 2.5 μM 8ds3ss (mean ± SD, n = 30 with three technical replicates). Linear fit (solid line) reveals c_{sat} (± SE). **C.** The minimum TRF2 concentration that exhibits phase separation with a variable number of dsTEL repeats per DNA substrate. The total concentration of dsTEL tracts was fixed to 20 μM. **D.** TRF2 has a reentrant phase behavior as a function of 8ds3ss concentration. **E.** The total volume of TRF1 or TRF2

condensates settled per micron squared area on the coverslip under variable 8ds3ss concentration (mean ± SD, n = 20 with two technical replicates). **F.** Example images of Cy3-TRF1 in the presence and absence of 8ds3ss. **G.** The total volume of TRF1 condensates settled per micron squared area on the coverslip in the presence or absence of 2.5 μM 8ds3ss (mean ± SD, n = 30 with three technical replicates). Linear fits (solid lines) reveal c_{sat} (± SE). **H.** An increase in 8ds3ss concentration inhibits TRF1 phase separation. **I.** Fusion of TRF2 (22 μM, left) and TRF1 (44 μM, right) droplets formed in the presence of 2.5 μM 8ds3ss. **J.** U2OS expressing sspB-mCherry-TRF2. TRF corelets were locally activated at a single telomere (solid circle) or away from any telomere (dotted circle). **K.** (Top) Example images show first and last frames of locally activated TRF1, TRF1^{A75P,} and TRF2 at telomeres (left; n = 3, 11, 10 cells analyzed, respectively, from one biological replicate) and away from telomeres (right; n = 3, 3, 4 cells analyzed, respectively, from one biological replicate) in U2OS cells. (Bottom) Quantification of change in intensity upon local activation, at and away from existing telomeres for WT TRF1, TRF1^{A75P}, and WT TRF2. The intensity of each locally activated telomere or region was normalized to the average intensity of all other telomeres within the same activated cell. P-values were quantified by one-way ANOVA with multiple comparisons.

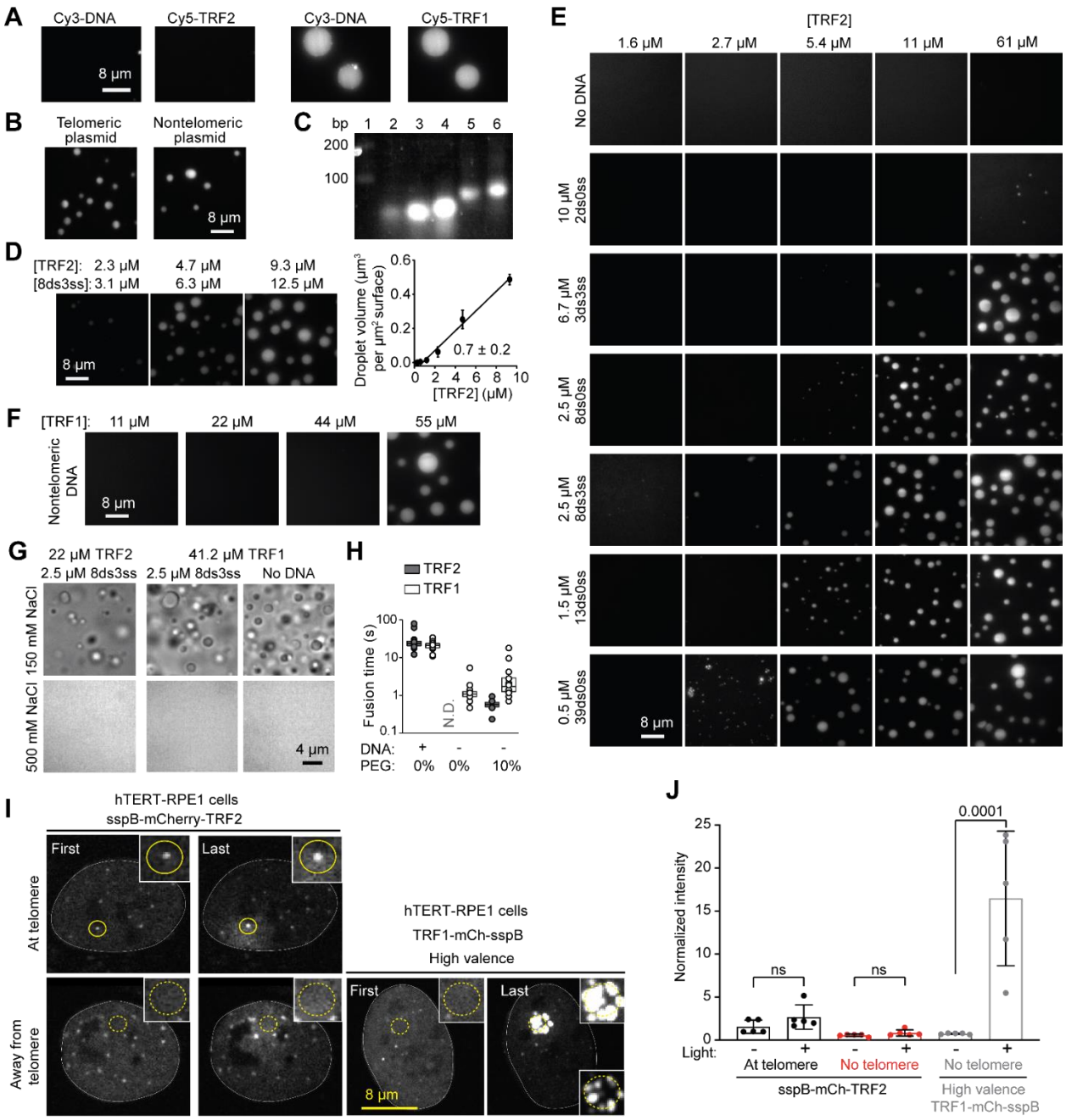

Figure 4. Phase separation of TRF2 and TRF1 in vitro**. A.** 10 μM TRF2 forms droplets with 270 ng/μL DNA plasmid containing either 39 telomeric repeats (pRST5 vector, modified) (left) or no telomeric repeats (pFastBac vector, Addgene 55221) (right). **B.** TRF2 does not stimulate phase separation with 2.5 μM of the same length of non-telomeric DNA (left), whereas non-telomeric DNA partitions into the TRF1 droplets (right). TRF1 and TRF2 concentrations were 51 and 61 μM; respectively. **C.** A 2% agarose gel stained with GelRed (1: Marker, 2: 8ds0ss sense oligo, 3: 8ds3ss sense oligo, 4: 8ds antisense oligo, 5: 8ds0ss after hybridization, and 6: 8ds3ss

after hybridization). The double-stranded DNA substrates migrate to a higher molecular weight than the individual ssDNA oligos. 75 ng of DNA were used in each sample. **D.** (Left) Condensates were formed under a constant ratio of TRF2 and telomeric DNA. (Right) The total volume of the condensates settled per micron squared area on the coverslip (mean ± SD, n = 20 with two technical replicates). Linear fit (solid line) reveals c_{sat} (± SE). **E.** Example images of Cy3-TRF2 in the presence and absence of telomeric DNA show that increasing the valency of the telomeric DNA substrates stimulates LLPS of TRF2 in vitro. The concentration of dsTEL tracts was fixed to 20 μM while varying the number of telomeric repeats in the DNA substrates. **F.** Example images show the formation of TRF1 droplets with 2.5 μM non-telomeric DNA. **G.** Brightfield images show that TRF2 and TRF1 droplets are dissolved at 500 mM NaCl. **H.** Fusion times of TRF2 and TRF1 droplets in the presence and absence of telomeric DNA and crowding agent (N.D.: not determined). TRF2 concentration was set to 22 μM, TRF1 concentration was set to 44 μM, + DNA indicates 2.5 μM 8ds3ss, and + PEG indicates 10% PEG-8000. The center and edges of the box represent the median with the first and third quartiles. **I.** sspB-mCherry-TRF2 and TRF1-mCh-sspB were expressed with the Corelet system in hTERT-RPE1 cells and localize to telomeres before light activation. sspB-mCherry-TRF2 was locally activated at a single telomere (top) and away from telomeres (bottom); shown are the first and last frames of activation. Insets show activated region. At high valence, local activation of TRF1-mCh-sspB can form droplets even away from telomeres (right), bottom inset shows alternate contrast of top inset. **J.** Quantification of change in intensity upon local activation, at and away from existing telomeres for sspB-mCh-TRF2 and TRF1-mCh-sspB in hTERT-RPE1 cells (n = 5 cells analyzed for each condition). Intensities were normalized to average intensity at non-activated telomeres in the same cell. P-values were calculated by one-way ANOVA with multiple comparisons.

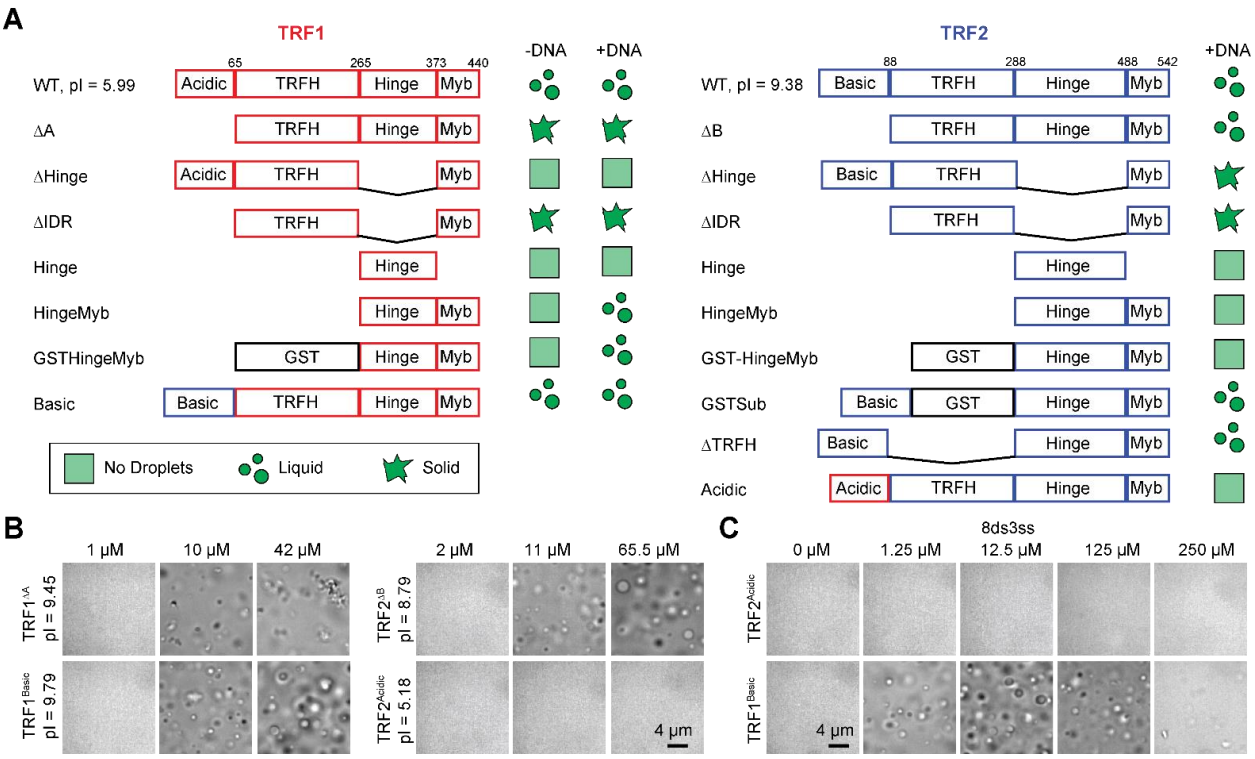

Figure 5. Differential drivers of TRF1 and TRF2 phase separation. A. Domain organization and condensate state of full-length, truncated, and engineered TRF1 and TRF2 constructs in the presence and absence of 2.5 μM 8ds3ss (pI: isoelectric point). **B.** Brightfield images taken in the presence of 2.5 μM 8ds3ss DNA show that TRF1$^{\Delta A}$, TRF1^{Basic}, and TRF2$^{\Delta B}$ form condensates, whereas TRF2^{Acidic} does not form condensates. **C.** TRF1^{Basic} phase separation exhibits reentrant response to 0 - 250 μM 8ds3ss, whereas TRF2^{Acidic} does not phase separate in any conditions (protein concentration was set to 20.1 μM).

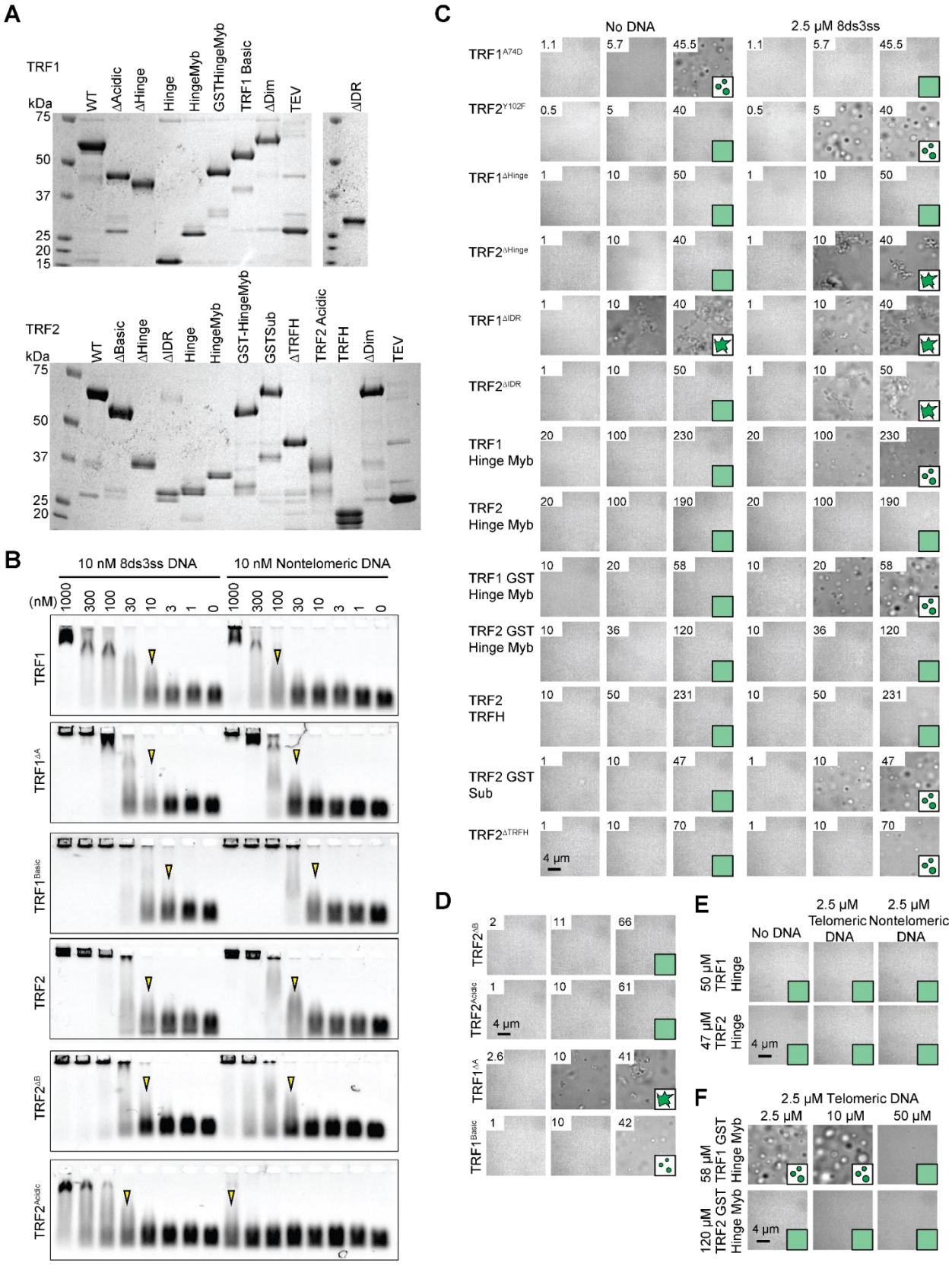
A
TRF1
kDa
WT
ΔAcidic
ΔHinge
Hinge
HingeMyb
GSTHingeMyb
TRF1 Basic
ΔDim
TEV
ΔIDR
75
50
37
25
20
15
TRF2
kDa
WT
ΔBasic
ΔHinge
ΔIDR
Hinge
HingeMyb
GST-HingeMyb
GSTSub
ΔTRFH
TRF2 Acidic
TRFH
ΔDim
TEV
75
50
37
25
20
B
10 nM 8ds3ss DNA
10 nM Nontelomeric DNA
(nM)
1000
300
100
30
10
3
1
0
TRF1
TRF1ΔA
TRF1Basic
TRF2
TRF2ΔB
TRF2Acidic
C
No DNA
2.5 μM 8ds3ss
TRF1A74D
TRF2Y102F
TRF1ΔHinge
TRF2ΔHinge
TRF1ΔIDR
TRF2ΔIDR
TRF1 Hinge Myb
TRF2 Hinge Myb
TRF1 GST Hinge Myb
TRF2 GST Hinge Myb
TRF2 TRFH
TRF2 GST Sub
TRF2ΔTRFH
4 μm
D
TRF2ΔB
TRF2Acidic
TRF1ΔA
TRF1Basic
4 μm
E
No DNA
2.5 μM Telomeric DNA
2.5 μM Nontelomeric DNA
50 μM TRF1 Hinge
47 μM TRF2 Hinge
4 μm
F
2.5 μM Telomeric DNA
2.5 μM
10 μM
50 μM
58 μM TRF1 GST Hinge Myb
120 μM TRF2 GST Hinge Myb
4 μm

Figure 6. Biochemical characterization and phase separation of the TRF1 and TRF2 mutants. A. Denaturing gel pictures of purified TRF1 and TRF2 mutants. The gels were stained with Coomassie. **B.** EMSA gels using Cy3 labeled 10 nM telomeric or nontelomeric DNA and decreasing protein concentration. Arrows indicate the minimum protein concentration for each condition with a noticeable shift in the DNA band. **C.** Brightfield images of native and truncated TRF2 and TRF1 constructs in the presence and absence of 2.5 µM 8ds3ss. **D.** In the absence of DNA, $TRF2^{\Delta B}$, $TRF2^{Acidic}$, and $TRF1^{\Delta A}$ do not form liquid droplets, whereas $TRF1^{Basic}$ forms small droplets only at high protein concentrations. **E.** Brightfield images of the hinge domain of TRF2 and TRF1 without DNA, and with 2.5 µM telomeric (8ds3ss) or non-telomeric DNA. **F.** Hinge-Myb of TRF1 and TRF2 were homodimerized with an N-terminal glutathione serin transferase (GST) tag. An increase in DNA concentration inhibits phase separation of TRF1 GST-Hinge-Myb similar to native TRF1 but does not affect TRF2 GST-Hinge-Myb. In C-D, the protein concentration is shown in µM at the top left corner. In C-F, the condensate state at the highest protein concentration is shown in the bottom right corner.

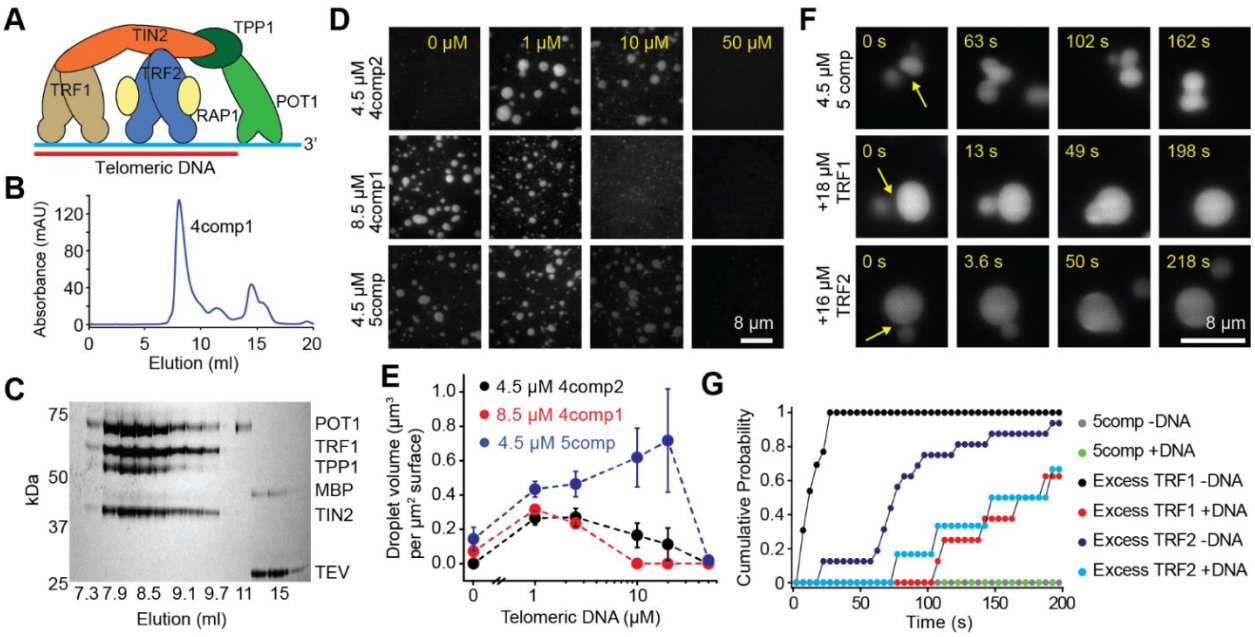

Figure 7. The shelterin complex phase separates in vitro**. A.** A schematic of the human shelterin complex. TRF1 and TRF2 are homodimers that bind to dsTEL, and POT1/TPP1 binds to ssTEL. TIN2 interconnects TRF1, TRF2, and TPP1, and RAP1 binds to TRF2. **B-C.** UV absorbance (B) and denaturing gel (C) show that 4comp1 elutes as a single complex from a gel filtration column. **D.** 4comp2 and 5comp exhibit reentrant responses to increasing DNA concentration similar to TRF2 droplets, while 4comp1 is inhibited by increasing DNA concentration similar to TRF1 droplets. **E.** The total volume of shelterin droplets settled per micron squared area on the coverslip under variable 8ds3ss concentration (mean ± SD, n = 20 with two technical replicates). **F.** In the presence of 2.5 μM 8ds3ss, 5comp droplets do not fuse on relevant time scales, whereas the addition of excess TRF1 or TRF2 reduces the fusion time. **G.** Cumulative probability of 5comp droplet fusion in the presence and absence of excess TRF1 or TRF2 after forming a contact at t = 0 s (n = 7, 4, 13, 16, 15 and 7 events from top to bottom; taken from two technical replicates).

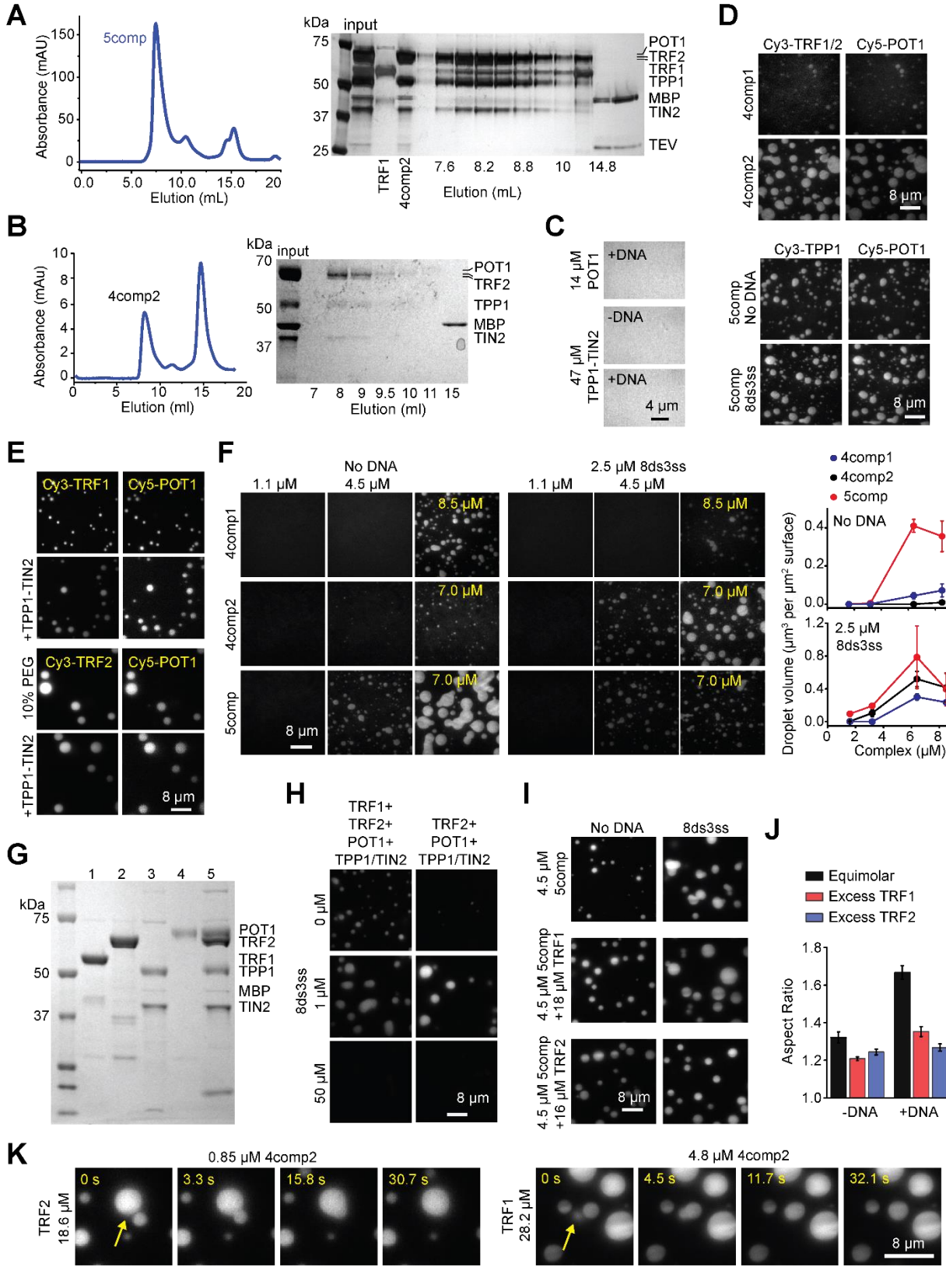
A
5comp
Absorbance (mAU)
Elution (mL)
kDa
input
POT1
TRF2
TRF1
TPP1
MBP
TIN2
TEV
TRF1
4comp2
Elution (mL)
B
4comp2
Absorbance (mAu)
Elution (ml)
kDa
input
POT1
TRF2
TPP1
MBP
TIN2
Elution (ml)
C
14 μM POT1
+DNA
-DNA
47 μM TPP1-TIN2
+DNA
4 μm
D
Cy3-TRF1/2
Cy5-POT1
4comp1
4comp2
8 μm
Cy3-TPP1
Cy5-POT1
5comp No DNA
5comp 8ds3ss
8 μm
E
Cy3-TRF1
Cy5-POT1
+TPP1-TIN2
10% PEG
Cy3-TRF2
Cy5-POT1
+TPP1-TIN2
8 μm
F
No DNA
2.5 μM 8ds3ss
1.1 μM
4.5 μM
8.5 μM
4comp1
4comp2
5comp
7.0 μM
8 μm
4comp1
4comp2
5comp
No DNA
2.5 μM 8ds3ss
Droplet volume (μm³ per μm² surface)
Complex (μM)
G
1 2 3 4 5
kDa
75
50
37
POT1
TRF2
TRF1
TPP1
MBP
TIN2
H
TRF1+ TRF2+ POT1+ TPP1/TIN2
TRF2+ POT1+ TPP1/TIN2
8ds3ss
0 μM
1 μM
50 μM
8 μm
I
No DNA
8ds3ss
4.5 μM 5comp
4.5 μM 5comp +18 μM TRF1
4.5 μM 5comp +16 μM TRF2
8 μm
J
Equimolar
Excess TRF1
Excess TRF2
Aspect Ratio
-DNA
+DNA
K
0.85 μM 4comp2
TRF2
18.6 μM
0 s
3.3 s
15.8 s
30.7 s
4.8 μM 4comp2
TRF1
28.2 μM
0 s
4.5 s
11.7 s
32.1 s
8 μm

Figure 8. Phase separation of shelterin is driven by TRF1 and TRF2. A. UV absorbance (left) and silver-stained denaturing gel (right) show that 5comp elutes as a single complex from a gel filtration column. **B.** UV absorbance (left) and silver-stained denaturing gel (right) show that 4comp2 elutes as a single complex from a gel filtration column. **C.** POT1 and coexpressed TPP1-TIN2 do not form higher-order structures. **D.** (Left) Equimolar POT1 partitions into droplets formed by 8.5 µM 4comp1 or 7.0 µM 4comp2 in the presence of 2.5 µM 8ds3ss. (Right) Equimolar TPP1 and POT1 partition into droplets formed by 5.6 µM 5comp in the absence or presence of 2.5 µM 8ds3ss. **E.** 0.5 µM Cy5-POT1 partitions into droplets formed by 15.7 µM Cy3-TRF1 or 16.6 µM Cy3-TRF2 in the absence of DNA and with or without the addition of 7.8 µM coexpressed TPP1 and TIN2. 10% PEG-8000 was added to trigger droplet formation of TRF2 in the absence of DNA. **F.** (Left) 4comp1 and 5comp complexes phase separate without DNA, whereas 4comp2 shelterin phase separates most strongly with telomeric DNA. (Right) The total volume of shelterin condensates settled per micron squared area on the coverslip as a function of complex concentration (mean ± SD, n = 20 with two technical replicates). **G.** A denaturing gel of individual shelterin components as well as 4comp2 purified from insect cells (Lane 1: TRF1, 2: TRF2, 3: coexpressed TPP1 and TIN2, 4: POT1, 5: 4comp2). **H.** Mixing of 4.5 µM TRF2, TRF1, TPP1-TIN2, and POT1 separately results in phase separation with similar DNA dependence as the co-purified 4comp2 or 5comp complexes. **I.** While the equimolar 5comp complex forms nonspherical condensates in the presence of 8ds3ss, adding excess TRF1 or TRF2 results in the formation of spherical droplets. **J.** Aspect ratios of 5comp droplets in the presence and absence of excess 18.0 µM TRF1 or 16.4 µM TRF2 with 2.5 µM 8ds3ss (mean ± SE; from left to right, n = 648, 1383, 1389, 611, 832, and 1159 droplets). **K.** 4comp2 droplets formed with 2.5 µM 8ds3ss fuse within 30 s in the presence of excess TRF1 or TRF2.

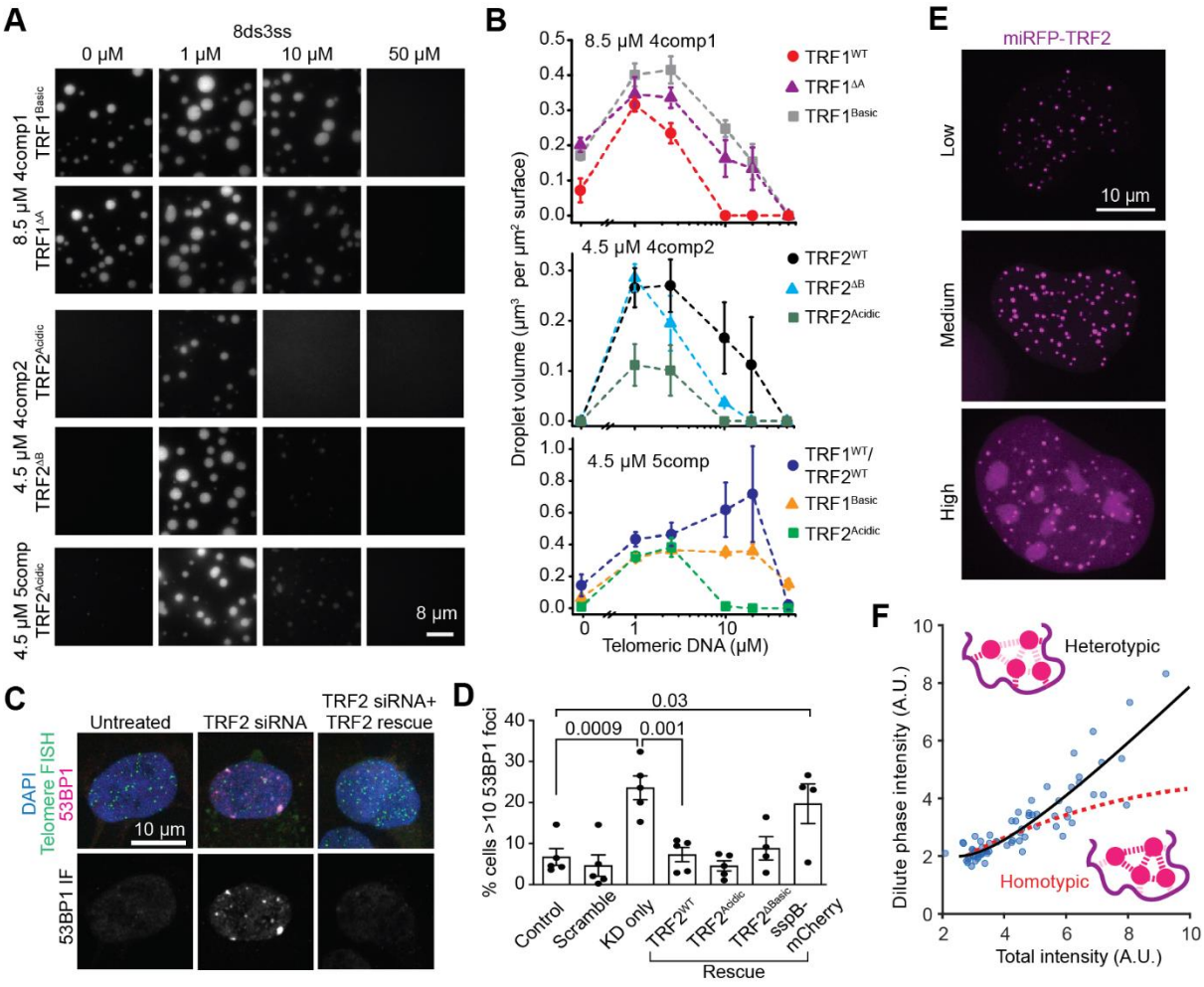

Figure 9. Telomeric condensates exhibit quantitative signatures consistent with multicomponent phase-separated liquids both in vitro **and in living cells. A.** Example images show phase separation of 4comp1, 4comp2, and 5comp assembled using N-terminal swap or truncation mutants of TRF1 and TRF2. **B.** The total volume of shelterin droplets assembled with native or mutant TRF1 and TRF2 settled per micron squared area on the coverslip under different 8ds3ss concentrations (mean ± SD, n = 20 with two technical replicates). **C.** 53BP1 staining (magenta) of hTERT-RPE1 cells that are treated or untreated with TRF2 siRNA. Telomeres are stained with a telomeric DNA FISH probe (green). Nuclei are labeled with DAPI (blue). **D.** The percentage of hTERT-RPE1 cells with more than 10 53BP1 foci per nucleus under knockdown and rescue conditions. Error bars represent SEM of five biological replicates for all conditions except for TRF2$^{\Delta B}$ and sspB-mCherry (four biological replicates). n > 1000 cells analyzed for all conditions. P-values were calculated by one-way ANOVA with multiple comparisons. **E.** Overexpression of miRFP-TRF2 leads to an increased dilute phase (nucleoplasmic) partitioning in U2OS cells. **F.** The dilute phase intensity increases nonlinearly as a function of the total intensity of the miRFP-TRF2 signal in U2OS cells (one biological replicate of 72 cells). The data fit to a nonlinear heterotypically stabilized model (black solid curve) but not to homotypic interactions (red dashed curve). The 'homotypic' curve is not a flat line due to the presence of endogenous protein (see [4]).

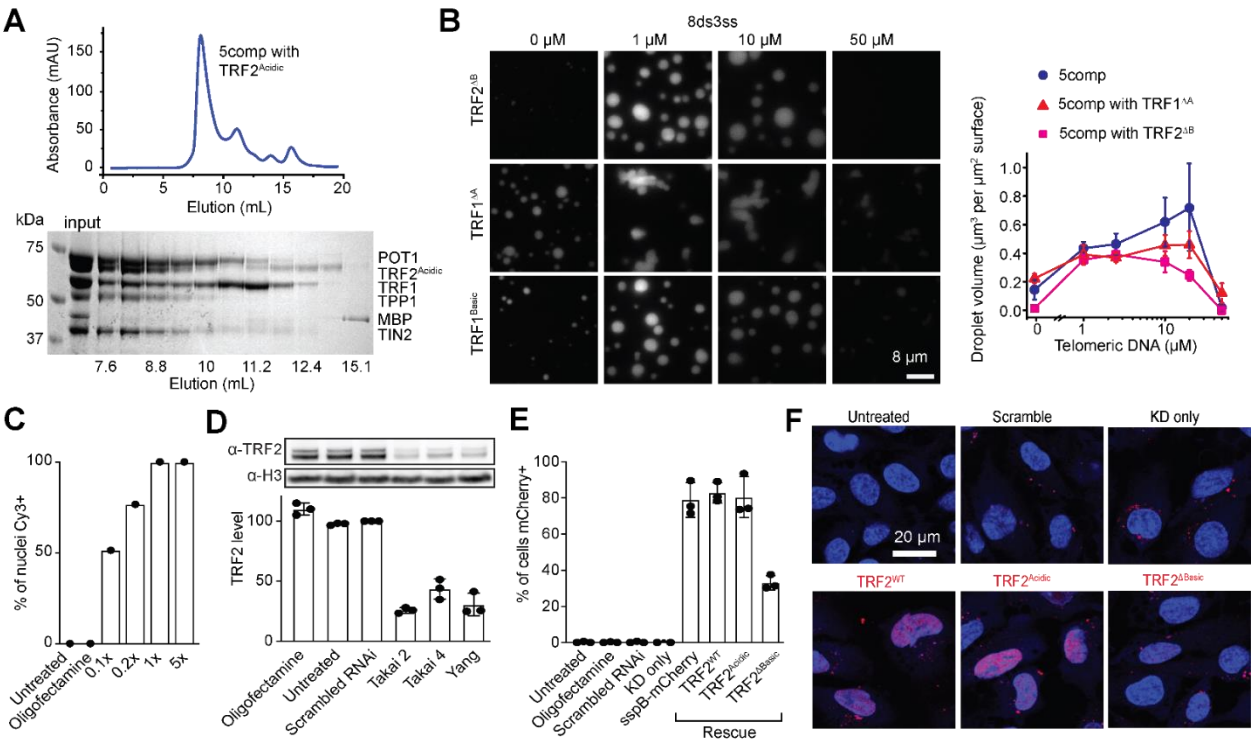

Figure 10. TRF2 knockdown efficiency in cells and in vitro **phase separation of shelterin complexes assembled with TRF1 and TRF2 mutants. A.** In vitro reconstitution of 5comp with TRF2^{Acidic} substituted for TRF2^{WT}. UV absorbance (left) and denaturing gel (right) show that 5comp with TRF2^{Acidic} elutes primarily as a single complex from a gel filtration column. **B.** (Left) Replacing TRF1 with TRF1$^{\Delta A}$ or TRF1^{Basic} and replacing TRF2 with TRF2$^{\Delta B}$ in 5comp shelterin does not strongly affect DNA-dependent condensation of shelterin. The complex concentration was set to 4.5 μM. (Right) The total volume of condensates settled per micron squared area on the coverslip as a function of complex concentration (mean ±SD, n = 20 with two technical replicates). **C.** Percentage of nuclei expressing Cy3-labeled control siRNA to estimate siRNA transfection efficiency in hTERT-RPE1 cells. **D.** (Top) Western blot comparing endogenous and knockdown TRF2 levels (anti-TRF2) and loading control (anti-H3). (Bottom) Quantification of TRF2 level for each condition in 3 biological replicates. **E.** Percentage of cells expressing mCherry constructs, indicating infection efficiency. **F.** Example images show the mCherry signal (red) in hTERT-RPE1 cells under different conditions. Nuclei of cells are labeled with Hoechst (blue).

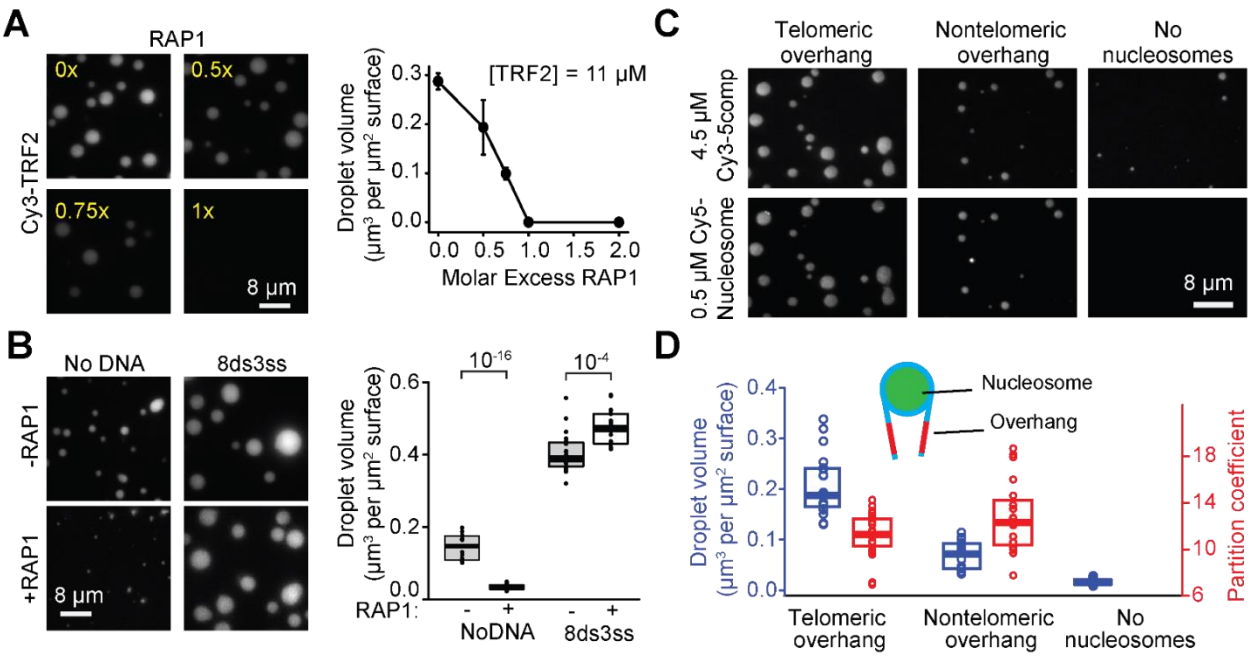

Figure 11. Telomere-associated proteins modulate phase separation of shelterin droplets. **A.** (Left) Increasing the molar ratio of RAP1 inhibits phase separation of TRF2 droplets. Droplets were formed in the presence of 2.5 μM 8ds3ss DNA. (Right) The total volume of TRF2 condensates settled per micron squared area on the coverslip as a function of RAP1 concentration (mean ± SD, n = 20 with two technical replicates). **B.** (Left) 5comp droplets formed with or without equimolar RAP1 and in the presence or absence of 2.5 μM 8ds3ss DNA. Complex concentration was set at 4.5 μM. (Right) The total volume of shelterin condensates settled per micron squared area on the coverslip in the absence or presence of RAP1. The center and edges of the box represent the median with the first and third quartile (n = 20 with two technical replicates). The p-values were calculated from a two-tailed t-test. **C.** Example images show phase separation of 4.5 μM 5comp in the presence and absence of nucleosomes wrapped with telomeric or nontelomeric DNA. **D.** Volume of droplets settled per micron squared area and partition coefficient of nucleosomes into 5comp droplets. The center and edges of the box represent the median with the first and third quartiles (n = 20 droplets with two technical replicates).

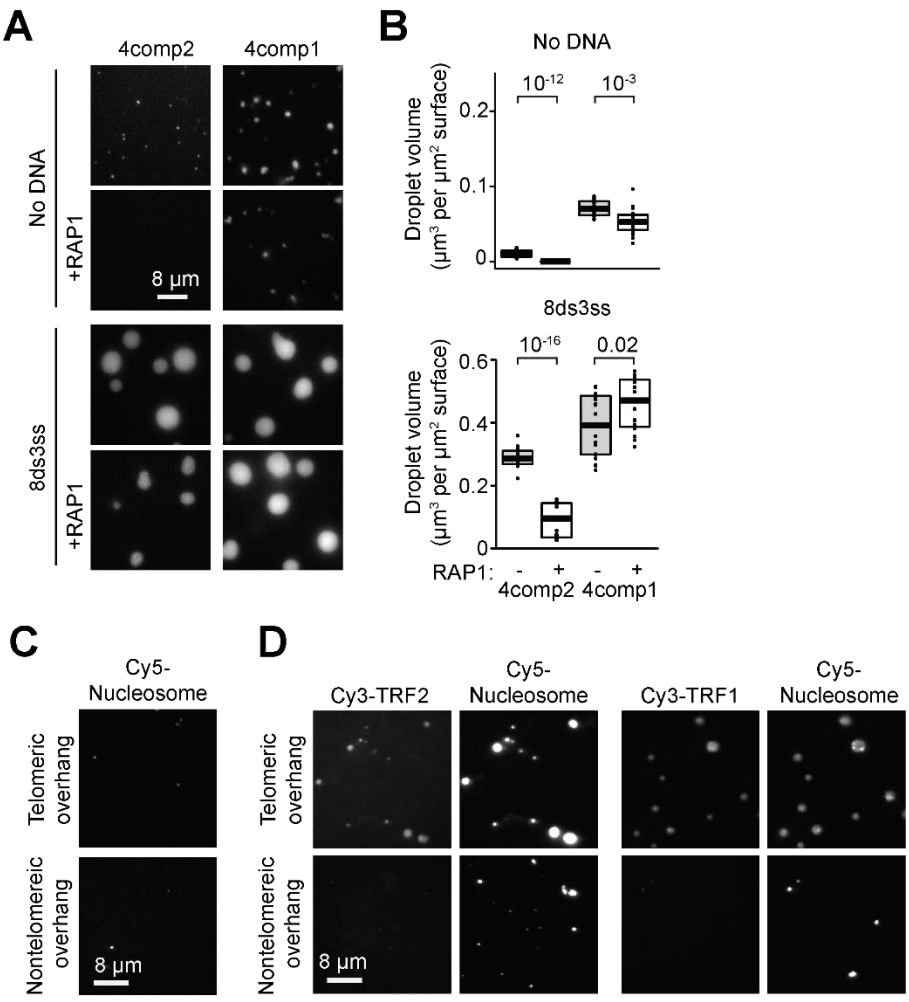

Figure 12. Phase separation of shelterin in the presence of RAP1 and nucleosomes. A. Shelterin droplets formed with or without equimolar RAP1 and in the presence or absence of 2.5 µM 8ds3ss DNA. Complex concentration was set at 4.5 µM for 4comp2 and 8.5 µM for 4comp1. **B.** The total volume of shelterin condensates settled per micron squared area on the coverslip in the absence or presence of RAP1. The center and edges of the box represent the median with the first and third quartile (n = 20 with two technical replicates). The p values were calculated from a two-tailed t-test. **C.** Without shelterin components, 0.5 µM mononucleosomes with telomeric or non-telomeric DNA overhangs do not phase separate in solution. **D.** 0.5 µM mononucleosomes with telomeric DNA overhangs (TTAGGGTTAGGG on each overhang) trigger droplet formation of TRF1 and TRF2 (at 42 and 37 µM, respectively), while 0.5 µM mononucleosomes wrapped with non-telomeric DNA do not promote phase separation of TRF2 or TRF1.

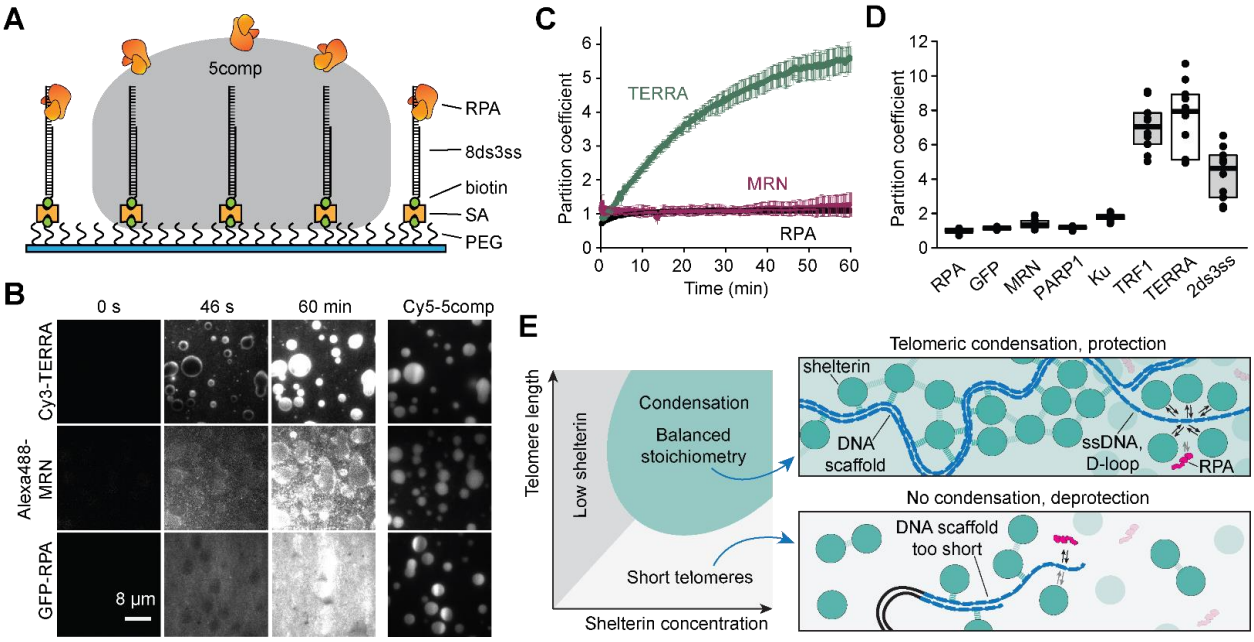

Figure 13. Shelterin droplets selectively recruit telomere-associated factors. A. 5comp droplets are settled onto PEG surfaces decorated with 8ds3ss. (PEG: polyethylene glycol, SA: streptavidin; not to scale). **B.** 100 nM Cy3-TERRA, 15 nM Alexa488-MRN complex, or 100 nM GFP-RPA are introduced to 7.6 µM Cy5-5comp droplets. TERRA partitions strongly into the droplets, while MRN and GFP-RPA are initially excluded from the droplets and uniformly distributed after 60 min incubation. **C.** Partitioning of 100 nM Cy3-TERRA, 15 nM Alexa488-MRN, or 100 nM GFP-RPA into 7.6 µM 5comp droplets over time (mean ± SD, n = 3 droplets per condition, one replicate). **D.** Partition coefficients of DDR proteins and telomere-associated factors in 7.6 µM 5comp droplets after 60 min incubation. The center and edges of the box represent the median with the first and third quartile (n = 10 droplets per condition, two technical replicates). **E.** (Left) Multicomponent phase diagram of telomere condensation with balanced stoichiometry. No condensation results at low shelterin concentrations and/or short telomeres. (Top) Telomere condensation, formed by both heterotypic (dark dashes) and homotypic (light dashes) interactions, selectively recruit telomere-associated factors while acting as a diffusion barrier against other components that target telomeric DNA, like RPA. The enrichment of shelterin, and thus POT1, outcompetes RPA binding to ssTEL. (Bottom) Shortened telomere scaffold cannot recruit enough shelterin to form a condensate, which could fail to protect the ssTEL overhang against RPA binding.

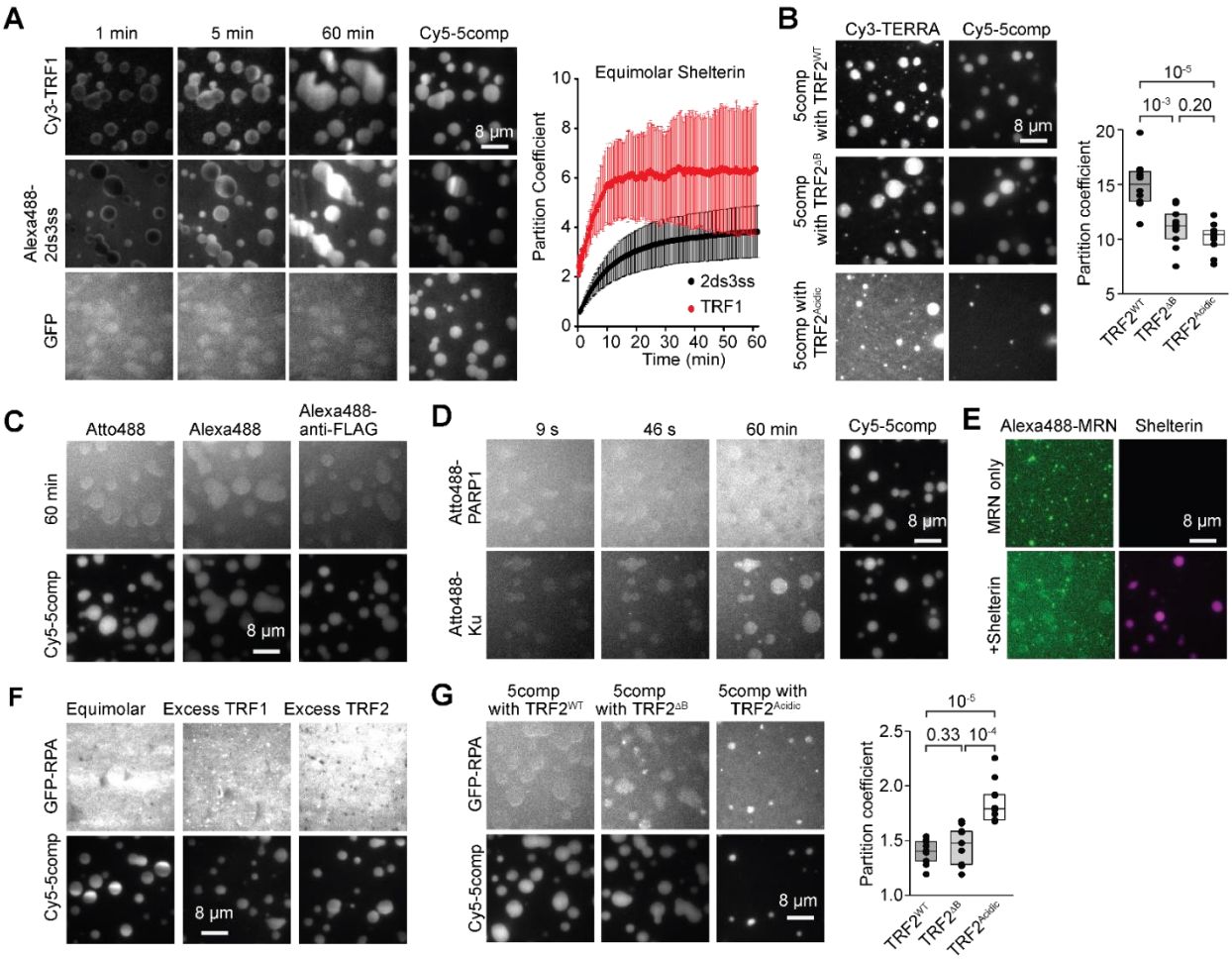

Figure 14. Partitioning of telomeric and non-telomeric factors into shelterin droplets settled on telomeric DNA. A. (Left) TRF1 (1.0 μM) and 2ds3ss DNA (2.5 μM) partition into 5comp shelterin droplets (7.6 μM) settled on surface-immobilized 8ds3ss DNA, while GFP (10.0 μM) is distributed uniformly over time. (Right) Partitioning of 2.5 μM 2ds3ss DNA or 1.0 μM TRF1 into equimolar 7.6 μM 5comp droplets settled on surface-immobilized 8ds3ss DNA over 1 h (mean ± SD, *n* = 3 droplets). **B.** (Left) 100 nM Cy3-TERRA is introduced to 5comp droplets assembled with Cy5-labeled TRF2^{WT} (7.6 μM complex; top), TRF2$^{\Delta B}$ (9.5 μM complex; middle), or TRF2^{Acidic} (9.5 μM complex; bottom) (mean ± SD, *n* = 3 droplets). (Right) The partition coefficient of TERRA in 5comp droplets after 60 min incubation. The center and edges of the box represent the median with the first and third quartile (n = 10 droplets per condition). The p-values were calculated from a two-tailed t-test. **C.** 100 nM Atto488 and 100 nM Alexa488 dye controls, and 100 nM Alexa488-anti-FLAG are distributed uniformly inside and outside of 7.6 μM 5comp shelterin droplets. **D.** 100 nM Atto488-PARP1 or Atto488-Ku are introduced to 7.6 μM 5comp droplets settled on surface-immobilized 8ds3ss DNA. After 60 minutes, PARP1 and Ku have partition coefficients of 1.2 and 1.8, respectively. **E.** 15 nM Alexa488-MRN forms small puncta on surface-immobilized 8ds3ss DNA in the absence (top) or presence (bottom) of 7.6 μM 5comp droplets. **F.** 100 nM GFP-RPA is introduced to Cy5-5comp droplets formed in

the presence or absence of excess TRF1 or TRF2 (7.6 μM equimolar shelterin (Top); 3.8 μM shelterin with 13.1 μM excess TRF1 (Middle); and 5.0 μM shelterin with 16.7 μM excess TRF2 (Bottom)). GFP-RPA is uniformly distributed after 60 min incubation. **G.** (Left) 100 nM GFP-RPA is introduced to 5comp droplets assembled with Cy5-labeled $TRF2^{WT}$ (7.6 μM complex; top), $TRF2^{\Delta B}$ (9.5 μM complex; middle), or $TRF2^{Acidic}$ (9.5 μM complex; bottom) (mean ± SD, *n* = 3 droplets). (Right) The partition coefficient of RPA into 5comp droplets after 60 min incubation. The center and edges of the box represent the median with the first and third quartile (n = 10 droplets per condition). The p-values were calculated from a two-tailed t-test.

Name	Sequence
2ds0ss	5′ GACCATGC[TTAGGG] TTATCATACAA 3′ CTGGTACG[AATCCC]$_2$ AATAGTATGTT
3ds3ss	5′ GACCATGC[TTAGGG] TTATCATACAAGTTAGGGTTAGGGTTAGGG 3′ CTGGTACG[AATCCC]$_3$ AATAGTATGTT
8ds0ss	5′ GACCATGC[TTAGGG] TTATCATACAA 3′ CTGGTACG[AATCCC]$_8$ AATAGTATGTT
8ds3ss	5′ GACCATGC[TTAGGG] TTATCATACAAGTTAGGGTTAGGGTTAGGG 3′ CTGGTACG[AATCCC]$_8$ AATAGTATGTT
13ds0ss	5′ GACCATGC[TTAGGG] TTATCATACAA 3′ CTGGTACG[AATCCC]$_{13}$ AATAGTATGTT
39ds0ss	5′ GACCATGC[TTAGGG] TTATCATACAA 3′ CTGGTACG[AATCCC]$_{39}$ AATAGTATGTT
Non-telomeric DNA	5′ CGTTCCGATGTCTTCAATCTGCAGGAGCTAGTCACGAACGAAGTACAAGGTTCTGTCCCAG 3′ GCAAGGCTACAGAAGTTAGACGTCCTCGATCAGTGCTTGCTTCATGTTCCAAGACAGGGTC TGTAAGTAGCTTCACCAGGTAGATG ACATTC
Bio-8ds3ss	5′ GACCATGC[TTAGGG] TTATCATACAAGTTAGGGTTAGGGTTAGGG 3′Bio-CTGGTACG[AATCCC]$_8$ AATAGTATGTT
601-Widom	5′ GAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACG 3′ CTCTTAGGGCCACGGCTCCGGCGAGTTAACCAGCATCTGTCGAGATCGTGGCGAATTTGC CACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGC GTGCATGCGCGACAGGGGGCGCAAAATTGGCGGTTCCCCTAATGAGGGATCAGAGGTCCG ACGTGTCAGATATATACATC TGCACAGTCTATATATGTAG
2x-Tel-Widom	5′ TCGAATTCTTAGGGTTAGGGTTACCCTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGT 3′ AGCTTAAGAATCCCAATCCCAATGGGACCTCTTAGGGCCACGGCTCCGGCGAGTTAACCA CGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGC GCATCTGTCGAGATCGTGGCGAATTTGCGTGCATGCGCGACAGGGGGCGCAAAATTGGCG GTTCCGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATACATCCTGTGCTTAGGG CAAGGCCTAATGAGGGATCAGAGGTCCGTGCACAGTCTATATATGTAGGACACGAATCCC TTAGGGTTAGGATCCAG AATCCCAATCCTAGGTC
TERRA RNA	5′[UUAGGG]$_{10}$

Table 1: The list of telomeric and non-telomeric DNA and RNA constructs (bio: biotin).

Movies Related to Part 1

Movie 1. Telomeres rarely coalesce due to their suppressed diffusivity in living cells. Time-lapse movie showing that telomeres marked by GFP-TRF1 in HeLa cells are mostly stationary due to their subdiffusive motion.

Movie 2. Specific local activation protocol using the Corelet system allows observation of liquid-like telomere coalescence in living cells. (Left) The shrinking of FUS_N Corelet droplets pull in FUS_N-miRFP-TRF1 marked telomeres to coalesce into a single telomere condensate. (Right) Telomeres relax back to more distal positions after detaching from FUS_N droplets.

Movie 3. Telomere coalescence is not due to FUS_N-FUS_N interactions as iLID fused to TRF1 demonstrates liquid-like telomere merging events. iLID-miRFP-TRF1 acts as a seed only when it is bound to FUS_N-mCherry-sspB upon light activation. The FUS_N droplet shrinks and pulls in the telomeres to merge upon light deactivation once FUS_N-mCherry-sspB is completely dissolved.

Movie 4. The fusion of TRF2, TRF1, and 5comp droplets in the presence of telomeric DNA. TRF2 and TRF1 droplets containing DNA fuse within 30 seconds, on average, while 5-component shelterin droplets containing DNA adhere and relax but do not become spherical over several hundred seconds. Droplets were imaged with a TIRF microscope at 4.3 frames/s. Droplets were formed with 2.5 μM 8ds3ss telomeric DNA using either 22 μM TRF2, 44 μM TRF1, or 4.5 μM 5comp.

Movie 5. RPA and TERRA diffuse into 5-comp droplets settled onto a DNA-coated slide. Droplets formed from 5.3 μM 5comp in physiological salt were settled onto a slide coated with 8ds3ss telomeric DNA for 10 min before 100 μM GFP-RPA and 100 μM 60nt-TERRA were flowed onto the slide. RPA is distributed uniformly across the surface over several minutes, while TERRA partitions strongly into the shelterin droplets. TIRF imaging began 55 s after the introduction of RPA and TERRA to the sample. Time-lapse images were acquired for 200 ms every 10 s.

Methods Related to Part 1

Cell Culture
Recombinant proteins were purified from Sf9-ESF *S. frugiperda* insect cells (RRID:CVCL_0549; female) grown at 27 °C in ESF 921 Insect Cell Culture Medium (Expression Systems, NC903611) supplemented with 1% fetal bovine serum (Corning, 35-010-CV) and 1% antibiotic-antimycotic (Thermo Fisher Scientific, 15240062). All human cell lines were incubated in and grown at 37°C with 5% CO_2. U2OS cells were obtained from the ATCC and hTERT-RPE1 cells (p53 -/-, Rb -/-) were cultured in DMEM (GIBCO, 11995065) with 10% FBS (Atlanta Biological, S11150H) and 1% streptomycin and penicillin (GIBCO, 15140122), grown at 37°C with 5% CO_2. The HeLa RMCE GFP-TRF1 cell line was cultured in DMEM with 10% FBS (Atlanta Biological, S11150H), 1% streptomycin and penicillin, and 1 µg/ml puromycin (Sigma, P7255). All cell lines were authenticated via ATCC's STR profiling.

Shelterin protein purification
Constructs for expressing individual components of the human shelterin complex were tagged with an N terminal ZZ affinity tag, TEV cleavage site, and YBBR labeling site and cloned into a Baculovirus vector. A construct expressing both TPP1 and TIN2 (TIN2 did not express on its own or without a solubility tag) and constructs expressing four- or five-component shelterin were cloned using a BigBac vector as described [119]. For the four-component shelterin BigBac construct, POT1 was given an N terminal YBBR tag, POT1 and TRF2 were each given an N terminal ZZ affinity tag and a TEV cleavage site, and TIN2 and TPP1 were each given an N-terminal His-MBP affinity tag and a TEV cleavage site. For a full list of constructed plasmids, see the Key Resources Table. Protein was purified from insect cells as previously described [119]. Briefly, plasmids containing genes of interest were transformed into DH10Bac competent cells (Berkeley MacroLab), and Bacmid DNA was purified using ZymoPURE miniprep buffers (Zymo Research, D4210) and ethanol precipitation.

Insect cells were transfected using Fugene HD transfection reagent (Promega, E2311). The virus was amplified in progressively larger cultures. 1 mL of the P1 virus was used to infect 50 mL of Sf9 cells at 1 million cells/mL for 72 h. 10 mL of the P2 virus was used to infect 1 L of Sf9 cells at 1 million cells/mL and expression proceeded for 72 h. Cells expressing the protein of interest were harvested at 4,000 *g* for 10 min and resuspended in 50 mL lysis buffer (50 mM HEPES pH 7.4, 1 M NaCl, 1 mM PMSF, 1 mM DTT, and 1 tablet of protease inhibitor (Sigma, 4693132001)). Lysis was performed using 15 loose and 15 tight plunges of a Wheaton glass dounce. The lysate was clarified using a 45 min, 360,000 *g* spin in a Ti70 rotor. The supernatant was incubated with 1 mL IgG beads (IgG Sepharose 6 Fast Flow, GE Healthcare, 17096902) for ZZ-tagged TRF1, TRF2, and POT1 constructs or 1 mL amylose beads (New England BioLabs, E8021S) for co-expressed TPP1 and TIN2 and shelterin constructs for 1 h. Beads were washed with 40 mL of labelling buffer (50 mM HEPES pH 7.4, 300 mM NaCl, 10 mM $MgCl_2$, 1 mM EGTA, 10% glycerol, 1 mM DTT). Beads were then collected and incubated with purified SFP protein (Addgene #75015)and a fluorescent dye functionalized with CoA (Lumidyne, custom synthesis) at room temperature for 30 min. Beads were washed in 40 mL labeling buffer, collected, and incubated with TEV protease (Berkeley Macrolab, Addgene #8827) for 1 h at room temperature to elute the protein. For shelterin protein preps, the protein was additionally incubated with 0.3 mL Ni-NTA beads (HisPur, Thermo Fisher Scientific, 88221) in 20 mM

imidazole to remove the His-MBP and TEV in solution. After 30 min of incubation at °C, the beads were pelleted and the unbound protein was collected from the supernatant.

For the TRF1 and TRF2 mutant proteins, all purification steps were carried out in 1M NaCl to prevent aggregation. After TEV cleavage, the mutant proteins were concentrated and resuspended to reduce NaCl concentration to 300 mM. Finally, the protein was concentrated using Amicon Ultra 30K concentrators, concentration was measured using Bradford reagent (Bio-Rad, 500-0006), and aliquots were snap-frozen in liquid nitrogen. Isoelectric points were calculated with ExPASy ProtParam.

4comp2 was purified using the BigBac system, where subunits were coexpressed from the same vector. 4comp1 was created by mixing known concentrations of purified TRF1, coexpressed TPP1 and TIN2, and POT1 on ice. 5comp consisted of purified 4comp2 mixed with purified TRF1. 4comp1 and 4comp2 were run through a Superdex 200 Increase 10/300 GL size exclusion column (Cytiva, 28-9909-44) to separate assembled complexes from individual proteins and subcomplexes. Mixing purified proteins on ice produced results comparable to coexpressing the components (Figure 8G-H).

DDR protein purification

GFP-RPA was purified as previously described [120]. GFP was expressed in Rosetta cells (Berkeley MacroLab) using the GFP plasmid (Addgene 54762). This culture was added to 1 L of LB media and grown for 3 h until OD_{600} reaches 0.7. Cells were induced with 0.2% L-arabinose and incubated for 4.5 h at 37 °C in a shaker. After harvesting cells at 4,785 *g* for 15 min in a JLA 8.1 rotor, 500 mL cell pellets were incubated with 40 mL lysis buffer (50 mM HEPES pH 7.4, 300 mM NaCl, 20 mM imidazole, 1 mM PMSF, 1 mM DTT, and 1 tablet of protease inhibitor (Sigma, 11836170001)). Cells were lysed with a sonicator and spun in a Ti70 rotor at 117,734 *g* for 30 min. The supernatant was incubated with 2 mL of washed Ni-NTA beads (HisPur, Thermo Scientific, 88221) for 1 h at 4 °C. Beads were collected in a Bio-Rad column and washed in lysis buffer. Protein was cleaved off the beads with TEV protease at room temperature for 1 h and concentrated in an Amicon Ultra 10K concentrator. Protein concentration was measured using Bradford reagent (Bio-Rad, 500-0006). Protein was aliquoted and snap-frozen in 10% glycerol.

Mre11-Rad50-Nbs1 (MRN) complex and Ku were purified from Sf21 insect cells as previously described [101], with the addition of a 3x FLAG tag on the C-terminus of Mre11 and a 3x HA tag on the C-terminus of Ku80. PARP-1 was purified from Rosetta *Escherichia coli* cells as previously described [121, 122], with the addition of an N-terminal His-SUMO-HA tag. To label MRN, anti-FLAG (Sigma-Aldrich, F1804) was labeled using an Alexa Fluor 488 antibody labeling kit (Thermo Fisher Scientific, A20181). Alexa488 anti-FLAG at 3x molar excess was then incubated with the MRN complex on ice for 10 minutes. Ku and PARP-1 were labeled by incubating 2 hr at room temperature with 5-fold molar excess maleimide-coupled Atto488 dye (Sigma-Aldrich, 28562). Free dye was removed using a 40kDa Zeba spin desalting column (Thermo Fisher Scientific, 87766).

Nucleosome preparation

Histones were expressed from pET-H2A, pET-H2B, pET-H3, and pET-H4 constructs [123] in

BL21(DE3)pLysS cells (Sigma-Aldrich, 69451) and purified according to previously described procedures [124]. Inclusion bodies were solubilized in DMSO and unfolding buffer (7M Guanidine HCl, 20 mM Tris-HCl, pH 7.5, 10 mM DTT) and purified using anion exchange chromatography (HiTrap SP HP, GE Life Sciences, 95056-076) in SAU buffer (7M deionized urea, 20 mM sodium acetate pH 5.2, 5 mM beta-mercaptoethanol (BME), 1 mM EDTA), eluting with a salt gradient from 0.2 to 0.6 M NaCl. Following dialysis of the peak fractions overnight in water + 2 mM BME, histones were buffer exchanged into unfolding buffer and concentrated. Octamers were assembled by mixing equimolar amounts of each histone and dialyzing overnight in refolding buffer (2M NaCl, 10 mM Tris pH 7.5, 1 mM EDTA, 5 mM BME). Next, octamers were purified in refolding buffer with a Superdex 10/300 GL column (GE Life Sciences, 17517501) using an Akta Pure chromatography system. Octamer formation in peak fractions was verified by Coomassie staining of SDS-PAGE gels. Nucleosome DNA sequences were amplified with Phusion polymerase (New England BioLabs, M0530L) from pGEM-3z/601 (Addgene plasmid #26656) [125] using the following primer pairs for the standard nucleosome: 5'-Cy5-CTGGAGAATCCCGGTGCCG-3' and 5'-ACAGGATGTATATATCTGACACG-3', and for the telomere-tagged nucleosome the primer pair 5'-Cy5-TCGAATTCTTAGGGTTAGGGTTACCCTGGAGAATCCCGGT-3' and 5'-CTGGATCCTAACCCTAACCCTAAGCACAGGATGTATATATCTGA-3' were used. Oligonucleotides were synthesized by IDT. PCR products were verified on SYBR-Safe (Thermo Fisher Scientific, S33102) stained gels and purified using the Genejet PCR purification kit (Thermo Scientific, K0701). Nucleosomal core particles (NCPs) were then reconstituted by mixing the DNA and octamer at a molar ratio of 1.1:1 and slowly dialyzing into low salt, according to the procedure of Chua et al. [126] except that the dialysis was stopped at 0.1 M KCl. NCPs were purified on a HiTrap DEAE-FF (Cytiva, 17515401) column, first binding to the column in TCS buffer (20 mM Tris, pH 7.5, 1 mM EDTA, 1 mM DTT) and then eluting with an increasing gradient of KCl in TES buffer (10 mM Tris, pH 7.5, 0.5 mM EDTA). Peak fractions were examined on 0.5% agarose gels for Cy5 fluorescence (Amersham Typhoon; GE Life Sciences, 29238583). Fractions containing NCPs were pooled and dialyzed overnight in TCS buffer, then concentrated using Millipore centrifugal filters.

Formation and labeling of DNA substrates

For all DNA substrates except 39ds0ss, ssDNA sequences were ordered from IDT. Solutions of equimolar complementary sequences suspended in annealing buffer (10 mM Tris pH 7.5, 50 mM LiCl) were mixed and incubated in a hot plate at 95 °C for 5 min. The sample was then removed from the hot plate and allowed to cool to room temperature over 2 h. Comparing the molecular weights of the ssDNA oligos with that of the annealed dsDNA on an agarose gel confirms that annealing efficiency is high (Figure 4C). The 39ds0ss substrate was created by PCR amplifying a region of telomeric repeats from a plasmid, and the purity and length of the resulting DNA was verified on a 0.8% agarose gel. DNA was Cy3 labeled using a Label IT kit (Mirus Bio, MIR 3600).

Imaging of condensates

Slides were incubated with wash buffer (50 mM HEPES pH 7.4, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM EGTA, 1 mM DTT, 1% pluronic) for 5 min. Samples were settled onto the coverslip for 25 min before imaging. I confirmed that all of the condensates were settled to the surface within 25 min, as I did not observe an increase in the number of condensates per viewing area of the

coverslip, and I could not detect freely diffusing condensates in the flow chamber after 25 min. Different buffers were used for phosphatase-treated (50 mM HEPES pH 7.4, 150 mM NaCl, 0.5 mM MnSO4, 1 mM DTT, 1% pluronic) and kinase treated (20 mM HEPES pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 100 μM ATP, 1 mM DTT, 1% pluronic) samples. Imaging was performed using a Nikon Ti-E Eclipse microscope equipped with a 100X 1.49 N.A. Plan Apo oil immersion objective (Nikon). The samples were excited in near-TIRF using 488, 561, and 633 nm laser beams (Coherent). The emission signal was passed through a filter wheel and detected by Andor Ixon EMCCD Camera (512x512 pixels). The effective pixel size was 106 nm after 1X magnification and 160 nm after 1.5X magnification. Image processing is described in the "Quantification and Statistical Analysis" section.

For experiments involving DNA bound to the surface of the slide, chambers were incubated with 1 mg/mL Biotin-BSA (Sigma-Aldrich, 9048-46-8) for 2 min, incubated with 1 mg/mL streptavidin (Thermo Fisher Scientific, 434301) for 2 min, washed twice with 20 μL wash buffer, incubated with 1 μM biotinylated 8ds3ss DNA (IDT) for 2 min and washed twice with 20 μL wash buffer. Shelterin droplets were formed in a test tube, flowed into the chamber, and settled on the coverslip for 10 min. 5 μL solution containing the protein or DNA being tested was introduced to the chamber, and the sample was imaged using time-lapsing for 1 s every 10 s for 1 hr.

Cell culture for Corelet imaging

All DNA fragments of interest were PCR-amplified using Phusion High-Fidelity DNA Polymerase (New England BioLabs, M0530L). The hTRF1 and hTRF2 gene fragments were synthesized by IDT as gBlocks, with synonymous codon optimization to reduce repetitive DNA tracts. These fragments and point mutants were cloned into a linearized FM5 lentiviral vector. FM5 lentiviral vectors carried standardized linkers to insert the PCR fragments using the In-Fusion HD cloning kit (Takara Bio, 638910) [81]. Corelet constructs, unless otherwise noted, were cloned into the pHR lentiviral vector and confirmed by GENEWIZ Sanger sequencing.

Lentiviruses were generated by plating Lenti-X 293T cells (Takara Bio, 632180) into 6-well plates to reach ~70% confluence at the time of transfection. 24-36 hours after plating the Lenti-X cells, the transfer plasmid (1.50 μg), pCMVdR8.91 (1.33 μg), pMD2.G (0.17 μg) were transfected into the cells using FuGENE HD incubated in OptiMEM (modified from [66]). Transfer plasmids for the 53BP1 counting assay were transfected into Lenti-X cells with the helper plasmids VSVG and PSP with the Transit293 transfection reagent (Mirus, MIR 2700), following the protocol listed in Sanders et al. 2014 [81]. The supernatant-containing viruses were harvested 48 hours after transfection and filtered with a 0.45 μm filter (Pall Life Sciences), then used immediately or stored at -80°C. U2OS and hTERT-RPE1 cells plated at low (10-20%) confluency in 96-well glass-bottom plates (Cellvis) were transduced for 2-3 days before the washout of the virus, replacement with fresh media, and subsequent live-cell imaging experiments. Virus used for the formation of TRF1-mCh-sspB droplets away from telomeres at exceedingly high TRF1 concentrations was concentrated 10x using the Lenti-X Concentrator (Takara Bio, 631231), following the manufacturer's protocol.

Live cell imaging for Corelet experiments

Cells plated on 96-well glass-bottom plates were incubated at 37°C and 5% CO_2 by an Okolab microscope stage incubator with 96-well insert during all imaging experiments. Confocal microscopy was performed on a spinning disk (Yokogawa CSU-X1) confocal microscope with an Andor DU-897 EMCCD camera on a Nikon Eclipse Ti body using a 100x oil immersion Apo TIRF objective (NA 1.49). The following wavelength lasers were used to image the respective constructs: constructs with mGFP (488 nm), mCherry (561 nm), miRFP (640 nm). Fixed samples in the 53BP1 counting assay also used the 405 nm laser to detect nuclei stained with Hoechst (Thermo Fisher Scientific, H3570) or DAPI (Vectashield, H-2000-10).

Estimation of telomere component concentration in vivo
Both estimate that there are on the order of thousands of TRF2 dimers in a cell. Telomeric puncta are estimated to be 60-300 nm in diameter, and because virtually all TRF2 localize at telomeres [9], the local concentration of dimers within TRF2 puncta would be hundred-micromolar.

FRAP assays for Corelet experiments
FRAP experiments were performed on a Nikon A1 laser scanning confocal microscope using a 60x oil immersion objective. A single telomere marked by shelterin proteins of interest was photobleached with the 488 nm and 640 nm laser each at bleaching power of ~ 400 kW cm^{-2}. The cell was imaged every 2 s for 10 s of pre-bleach, and every 2 s post-bleach for 200 s. FRAP data were normalized by using the normalization method [127]. The first post-bleach point was set to zero.

Optogenetic telomere coalescence
Local activation was performed by using a Mightex Polygon digital micromirror device (DMD) to pattern blue light (488nm) stimulation from a Lumencor SpectraX light engine using Nikon Elements software. U2OS cells expressing the optogenetic telomere coalescence constructs FUS_N-miRFP-TRF1, NLS-GFP-iLID-Fe and FUS_N-mCherry-sspB were imaged using a specific local activation protocol, as follows. Pre-activation, imaging the mCherry (541 nm beam) and miRFP (640 nm beam) channels every 5 s for 15 s. Activation, wherein an elliptical region of interest (ROI) was used to locally activate two telomere foci to nucleate and grow FUS_N Corelet droplets using the 485 nm DMD laser every 5 s for 6 min. A second activation sequence used a smaller, circular ROI aimed at the junction between two FUS_N Corelet droplets every 5 s for 4 min to encourage them to fuse. Finally, the FUS_N droplet was deactivated for 10 min by only imaging the mCherry and miRFP channels every 5 s, which allows the droplets to dissolve and pull together any attached telomeres. The second set of telomere coalescence constructs (iLId-miRFP-TRF1) uses a similar local activation protocol but only a single circular activation ROI for 3 min and a longer deactivation sequence (15-30 min).

Corelet experiments
$TRF1^{WT}$, $TRF2^{WT}$, and TRF1 mutant Corelet experiments were imaged every 5 s and followed this protocol: 15 s pre-activation (561 and 640 nm lasers) and 10 min of activation for local activation (488, 561, and 640 nm lasers). Each locally activated telomere and region away from telomere was normalized by subtracting the background from the ROI and divided by the average intensity of all other telomeres in the same cell minus the background. The first and last frames of activation were quantified.

siRNA TRF2 knockdown

Endogenous TRF2 levels were knocked down using siRNAs made by IDT with sequences from Takai et al. 2003 [92] and Yang et al. 2015 [128]. siRNA transfection efficiency was estimated by transfecting a Cy3 labeled control siRNA (ThermoFisher Scientific, AM4621). To quantify transfection efficiency, three biological trials were transfected, fixed, stained with Hoechst or DAPI, and imaged. Nuclei were segmented with DAPI/Hoechst channel in FIJI and the intensity of Cy3 within each nucleus was recorded. Background intensity was subtracted from each, and the percent of cells with Cy3 signal at least 200 A.U. above background was plotted. The 1x condition was used for the 53BP1 counting assay. TRF2 knockdown efficiency was then validated by western blots (Figure 10B).

Western blot analysis

hTERT-RPE1 cells were plated on 6-well plates 24 hrs before siRNA treatment, and cells with or without siRNA treatment (siRNA #2 from Takai et al., 2003 [92] and a control scramble siRNA from Yang et al., 2015 [128] were grown for 48 hr before harvesting. Cell pellets were resuspended in 300ul RIPA buffer (BCA, 89901) with protease inhibitor (Sigma Aldrich, 4693132001). 1ul of benzonase (Sigma Aldrich, E1014-25KU) was added to each sample and left on ice for 30 min, each sample was spun down for an additional 30 min at 4°C, 30ul of lysate was resuspended in sample buffer (Thermo Fisher, NP0007), boiled at 100°C for 5 min with 15ul of the mix loaded for SDS-PAGE. Samples were run on a NuPAGE 4-12% Bis-Tris protein gel (Thermo Fisher, NP0322BOX) and transferred onto Trans-Blot Turbo Mini 0.2 um PVDF transfer pack (Bio-Rad, 1704156) for 30 min. Membranes were blocked for 2 hr with 5% NFDM in 1X TBST (Fisher Scientific, AAJ62938K2), and incubated in block with the anti-TRF2 antibody 1:2000 (Novus Biologicals, NB110-57130) and anti-Histone H3 antibody 1:2000 (Abcam, ab10799) for the loading control overnight at 4°C. Membranes were washed three times 5 min each with 1X TBST, incubated with either the Peroxidase AffiniPure Goat anti-mouse IgG 1:10,000 (Jackson ImmunoResearch, 115-035-062) or the Peroxidase AffiniPure Goat anti-rabbit IgG secondary antibodies 1:10,000 (JacksonImmunoResearch, 111-035-144) for 30 min at room temperature. Membranes were washed three times 5 min each with 1X TBST and developed using the SuperSignal West Pico PLUS Chemiluminescent Substrate (Thermo Fisher, 34577), following the manufacturer's protocol. To determine the knockdown efficiency, the background intensity was subtracted from each band intensity, then normalized relative to loading control and plotted as a ratio relative to scrambled RNAi (set at 100% for each trial).

53BP1 foci counting assay

The siRNA #2 from Takai et al., 2003 [92] and a control scramble siRNA from Yang et al., 2015 [128] were used for the 53BP1 counting assay that used an IF-FISH protocol adapted from both the de Lange lab's IF-FISH protocol and PNA bio's FISH protocol. hTERT-RPE1 cells plated on glass-bottom 96 well plate 24 hr before transfection and were transfected twice with Oligofectamine (Thermo Fisher, 12252011) according to the manufacturer's protocol (the second transfection was 24 hr after the first). Cells were transduced with 50-70ul of the rescue construct lentiviruses simultaneously with the siRNA treatment. Cells were then fixed with 4% paraformaldehyde for 5 min 48 hours after the first siRNA transfection, washed three times 5 min each with 1X TBST, and permeabilized for 15 min in 0.5% Triton X-100 buffer, incubated with block (10% goat serum and 0.1% Triton X-100 in 1X TBST) at room temperature for 1 hr, and incubated in block with anti-53BP1 antibody 1:50 (Novus Biologicals, NB100-305)

overnight at 4°C. Cells were then washed four times 5 min each with 1X TBST, incubated with Goat anti-Rabbit IgG secondary antibody conjugated to Alexa fluor 647 (Thermo Fisher, A-21245) for 2 hours, washed four times, and fixed again with 4% paraformaldehyde. After three washes of 5 min each with 1X TBST, cells were dehydrated in 70%, 85%, 100% cold ethanol for 5 min each, air-dried for 15 min, denatured for 10 min with the hybridization buffer at 80°C. The hybridization buffer contained 70% formamide, 0.5% blocking reagent (Millipore Sigma, 11096176001), 20mM Tris-HCl pH 7.5, and a FITC labeled C-rich telomere probe (PNA bio, F1009). The samples were then incubated in the dark at room temperature for 2 hr, washed twice with 70% formamide, 10mM Tris-HCl pH 7.5 for 15 min each, washed three times for 5 min each with 1X TBST, left to air dry before mounting in Vectashield Plus Antifade Mounting Medium with DAPI (H-2000-10). 2x2 tiled images were taken from 31 z-stacks of 0.2um spacing on a spinning disk (Yokogawa CSU-X1) confocal microscope. 3D Objects Counter [129] and 3D Multicoloc included in the 3D ImageJ suite [130] were used to detect the nuclei, 53BP1, telomeres, and the number of colocalizations per stack. 3D segmented data was parsed in Python 3.7.10 with a custom Python script that counts the number of 53BP1 foci per nucleus.

miRFP-TRF2 dilute phase vs. total concentration

U2OS cells were transduced with 50 µl of miRFP-TRF2 lentivirus to result in differential overexpression levels. Images were taken from 11 z-planes of 0.5 µm spacing. Total concentration was calculated by taking the average intensity of the miRFP-TRF2 signal in an entire segmented nucleus, and dilute phase calculated by taking the average intensity of miRFP-TRF2 in the nucleoplasm, with bright telomeres masked out. c_{dil} is measured as the 'background' concentration when a condensed phase is present. In a single-component phase separating system, c_{dil} will saturate at a single 'saturation concentration' (c_{sat}), while in a multicomponent phase separating system, c_{dil} may vary as a function of total system concentration.

In vitro droplet image processing

To calculate the saturation concentration (c_{sat}) of the proteins, droplets were identified using the Phansalkar function in Fiji (ImageJ 1.52p) with a 30-pixel radius and a minimum condensate size of 10 pixels. The volume of the condensates was estimated from 2D projections by taking the semi-principal axis in the z-plane as the geometric average of semi-principal axes in the XY plane. The total volume of the condensates settled per micron squared on the coverslip was quantified. Conditions that resulted in measurable condensate volumes were fit to linear regression in Origin. The x-intercept of the linear regression represents c_{sat}, the minimum protein concentration that results in condensate formation. c_{sat} for TRF2 in the presence of different DNA constructs (Figure 5D) was determined as the lowest protein concentration for which condensates are visible (Figure 4D). The aspect ratio was calculated using the Phansalkar function with a 30-pixel radius to detect particles and the Fit Ellipse function to determine the major axis length by the minor axis length in Fiji. Fusion times were calculated as the time between the last frame where two droplets appear separated (i.e. no overlap) and the first frame where the fused droplet appears spherical (aspect ratio ~ 1).

Image segmentation for time-lapse imaging of telomeres

All images were analyzed in Fiji (ImageJ 1.52p) and MATLAB 2019b (Mathworks). The first frame of each movie was used to calculate inter-telomere spacing. Briefly, nuclei were segmented using Otsu's method; telomeres were then segmented by filtering using an LoG

(Laplacian of Gaussians) kernel and applying a two standard deviation threshold. Average pairwise distance and nearest neighbors were then calculated (*pdist2*) based on the weighted centroids of all telomeres within each nucleus (extracted from the punctate mask *regionprops*). The local concentration of TRF2 at telomeres was estimated to be 400 μM from the measured radius of telomeres (~100 nm, [63, 131]) and estimated number of TRF2 at each telomere in cells (~1,000 on average) from immunoblotting [80] and superresolution imaging assays [63].

Mean Squared Displacement measurements

Time-lapse movies were taken of U2OS cells with GFP-TRF1 or FUS_N-miRFP-TRF1 overexpression and HeLa RMCE GFP-TRF1. Cells were plated 24 h before imaging on 96-well glass-bottom plates coated with fibronectin to reach high confluency. Each movie was 1 h long, imaging 1 s per frame.

Mean Squared Displacement analysis

To analyze telomere movement, images were registered to correct for whole-cell movement using StackReg plugin in Fiji. Then, Trackmate was used to track telomere movement with subpixel resolution using a Laplacian of Gaussian detector and object diameter of 500 nm. Trajectories of telomeres were then created using LAP tracking with maximum linking and gap-closing distances of 500 nm and zero-gap frames. Trajectories were only used if they spanned at least half the number of frames of the movie, then coordinates exported to MATLAB to calculate mean squared displacement.

Integrated intensity predictions and measurements

Telomeres were segmented using an LoG filter threshold method. Their respective total integrated intensity was calculated by summing over the intensity per pixel in the identified region. Since the integrated intensity should be directly proportional to the volume [132], the average integrated intensity of each telomere was calculated pre-coalescence (defined as all frames wherein two puncta were identified) and summed; the error was estimated and propagated by taking the standard error of the mean over pre-coalescence frames. The resulting summed integrated intensity and error bar were used as the independent prediction of the integrated intensity post-coalescence (defined as all frames wherein only one object was detected).

Statistical analysis

Statistics for the Corelet experiments and 53BP1 counting assay were performed using GraphPad PRISM version 9.1.0 software (GraphPad). Statistical significance (when reported) was calculated by one-way ANOVA with multiple comparisons or two-tailed t-test as noted in the figure legends. Number of replicates, size of n and precision measures (mean, median, ± SE and ± SD) are noted in the figure legends and captions.

Part 2: SARS-CoV-2 Nucleocapsid Protein forms Condensates with Viral Genomic RNA

Preface

The SARS-CoV-2 virus consists of a 30 kb single-stranded RNA genome packaged into a 100 nm diameter membrane enveloped virion. SARS-CoV-2 encodes for multiple proteins involved in viral assembly and propagation [46] and infects human cells by binding its spike (S) protein to the ACE2 receptor on host cells [133-135]. While the majority of current efforts to treat COVID-19 have focused on targeting this interaction [136, 137], not much work has been done to stop the proliferation of the virus in host cells following infection. Condensation of the viral genome into a virion is primarily driven by the N protein [33], which is the most abundant viral protein in infected cells [30, 138]. A large pool of free N protein is expressed early in infection [40], and only a small fraction is transferred into mature virions [138]. The N protein accumulates at the replication transcription complex (RTC) [28, 29] where it enhances replication and the transcription of viral RNA [31, 32]. The N protein also restructures viral genomic RNA into shell-shaped structures (~15 nm in diameter), which contain approximately 12 N proteins and 800 nucleotides of viral RNA [33, 34]. These vRNPs form asymmetric "beads on a string" structures which then bind to the viral M protein on the surface of the ERGIC to trigger the budding of the vRNP complex.

The mechanism by which N remodels the viral RNA and packages it into a viral particle is not well understood, but recent studies proposed that the replication and packaging of viruses involve LLPS [2, 41, 139]. Nucleic acids are highly involved in the formation of biomolecular condensates because they can scaffold multivalent interactions [4, 140, 141]. Coronaviruses are involved with phase-separated structures such as stress granules [40, 46] and replicate at dynamic clusters associated with the ERGIC, suggesting that phase separation may play a critical role in the replication and packaging of SARS-CoV-2.

Previous studies in HIV-1, negative-sense RNA viruses, and SARS-CoV showed that nucleocapsid proteins drive the formation of phase-separated condensates in the cytosol [42, 43, 142, 143]. The N protein of SARS-CoV-2 also contains many of the characteristic domain features common in phase separating proteins. It contains a well-conserved N-terminal domain (NTD) and a C-terminal domain (CTD, Figure 15A), and 40% of its primary sequence is predicted to be part of intrinsically disordered regions (IDRs, Figure 15B). The NTD (aa: 44-174) interacts nonspecifically with RNA and recognizes a nucleotide sequence in the 3' end of the viral genome [31]. The CTD (aa: 257-366) mediates dimerization [144], but the N protein can also self-associate into tetramers and higher oligomers [145, 146]. The NTD and CTD are separated by an IDR that contains a serine/arginine-rich (SR) motif, which has been associated with phase separation in other ribonucleoproteins [31, 147]. The N and C-terminal IDRs are less conserved but contain arginine- and lysine-rich disordered regions (Figure 15B), which may facilitate additional interactions with the negatively-charged RNA backbone [148] and drive biomolecular condensation of RNA [149, 150]. The N-terminal IDR contains a predicted prion-like domain (PLD, Figure 15A) that can potentially trigger protein demixing [31, 147, 151]. The C-terminal IDR of SARS-CoV N mediates binding to the M protein [152, 153]. The N protein is

highly positively-charged (+24 in pH 7.4) [44], and the CTD and disordered regions also interact with negatively-charged RNA and promote vRNP packaging [154]. The precise roles of these domains in the phase separation of N protein remain to be elucidated.

In this study, I purified the SARS-CoV-2 N protein from human embryonic kidney (HEK293) cells and observed that N protein forms biomolecular condensates with both homopolymeric and viral genomic RNA under physiological salt conditions in vitro. I also showed that the N protein forms liquid condensates in mammalian cells. Cross-linking mass spectrometry (CLMS) identified two regions flanking the CTD with interactions enriched within the condensed phase, and the deletion of one of these regions fully abrogated condensate formation in vitro. Together, my results indicate that the N protein phase separates with genomic RNA of SARS-CoV-2, which may play an important role in the packaging of new viral particles in host cells.

Results

The N protein phase separates with RNA

We first asked whether the N protein forms bimolecular condensates in the presence or absence of viral RNA in vitro. To address this, I expressed wild-type (WT) N protein in HEK293 cells and purified it in a high salt buffer (1 M NaCl) to eliminate the retention of RNA from human cells [155] (Figure 16A-B). Consistent with recent studies, purified N protein eluted from gel filtration as an oligomer in 300 mM NaCl [144] and had a high affinity for binding to various RNA substrates [156, 157] (Figure 16C-F). The protein was labeled with a fluorescent dye (LD655) at the C-terminal ybbR tag and introduced to a flow chamber in the presence or absence of RNA. Phase separation was monitored by the settling of N or N-RNA condensates onto the coverslip within 25 min under highly inclined and laminated optical sheet (HiLO) excitation (see Methods for details). In the absence of RNA, N protein did not form condensates in physiological salt (150 mM NaCl) (Figure 15C). Similarly, 2-kb long polyC RNA homopolymer did not form any condensates in the absence of N protein (Figure 17A-B). Mixing of 50 ng/µL polyC RNA and the N protein resulted in the formation of condensates in 150 mM NaCl (Figure 15C). LD655-N and Cy3-polyC colocalized well in the droplets, suggesting that phase separation of N protein is mediated by RNA (Figure 17C). The analysis of the condensates settled on the coverslip revealed a saturation concentration (c_{sat}) of 3.3 ± 1.3 µM for N protein in the presence of 50 ng/µL polyC RNA (± s.e., Figure 15C), consistent with the abundance of N protein in infected cells [158]. The partition coefficient of N protein into condensates was 13 ± 2 (± s.d.).

Although N protein is unable to phase separate without RNA in 150 mM salt, 24 µM N protein efficiently formed condensates at lower salt with a half-maximal inhibition constant (IC_{50}) of 90 ± 1 µM NaCl (± s.e.) (Figure 15D). The addition of 50 ng/µL polyC RNA increased IC_{50} to 152 ± 6 mM NaCl (Figure 15D). The ability of N protein to phase separate without RNA and sensitivity of these condensates to salt indicates that these condensates may be driven, in part, by electrostatic interactions among N proteins, as observed for other phase separating systems [5, 159-161].

Condensate formation of N protein and RNA was dependent on protein-RNA stoichiometry. At 18 µM N protein, condensate formation was not observed in the presence of 0-5 ng/µL polyC RNA. Increasing the RNA concentration promoted phase separation with an optimal RNA

concentration of 100-500 ng/μL, at which charge neutralization occurs with the positively-charged N protein [47]. A further increase in RNA concentration dissolved these condensates (Figure 15E) [81, 82]. This reentrant phase separation behavior is characteristic of heterotypic RNA and protein interactions in phase separating systems [82].

We also showed that N-polyC condensates exhibit liquid-like, rather than solid-like, material properties. First, these condensates were nearly spherical with an aspect ratio of 1.3 ± 0.6 (mean ± s.d.). Second, I observed the fusion of condensates with a mean fusion time of 90 ± 30 s (mean ± s.d.) after they come into contact (Figure 15F). Finally, FRAP experiments showed that 90 ± 10% of N protein in condensates can slowly exchange with the solvent with the half recovery time of 80 ± 30 s (mean ± s.d.). In comparison, polyC RNA exhibited slower recovery and a lower mobile fraction (Figures 15G-H and 18), indicating that the RNA may be stabilized by a network of interactions with multiple N proteins in the condensates.

Similar to polyC RNA, polyA and polyU RNA substrates produced spherical condensates at micromolar concentrations of the N protein (Figure 19A) [162]. However, combining N protein with polyG that forms Hoogsteen basepairing [59] or polyAU that forms Watson-Crick base-pair interactions [163] led to the formation of non-spherical condensates (Figure 19A). These results indicate that the N protein forms liquid condensates with some of the homopolymeric RNA substrates, whereas other homopolymeric RNA substrates that form basepairing interactions result in asymmetric condensates [140, 164-166].

Next, I sought to characterize how the N protein interacts with SARS-CoV-2 genomic RNA. The SARS-CoV-2 RNA genome was reverse transcribed and assembled into a DNA plasmid [167]. Using this plasmid, I generated six 5 kb fragments (Figure 19B) and two 1 kb fragments of the viral RNA genome via in vitro transcription [35]. In silico methods predict that these RNA fragments can form intra- and inter-molecular base-pairing interactions and contain extensive secondary structure elements (Figure 20) [164, 168, 169]. The N protein formed non-spherical condensates with viral RNA fragments across a wide range of protein and viral RNA concentrations (Figures 19B-C and 20B-D). These condensates were dissolved by increasing the salt concentration, but they neither changed shape over time, fused with each other, nor were strongly affected by raising the temperature from 20 ℃ to 37 ℃ (Figure 21). I also did not detect fluorescence recovery of either the LD655-labeled N protein or Cy3-labeled viral RNA in FRAP assays (Figures 19D and 22). I concluded that N protein forms solid-like condensates with viral RNA in vitro and that basepairing interactions of the RNA may influence the material properties of the condensate [140, 164, 165].

Cross-linking mass spectrometry (CLMS) identifies N protein interaction sites

To understand the mechanism of phase separation of the N protein, we performed CLMS to identify interactions between different domains of the full-length N protein in the absence of RNA [170]. CLMS detects protein-protein contacts by covalently capturing closely positioned lysine residues with bifunctional reagents. We first crosslinked the soluble (not phase separated) N protein in 300 mM KAc using a bifunctional crosslinker bis(sulfosuccinimidyl) suberate (BS3) (Figure 24A) [155]. We detected that the N-terminal half of the protein, including NTD, makes diverse contacts throughout the entire protein (Figure 24A). There was also an abundance of

contacts between the regions immediately flanking the CTD on either side, referred to as R1 (aa: 235-256) and R2 (aa: 369-390).

Next, we performed quantitative CLMS measurements [171] comparing the soluble N protein in 300 mM KAc with phase-separated N protein in 100 mM KAc (Figure 23A). The soluble N protein was crosslinked with heavy (D12) BS3 whereas the phase-separated protein was crosslinked with light (H12) BS3. As a result, the crosslinked precursor ions from high and low salt conditions were spaced by 12 Da (Figure 23A). Interactions between specific regions that promote condensate formation are implicated by the ratio of the crosslinked precursor ion signal and its corresponding isotopic doublet (Figure 23A) [171]. This experiment was repeated by reversing the labels, such that (H12) BS3 was used to crosslink the soluble N protein and (D12) BS3 was used for the phase-separated N protein. Across two independent experiments, 29 unique crosslinks were enriched and 30 crosslinks were depleted upon phase separation (Figures 23B and 24B). Although lysine residues are distributed throughout the N protein (Figure 24A) and depleted crosslinks spanned the entire primary sequence, the analysis of the crosslink fold change (Table 2) revealed that nearly all of the enriched interactions are concentrated in regions R1 and R2 (Figure 23C). These results suggest that the interactions involving the amino acids in regions R1 and R2 may be important in driving phase separation (Figures 23D and 24C).

Our mass spectrometry (MS) analysis also found phosphorylation sites on the N protein (Figure 23E). While some of these sites have been identified in previous studies [40, 172, 173], we also identified several novel sites (Table 3). Although one of the phosphorylation sites (S176) is involved in a crosslink, the phosphorylated and unphosphorylated peptides were both strongly depleted in condensates, suggesting that S176 phosphorylation does not play a major role in phase separation (Tables 2 and 3, Figure 24B). Additionally, MS identified native proteins that co-purified with the N protein in 1 M salt (Table 4). Consistent with the recruitment of N to stress granules in cells [31, 46-48], two of the most frequently identified proteins were stress granule proteins G3BP1 and G3BP2, with R2 of the N protein interacting with G3BP1 (Figure 24D).

The C-terminal region and phosphorylation modulate phase separation

The quantitative CLMS experiments show that interactions between R1 and R2 are correlated with the formation of condensates. However, I could not exclude the possibility that some of the changes in pairwise interactions are due to differences in protein structure or electrostatic interactions induced by differences in salt concentrations used for the soluble and condensate phase. To directly test the predictions of the CLMS results, I determined how different domain deletions affected phase separation of the N protein with polyC RNA under the same salt concentration (Figures 25A and 26A). Phase separation was only moderately reduced in the ΔSR, ΔPLD, and ΔR1 mutants (Figures 25B and 27), suggesting that these regions are not essential for phase separation. In comparison, deletion of the R2 sequence fully abolished the formation of condensates with polyC RNA (Figures 25B and 27A-D). Similarly, ΔR2 was unable to form condensates with viral RNA, whereas other deletion mutants phase separated with the same RNA fragments (Figure 27E). ΔR2 maintained a high affinity to bind polyC and viral RNA (Figure 26C), excluding the possibility that the disruption of phase separation is due to the lack of protein-RNA interactions. These results show that protein-protein interactions driven by the

R2 motif are required for phase separation of N protein with RNA, consistent with the proposed role of this sequence in the oligomerization of N protein [144].

Recent studies proposed that the N protein is phosphorylated early in infection, and this results in localization of N protein with the RTC, where it enhances transcription of subgenomic RNA [38]. However, nucleocapsid formation does not require phosphorylation and N protein in SARS-CoV viruses is hypophosphorylated [38, 174]. The underlying molecular mechanism and kinases and phosphatases responsible for post-translational modification of N protein remain poorly understood [30]. To understand how phosphorylation affects phase separation, I treated the full-length N protein with casein kinase 2 and λ protein phosphatase (see Methods). While kinase treatment did not alter the migration of N protein on a denaturing gel, phosphatase treatment resulted in a reduction in molecular weight (Figure 26A), suggesting that the N protein expressed in human cells is phosphorylated. Phosphatase treatment resulted in phase separation at slightly lower concentrations in comparison to kinase-treated N protein (Figure 25C-D). Similar to WT N, dephosphorylated ΔSR or ΔR1 had only a moderate increase in phase separation (Figures 26 and 28). However, unlike untreated ΔR2, dephosphorylated ΔR2 exhibited robust phase separation with polyC RNA (Figure 25E-F). These results suggest that phosphorylated N is capable of phase separation with RNA, primarily due to interactions between R2 motifs. In the absence of the R2 motif, phosphorylation negatively regulates phase separation of the N protein.

Targeting phase separation of the N protein with small molecules

We next sought to identify small molecules that could interfere with the phase separation of the N protein. The condensates formed by the N protein in the absence or presence of polyC or viral RNA dissolve with the addition of 10% 1,6 hexanediol, indicating that phase separation is driven, at least partially, by hydrophobic interactions (Figure 29A) [5]. In comparison, lipoic acid that increases the liquidity of stress granules [175] did not affect phase separation (Figure 29A). Using a high-throughput microscopy platform, we also tested whether any of the 1200 compounds in an FDA-approved drug library alters phase separation of the N protein with polyC RNA in vitro. While none of the compounds fully dissolved the condensates, we identified several compounds that affected their number, size, and shape (Table 5). Nelfinavir mesylate and LDK378 produced larger but fewer condensates, whereas crystal violet, tolcapone, and chlorhexidine enhanced phase separation by increasing the number and size of the condensates (Figures 29B and 30A-B). Nilotinib resulted in a 50% increase in condensate volume and altered the shape of the condensates (Figure 29B), which may be due to changes in condensate fusion or maturation. While most drugs did not have a substantial effect on condensates formed by N and viral genomic RNA in vitro, nilotinib addition resulted in the formation of thread-like filaments (Figure 30C-D). Morphologically, these filaments were similar to those formed during the liquid-to-solid transition of FUS condensates [176], suggesting that nilotinib increases the viscosity of the condensates.

We then tested these drugs in human pulmonary epithelial (Calu-3) and African Green Monkey kidney (Vero-E6) cells infected with SARS-CoV-2 (see Methods), as these cell lines supported high levels of infection [177]. The cells were treated with different concentrations of drugs 1 h before infection, and the inhibition of SARS-CoV-2-mediated cell death under drug treatment was quantified using the cytopathic effect (CPE) inhibition assay [177]. The toxicity of the drugs

was quantified by measuring the viability of uninfected cells under drug treatment. Among the molecules we identified, nelfinavir mesylate resulted in the highest percent CPE inhibition in both cell lines without significantly affecting cell viability (Figures 29C and 31), with the exception that the highest dose (40 μM) was toxic to the cells.

We also tested the effectiveness of these drugs using a median tissue culture infectious dose ($TCID_{50}$) assay. The supernatant from Vero-E6 cells that had been cultured with virus and drugs was added to uninfected cells, and the viral titer was measured by observing CPE in these cells. In this assay, nelfinavir mesylate, nilotinib, and LDK378 each reduced SARS-CoV-2 titer viability, with nilotinib and LDK378 producing an inhibitory effect similar to remdesivir, which strongly inhibits proliferation of SARS-CoV-2 in infected cells (Figures 29C and 31C) [178]. Collectively, our drug screen identified compounds that affect the condensate formation of N protein and RNA in vitro, inhibit virus-mediated cell death, and reduce viral titer.

The N protein phase separates in mammalian cells

To test whether N protein also phase separates in mammalian cells, I expressed N-GFP in HEK293T cells. I observed the formation of distinct puncta in the cytoplasm in N-GFP expressing cells, which were not observed in control cells expressing GFP only (Figure 32A). FRAP imaging of these puncta revealed a recovery signal with 60% mobile fraction and 6.3 ± 0.1 s recovery lifetime, suggesting that N protein is capable of forming liquid condensates in cells (Figures 32B and 33A-D). The recovery of N-GFP was substantially slower than GFP only (Figure 32C-E) but an order of magnitude faster than that of N-polyC condensates in vitro (Figure 15H) [30, 155], suggesting that these condensates are less viscous than N-polyC condensates formed in vitro. This difference is not due to the absence or low stoichiometry of RNA in N-GFP puncta in live cells because N condensates formed in the absence of RNA in vitro also exhibited an order of magnitude slower recovery than N-GFP puncta in live cells (Figure 34). I next tested whether deletion of the R2 region disrupts phase separation of N protein in cells. Surprisingly, cells expressing the R2 deletion mutant (ΔR2-GFP) still formed puncta and did not display significantly different FRAP dynamics than N-GFP in live cells (Figure 32). This may be due to macromolecular crowding of the cellular environment because the addition of 10% PEG triggered phase separation of ΔR2 without RNA in vitro (Figure 35).

Additionally, I sought to determine whether the N RNA transcript was recruited to the N condensate. I transfected U2OS cells stably expressing N-Clover [155] with a construct producing N-MS2 RNA (see Methods) and performed fluorescence in situ hybridization (FISH) assays against the MS2 RNA sequence. In the absence of N-MS2 transfection, the MS2 probe did not localize to the N-Clover puncta in any of the cells (N = 55 cells, two independent experiments) (Figure 32F). In comparison, the MS2 probe was strongly localized to all of the N-Clover condensates in ~18% of the cells transfected with N-MS2 (N = 10 in 55 cells, two independent experiments, Figure 32F). In other cells, I observed that the MS2 probe was uniformly distributed inside and outside of the N puncta, suggesting that these cells were not expressing N-MS2 at high enough levels to recruit the probe. I also performed RNA FISH using a d(T)20 probe that binds the polyA tails of RNA transcripts. Consistent with a previous report [155], the d(T)20 probe was not recruited to N-Clover condensates (Figure 33E). These results suggest that N-Clover condensates do not nonspecifically recruit nucleic acids, but selectively recruit viral RNA. Differences between my in vitro and in vivo phase separation might be

attributed to macromolecular crowding of the cytosol or the interaction of the N protein with other cellular proteins, such as G3BP1 (Figure 24D) [46-49].

Discussion

In this study, I showed that the SARS-CoV-2 N protein phase separates with nonspecific RNA sequences in vitro and the viscosity and shape of these condensates depend on the structure of the RNA substrate. Concurrent studies have shown that the SARS-CoV-2 N protein expressed in bacteria forms biomolecular condensates with RNA in vitro [30, 35, 36, 39, 44, 47, 48, 149, 155, 166, 179, 180]. I purified N protein from mammalian cells to recapitulate the post-translational modifications that occur in infected human cells [40]. Consistent with Carlson et al. and Chen et al., I observed that the N protein forms liquid condensates with various RNA substrates in physiological salt [30, 44]. Condensates of N with homopolymeric RNA that do not form Watson-Crick or Hoogsteen basepairing interactions recovered from photobleaching and relaxed to a spherical shape, suggesting that they are dynamic, liquid-like compartments. In agreement with Iserman et al., I found that RNA structure affects the material properties of the condensate [35]; RNA capable of Watson-Crick or Hoogsteen basepairing produces irregularly-shaped condensates [166, 180]. N protein also formed solid-like condensates with long fragments of SARS-CoV-2 genomic RNA. These condensates had asymmetric shapes and did not relax over time or with increased temperature [30, 47]. Viral RNA may drive the formation of abnormal condensate shapes by forming networks of intermolecular interactions, as previously reported [140, 164, 165, 181].

We also investigated the mechanisms underlying condensate formation using CLMS and protein engineering [155]. Previous studies reported that prion-like and SR motifs are common features of phase separating proteins [147, 151], and the linker sequence outside the SR motif (amino acids 210-247) is essential for phase separation of N expressed in bacteria [155]. The SR motif was required for N/N interactions in SARS-CoV and for forming puncta in SARS-CoV infected cells [182]. In addition, the SR motif is highly phosphorylated in human cells, and phosphorylation has been shown to promote interactions with host proteins, nuclear targeting, and transcription of the viral genome, and either enhances or inhibits oligomerization in SARS-CoV [31, 38, 174, 183]. In SARS-CoV-2, SR phosphorylation is reported to make N condensates more liquid [30, 155], and the deletion of the SR domain enhanced N phase separation in cells [48, 155]. However, I observed that the SR region is not necessary for in vitro condensate formation of N protein expressed in human cells. Although the deletion of the entire C-terminal IDR was reported to enhance phase separation [30, 144], I found that deletion of the R2 motif in this region fully disrupts phase separation of N protein with both homopolymeric and viral RNA in vitro, but dephosphorylation of this mutant recovered phase separation. The C-terminal region interacts with the M protein, which was shown to drive phase separation of the N protein in the absence of RNA [144]. Because the C-terminal region plays an important role in phase separation, binding of the N protein to the membrane-associated M protein through this region can alter the material properties of the N-RNA condensates and initiate virion assembly.

It remains to be demonstrated whether condensate formation of N protein and the viral genome is essential for the propagation of SARS-CoV-2 in human cells. The N protein is the most abundantly expressed viral protein in infected cells [38], and I observed that it forms condensates

in cells and recruits the viral RNA to these condensates. Phase separation has been implicated as a possible mechanism in the viral life cycle. For example, several viruses have been shown to replicate in viral inclusion bodies that are characterized as phase-separated condensates [42, 143]. A recent in vitro work has shown that nucleoproteins and phosphoproteins of the measles virus form liquid-like membraneless organelles and triggered nucleocapsid formation [41], suggesting that phase separation could be a general mechanism for viral replication. In the case of SARS-CoV-2, phase separation of N protein can form membrane-less compartments and function as a selectivity barrier to control the entry of certain agents into these compartments. Recent in vitro studies have shown that N-RNA condensates recruit the components of the SARS-CoV-2 replication machinery [47]. This mechanism may increase the efficiency of replication of the viral genomic RNA by increasing the local concentration of the replication machinery at the RTC complex.

N condensates may also sequester the viral assemblies from the immune response of the host cell [30]. For example, N protein interacts with stress granule proteins G3BP1 and G3BP2 in vitro [46, 48, 166], localizes to stress granules in cells [31, 46-48], and inhibits stress granule formation in cells overexpressing N protein [49, 50] and in SARS CoV-2 infected cells [51]. N mutants that do not interact with G3BP1 failed to suppress stress granule formation in SARS CoV-2 infected cells, and these cells were impaired in mediating SARS CoV-2 viral-like particle (VLP) production [184]. LLPS of SARS CoV-2 N protein has also been shown to inhibit poly-ubiquitination and aggregation of mitochondrial antiviral-signaling protein (MAVS, also known as IFN-β promoter stimulator I) and thereby suppress the antiviral immune response [51]. Lys375 mutation or a peptide targeting the CTD of N protein disrupted LLPS of N with RNA in vitro, inhibited SARS CoV-2 replication, and rescued antiviral immunity in mice [51]. Future work will be required to dissect the precise role of N-mediated phase separation in suppression of the innate immune response in infected organisms.

The phase separation of N protein with viral RNA may also facilitate the compaction and packaging of the genome into the nascent viral particle (Figure 32G). Consistent with this possibility, recent studies have shown that phase separation plays an essential role in the formation of heterochromatin regions in the nucleus of mammalian cells [6, 7]. However, this model raises several questions. Because N protein can phase separate with nonspecific RNA substrates, it remains unclear how these condensates may exclude subgenomic viral RNA and other RNA from the host cell. Recent in vitro work has shown that N protein has a higher affinity to bind 5’ and 3’ untranslated regions of the genomic RNA [35, 36], which may serve as a mechanism to trigger genome packaging and exclude other RNA from the condensates. In addition, condensates observed in vitro and in vivo are large structures that can potentially contain thousands of N proteins and RNAs, and it is not clear how a virion with a single genomic RNA can bud from these structures. A recent modeling study proposed that the presence of high-affinity sites in genomic RNA can trigger the formation of single-genome condensates [149]. Alternatively, interaction with the M protein or dephosphorylation of N protein, which could trigger the liquid-to-solid transition of N-RNA condensates [30, 155], at the viral assembly sites may trigger budding of a single genomic copy in virions. A condensate formed by N and a single genomic copy is expected to be stabilized once encapsulated in a lipid bilayer of the virion, and this would allow the virus to maintain its genome organization in a dilute phase. Testing of these models requires in-depth studies of N protein and genomic RNA in SARS-CoV-2 infected cells.

Phase separation could also provide a macroscopic readout to study N protein and RNA interactions [175] and suggest novel strategies to disrupt genome packaging and viral propagation in infected cells. I performed a screen of an FDA-approved drug library and identified several compounds that altered the size, number, and shape of N/RNA condensates in vitro. In particular, nelfinavir mesylate binds the SARS-CoV-2 protease [185-188], and both nelfinavir mesylate and nilotinib block SARS-CoV viral production [189, 190]. I showed that nelfinavir mesylate, nilotinib, and LDK378 each inhibit the proliferation of the virus in the host cell, which may be related to changes they induce on N condensates. Future work in infected cells is needed to address whether these drugs reduce virus-mediated cell death by interfering with the functions of N-RNA condensates.

Figures Related to Part 2

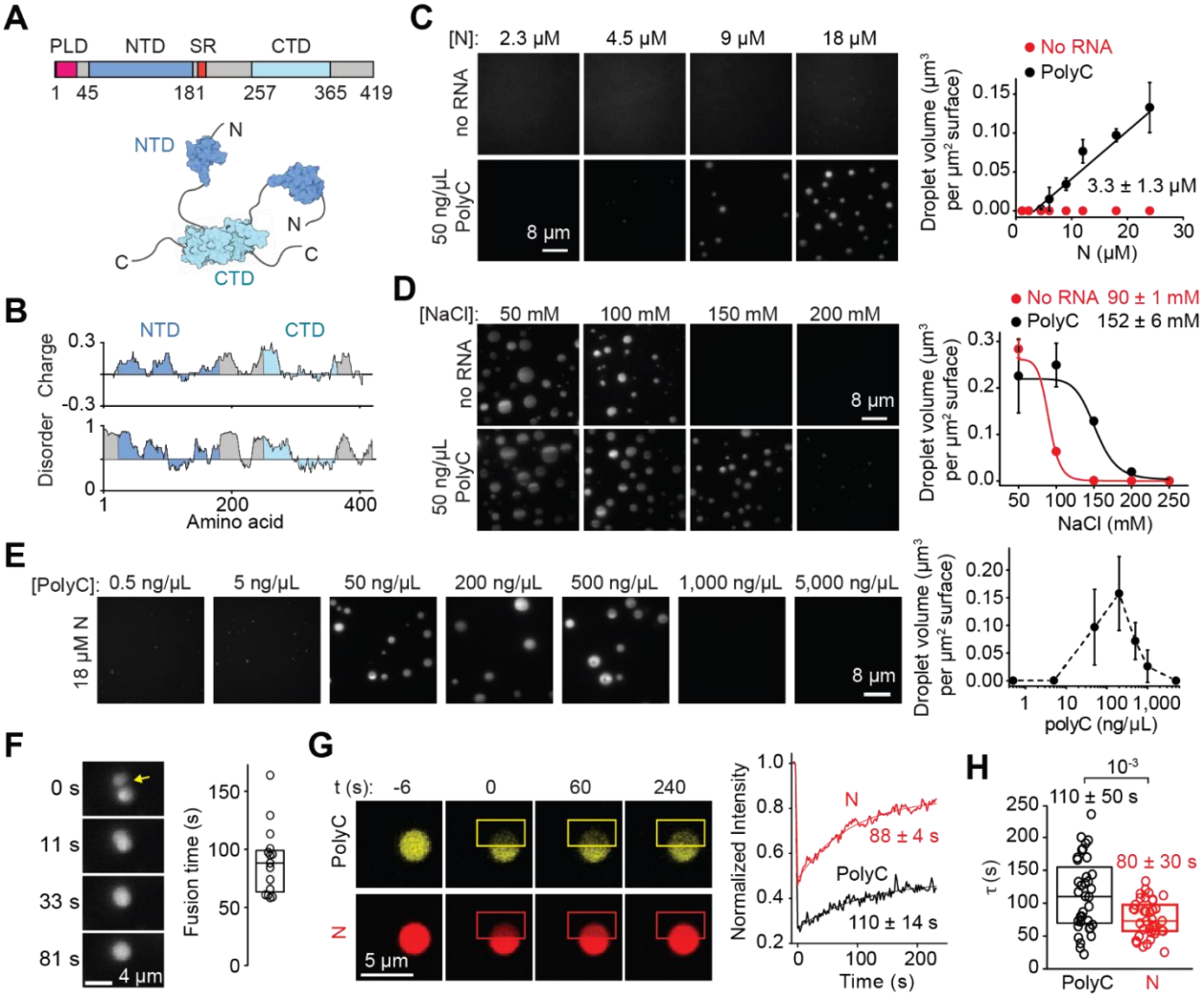

Figure 15. The SARS-CoV-2 N protein phase separates with RNA in vitro**. A.** Domain organization and the schematic of the N protein dimer. **B.** Sliding window plot of charge distribution (EMBOSS) and disorder prediction (IUPred2A) for the N protein. Charge y-axis represents mean charge across a 30-residue sliding window. Disorder prediction (1, disordered; 0, ordered) was calculated using the "long disorder" setting, encompassing a 30-residue sliding window. **C.** (Left) Images of the LD655-labeled N protein in the presence and absence of polyC RNA in 150 mM NaCl. (Right) The total volume of N-RNA condensates settled per micron squared area on the coverslip with 50 ng/μL polyC RNA (mean ± s.d., n = 20 with two technical replicates). A linear fit (solid line) reveals c_{sat} (± s.e.), the minimum N protein concentration for condensate formation (see Methods). **D.** (Left) Condensates formed by 24 μM LD655-labeled N protein in the presence or absence of 50 ng/μL polyC RNA dissolve by increasing NaCl concentration. (Right) The total volume of N condensates settled per micron squared area on the coverslip with increasing salt concentration (mean ± s.d., n = 10). Solid curves represent a fit to a dose-response equation to determine IC_{50} (± s.e.). **E.** The stoichiometry of the N protein and

RNA affects phase separation. (Left) Example pictures show that Cy3-labeled N protein forms condensates with different concentrations of polyC RNA. The N protein concentration was set to 18.5 μM. (Right) The total volume of N-polyC condensates settled per micron squared area on the coverslip under different RNA concentrations (mean ± s.d.; n = 20, two technical replicates). **F.** (Left) The fusion of N-polyC condensates formed in the presence of 18.5 μM LD655-labeled N and 50 ng/μL polyC RNA. (Right) Fusion time of N-polyC condensates (mean ± s.d., n = 15 fusion events). The center and edges of the box represent the median with the first and third quartiles. **G.** (Left) Representative FRAP imaging of an N-polyC condensate. The image of a condensate before the time of photobleaching (0 s) shows colocalization of Cy3-polyC and LD655-N in the condensate. Rectangles show the photobleached area. (Right) Fluorescence recovery signals of N and polyC in the bleached region. Solid curves represent a single exponential fit to reveal the recovery lifetime (τ, ±95% confidence interval). **H.** The distribution of fluorescence recovery lifetimes of N and polyC in droplets (n = 37). The center and edges of the box represent the median with the first and third quartiles. The p-value was calculated from a two-tailed t-test.

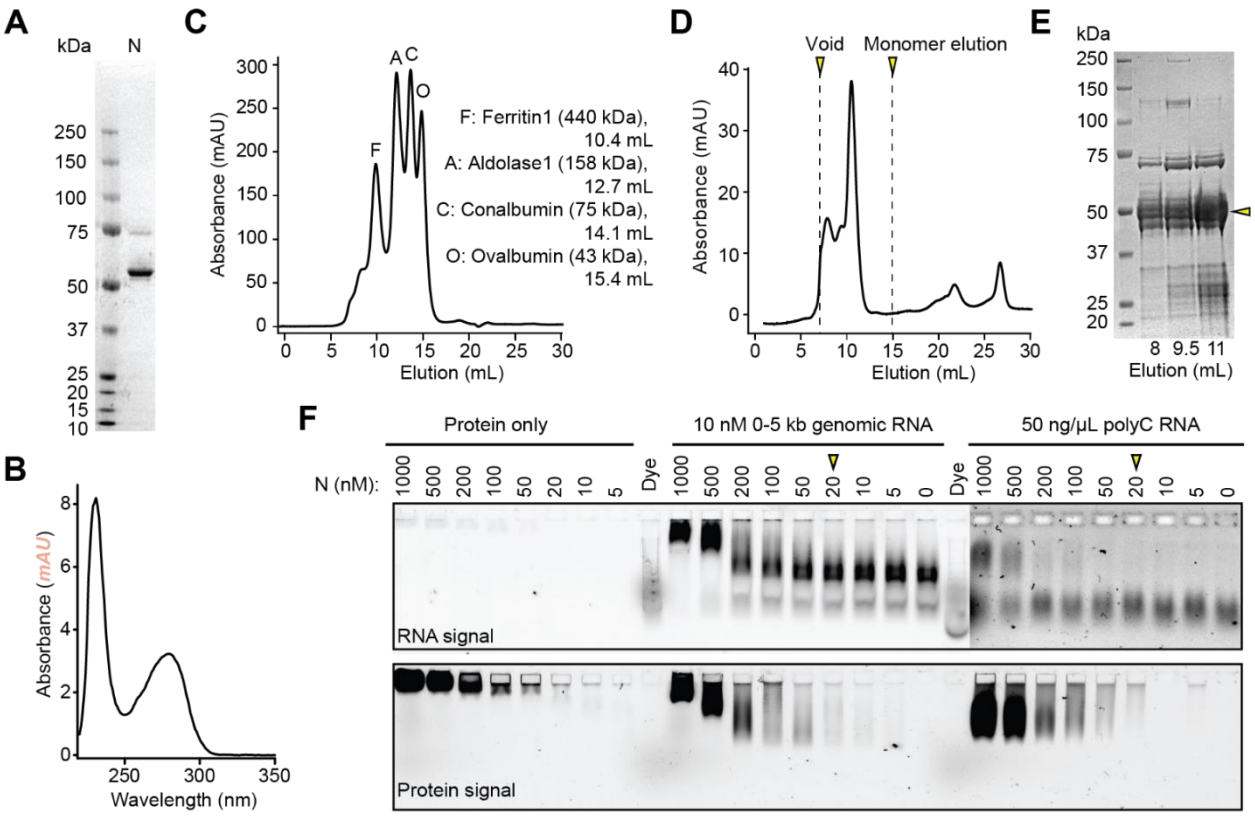

Figure 16. Gel filtration and EMSA analysis of the WT N protein purified from HEK293S GNTI- cells. A. The coomassie-stained denaturing gel of the N protein purified from affinity chromatography. **B.** UV absorbance of the N protein purified from affinity chromatography shows no evidence for the presence of contaminating nucleic acids**. C.** UV absorbance of protein standards eluting from a gel filtration column. **D.** UV absorbance of the N protein eluting from a gel filtration column. Arrows mark the void volume and expected elution volume for an N monomer. **E.** The coomassie-stained denaturing gel of the eluents from the gel filtration column. **F.** EMSA using no RNA, 10 nM 0-5 kb viral RNA, and 50 ng/µL polyC RNA and decreasing concentration of N protein. The protein was labeled with LD655. RNA was labeled with Cy3. Arrows indicate the minimum protein concentration for each condition with a noticeable shift in the protein band.

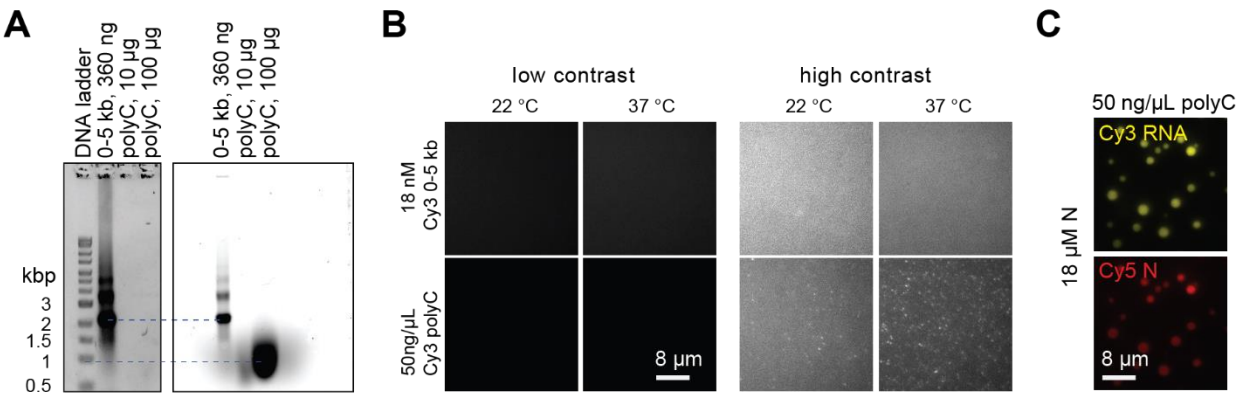

Figure 17. RNA substrates do not form condensates in the absence of N protein. A. An agarose gel picture of 0-5 kb viral RNA and polyC RNA substrates. The gel was stained with GelRed (left), and the RNA substrates were labeled with Cy3 (right). The ladder corresponds to the length of double-stranded DNA. Estimated lengths of 0-5 kb viral RNA and polyC RNA are 6 kb and 2 kb, respectively. **B.** Representative pictures show that 18 nM 0-5 kb viral RNA and 50 ng/µL polyC RNA do not form condensates in the absence of N protein. The assay was performed in 150 mM NaCl. **C.** Two-color imaging shows colocalization of LD655-labeled N protein and Cy3-labeled polyC RNA in condensates.

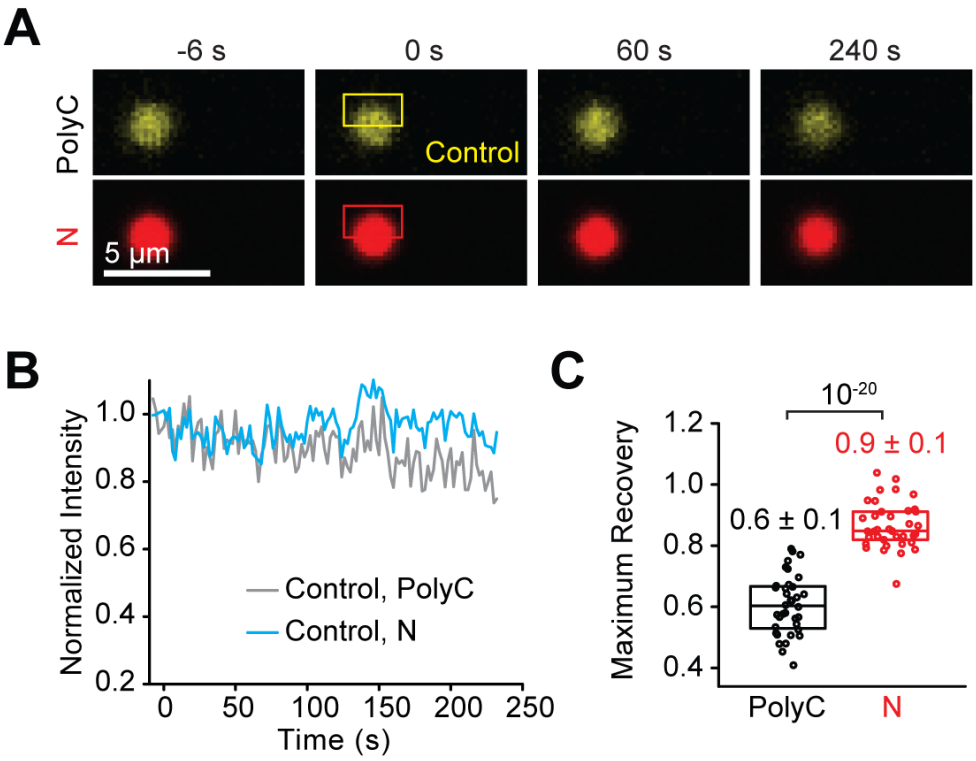

Figure 18. FRAP analysis of N-polyC condensates in vitro**. A.** Control experiments show changes in the fluorescent signal of N and polyC without photobleaching. **B.** The changes in the integrated fluorescent intensities of regions highlighted with yellow and red rectangles in A. **C.** The maximum fractional recovery of N and polyC in condensates after photobleaching (n =36, mean ± s.d.). The center and edges of the box represent the median with the first and third quartiles. The p-value was calculated from a two-tailed t-test.

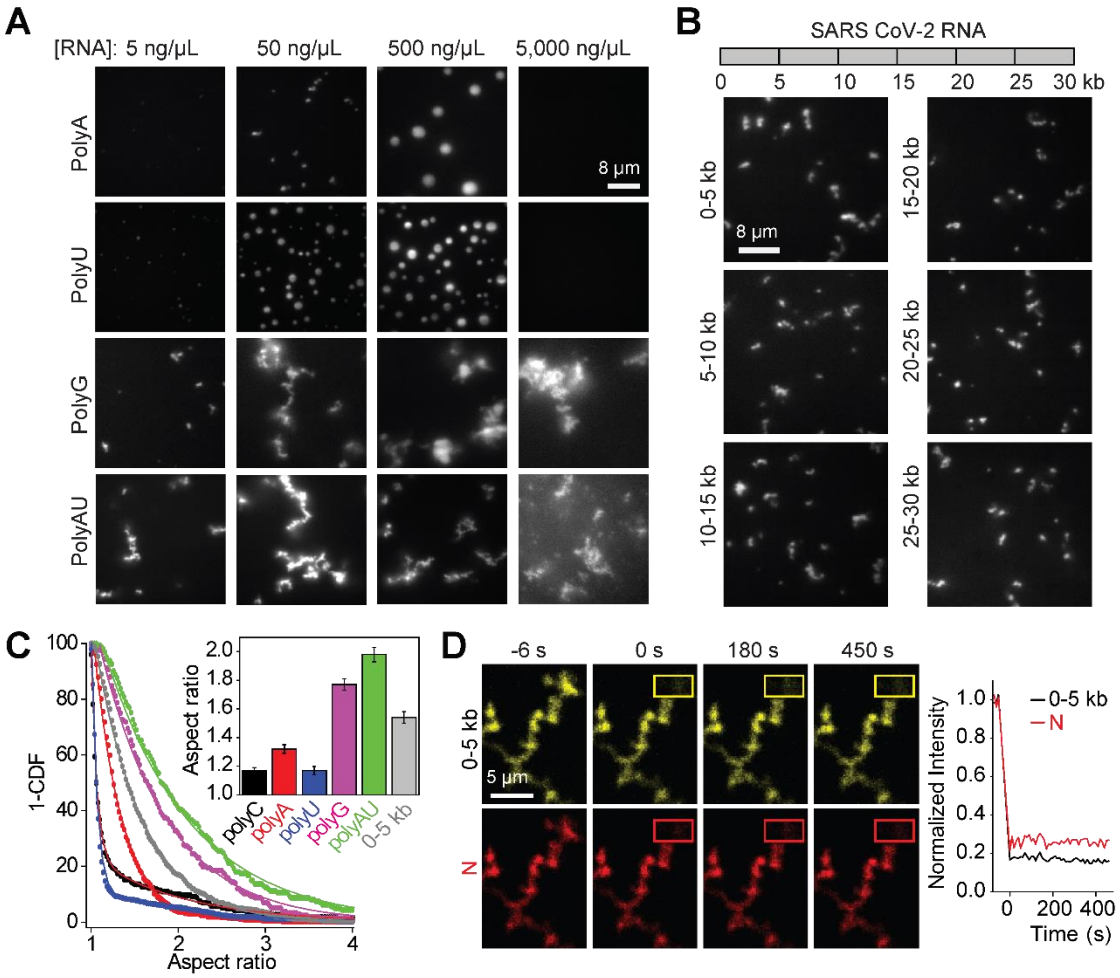

Figure 19. The N protein forms asymmetric condensates with viral RNA. A. The N protein forms spherical condensates with polyA and polyU RNA but forms asymmetric condensates with substrates that form Hoogsteen (polyG) and Watson-Crick (polyAU) basepairing interactions. The N protein concentration was set to 18.5 μM. **B.** (Top) SARS-CoV-2 genomic RNA was divided into six 5 kb sections. (Bottom) The formation of asymmetric N condensates in the presence of 18 nM RNA. The N protein concentration was set to 18.5 μM. **C.** The inverse cumulative distribution (1-CDF) of the aspect ratio of individual N condensates formed with different RNA substrates. The concentrations of N protein, RNA homopolymers, and 0-5 kb viral RNA were set to 18.5 μM, 50 ng/μL, and 18 nM, respectively. Solid fits represent a fit to exponential decay. (Insert) Decay constants of the exponential fits (± s.e.). **D.** (Left) Representative FRAP imaging of an N and 0-5 kb viral RNA condensate. The image of a condensate before the time of photobleaching (0 s) shows colocalization of Cy3-labeled 0-5 kb viral RNA and LD655-N in the condensate. Rectangles show the photobleached area. (Right) N and 0-5 kb viral RNA do not exhibit fluorescence recovery in the bleached region (n = 16).

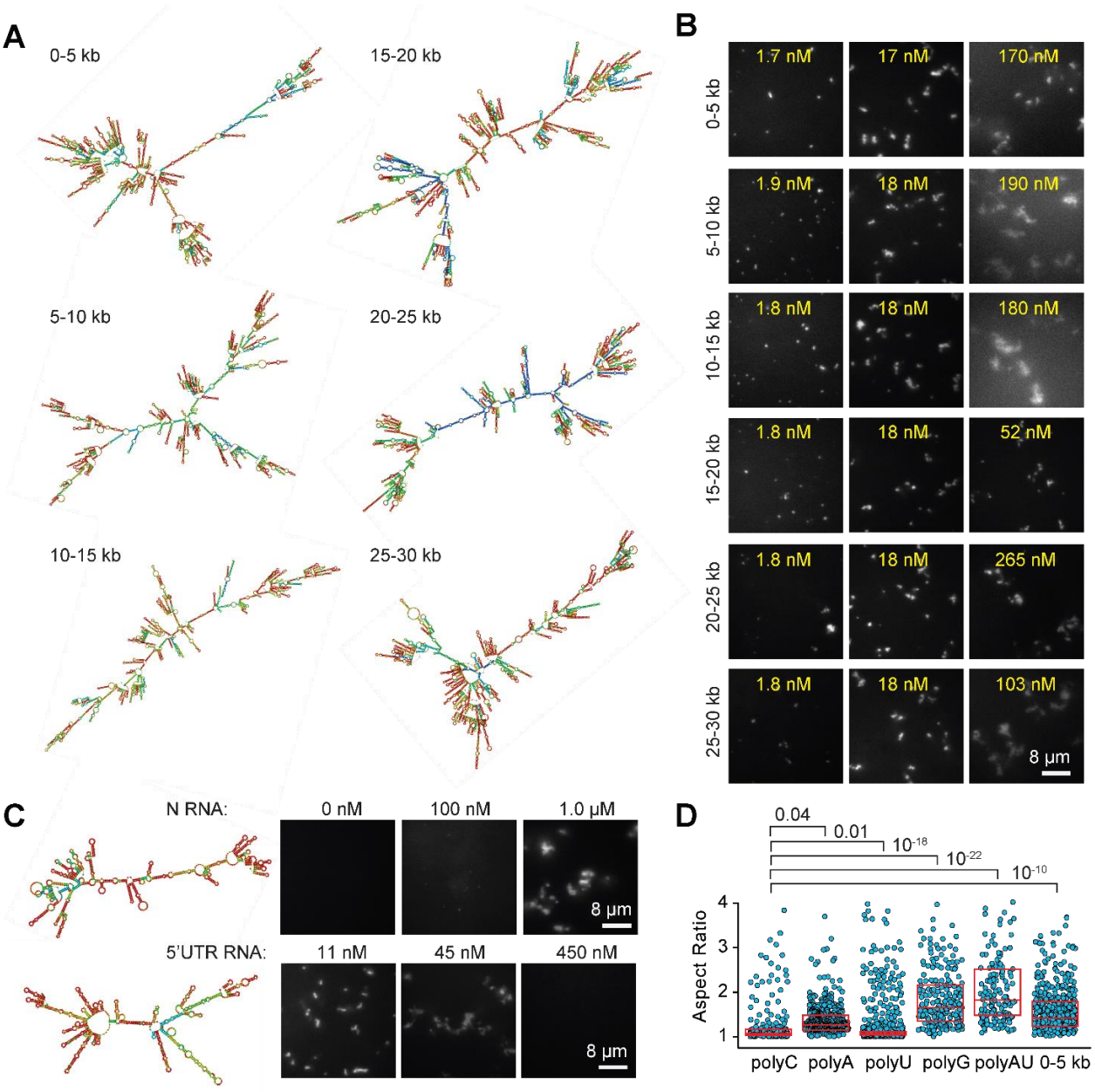

Figure 20. The N protein forms asymmetric condensates with viral RNA. A. Structure prediction of each section of SARS CoV-2 genomic RNA. **B.** The formation of asymmetric N condensates under different viral RNA concentrations. The N protein concentration was set to 18.5 μM. **C.** Structure prediction of viral RNA (left) and the formation of asymmetric N condensates under different RNA concentrations (right). The N RNA is the 1.3 kb long genomic RNA fragment that encodes the SARS CoV-2 N protein. The 5' UTR RNA is the first 1,000 bases of the SARS-CoV-2 genome. The N protein concentration was set to 18.5 μM (IVT: in vitro transcribed; UTR: untranslated region). **D.** The distribution of aspect ratios of individual condensates formed with different RNA substrates. The N protein concentration was set to 18.5 μM, and RNA concentration was set to 50 ng/μL for RNA homopolymers and 18 nM for 0-5 kb

viral RNA. The center and edges of the box represent the median with the first and third quartiles. P values are calculated from two-tailed t-tests.

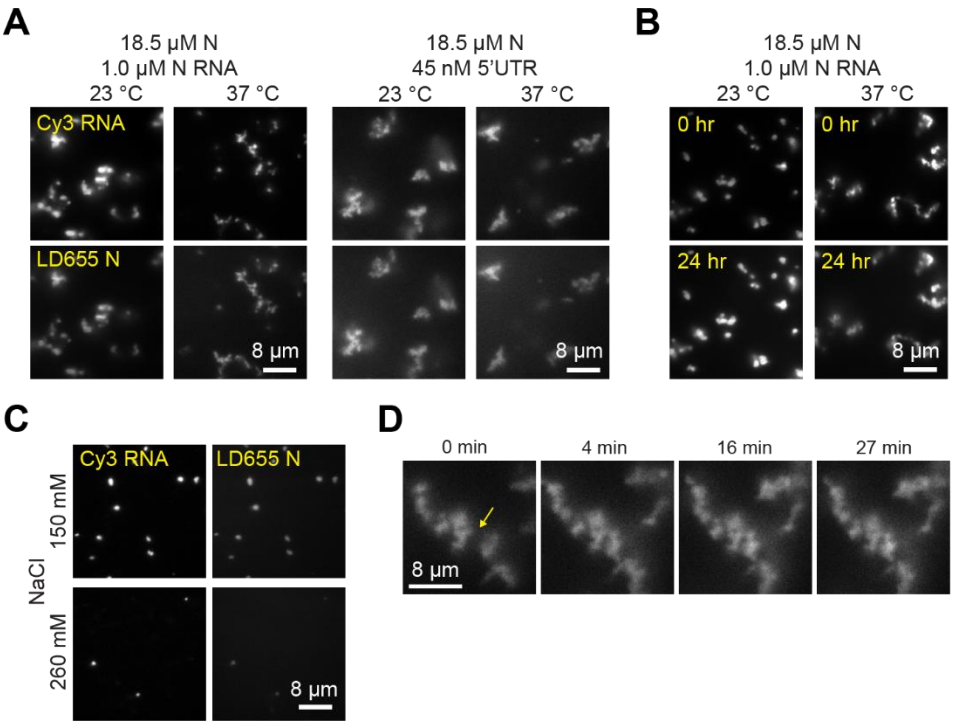

Figure 21. Condensates formed by the N protein and viral RNA do not fuse or change shape but are dissolved at higher salt. A. Condensates formed by the N protein with in vitro transcribed N RNA or the 5'UTR of the viral RNA are not sensitive to an increase of temperature to 37 ℃. **B.** Condensates formed by the N protein with 5'UTR of viral RNA do not change shape over 24 h. **C.** Condensates formed by 11.5 µM N protein and 36 nM 5'UTR RNA are dissolved in the presence of 260 mM NaCl. **D.** Condensates formed by 18.5 µM N protein and 18 nM 0-5 kb viral RNA do not fuse after contact (yellow arrow).

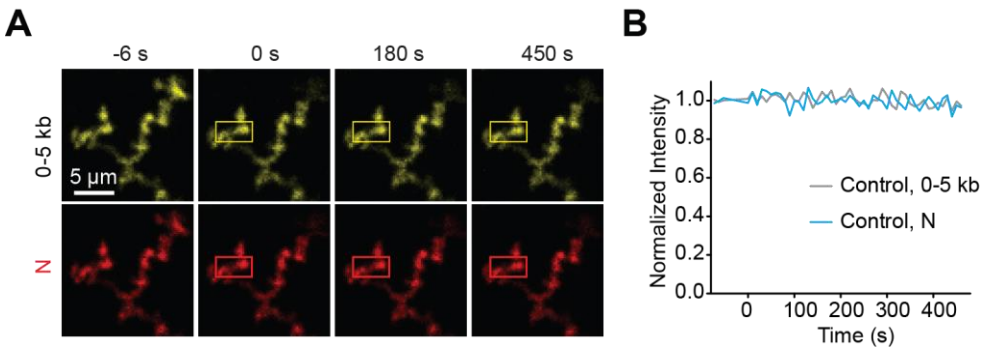

Figure 22. FRAP analysis of N and 0-5 kb viral RNA condensates in vitro**. A.** Control experiments show snapshots of N and 0-5 kb viral RNA without photobleaching. **B.** The changes in the integrated fluorescent intensities of the regions that are highlighted with yellow and red rectangles in **A.**.

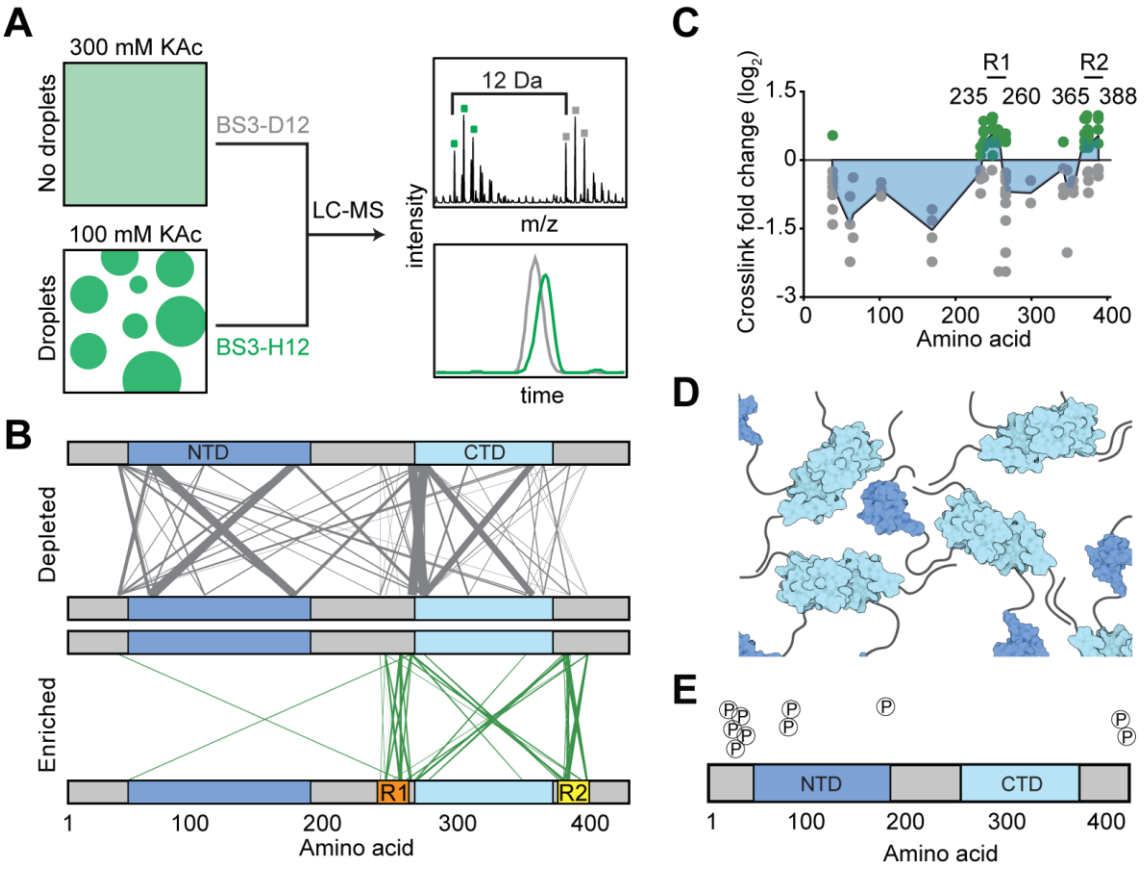

Figure 23. CLMS reveals inter-domain interactions of the N protein. A. Schematic of the CLMS experiment. (Left) A high salt (300 mM KAc) buffer disrupts N condensates, whereas a low salt (100 mM KAc) buffer promotes condensate formation. (Right, top) Example of an individual crosslinked peptide in quantitative CLMS analysis. Precursor ions from the high salt (gray) and low salt (green) BS3 crosslinking conditions show the 12 Da shift between light (H12) and heavy (D12) crosslinkers. (Right, bottom) Ion chromatograms from the first three isotopes of each doublet were extracted and expressed as the ratio of peak areas. **B.** The plot of crosslinks depleted and enriched in the condensate condition. The width and transparency of the lines scale with the number of times the crosslink was detected across 3 independent experiments. **C.** Fold changes of crosslink abundance upon condensate formation of N. As crosslinks contain two positions, fold change information is plotted at both positions. Only crosslinks with p-values less than 0.05 are included. Green and grey dots represent crosslinks enriched and depleted in the condensate condition, respectively. The blue area represents a plot of median crosslink fold change. **D.** Model for how multiple N dimers could phase separate via their disordered regions. **E.** Phosphorylation sites detected by the CLMS experiment in 300 mM KAc.

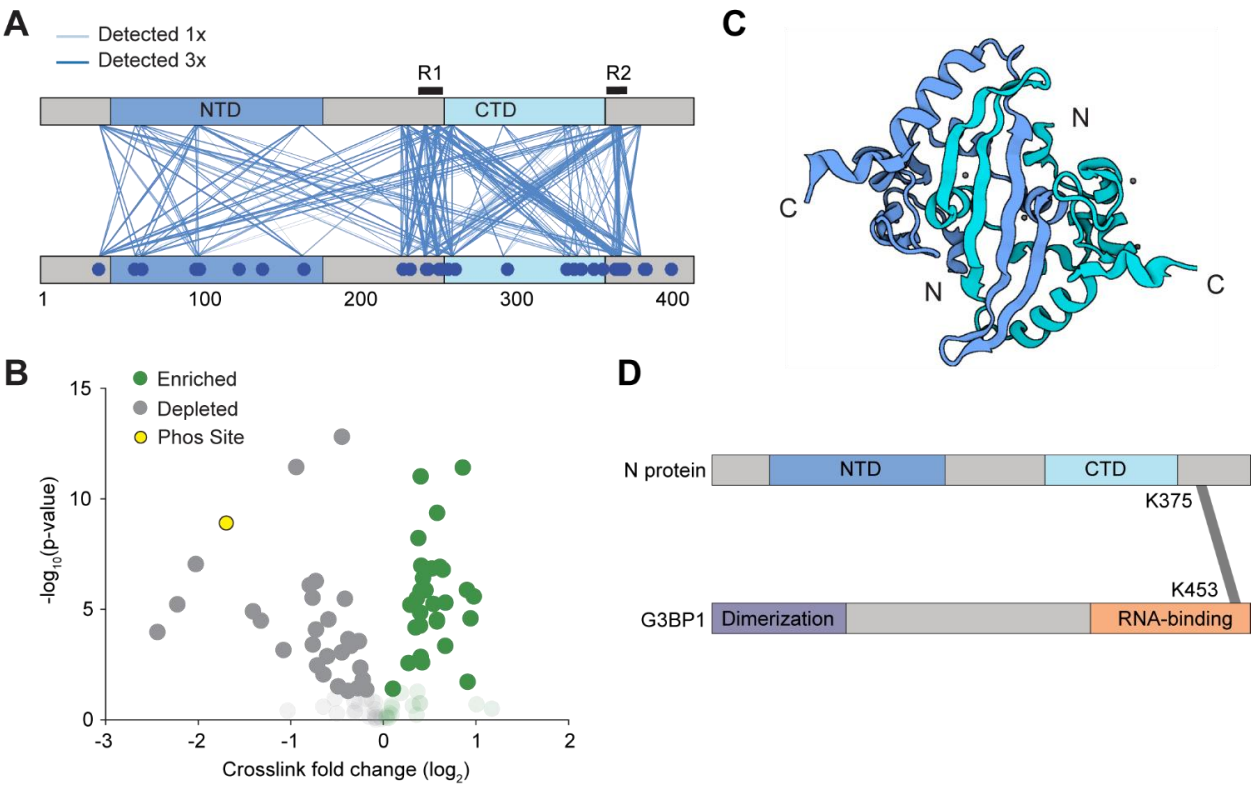

Figure 24. Interactions between N proteins and between N and G3BP1. A. In an initial qualitative experiment, pairwise interactions were detected by CLMS at 300 mM salt. Blue dots depict the positions of lysine residues. Lines depict a unique crosslink detected. The regions of N protein interactions flank the CTD. **B.** Volcano plot of the quantitative CLMS data comparing the condensate and no condensate condition. Opaque data points have a p-value below 0.05, transparent data points have a p-value greater than 0.05. Green represents unique crosslinks that are enriched in the condensate condition. The yellow markers represent the K169-K65 and K169 (phosS176)-K65 crosslinks. **C.** The structure of the CTD of the SARS-CoV-2 N protein was plotted with BioRender (PDB 6WJI [191]). Because the N and C termini of the protomers are positioned away from each other, R1 and R2 within the same dimer are unlikely to interact with each other. **D.** Mass spectrometry identified that the RNA binding domain of the stress granule protein G3BP1 interacts with the R2 region of N.

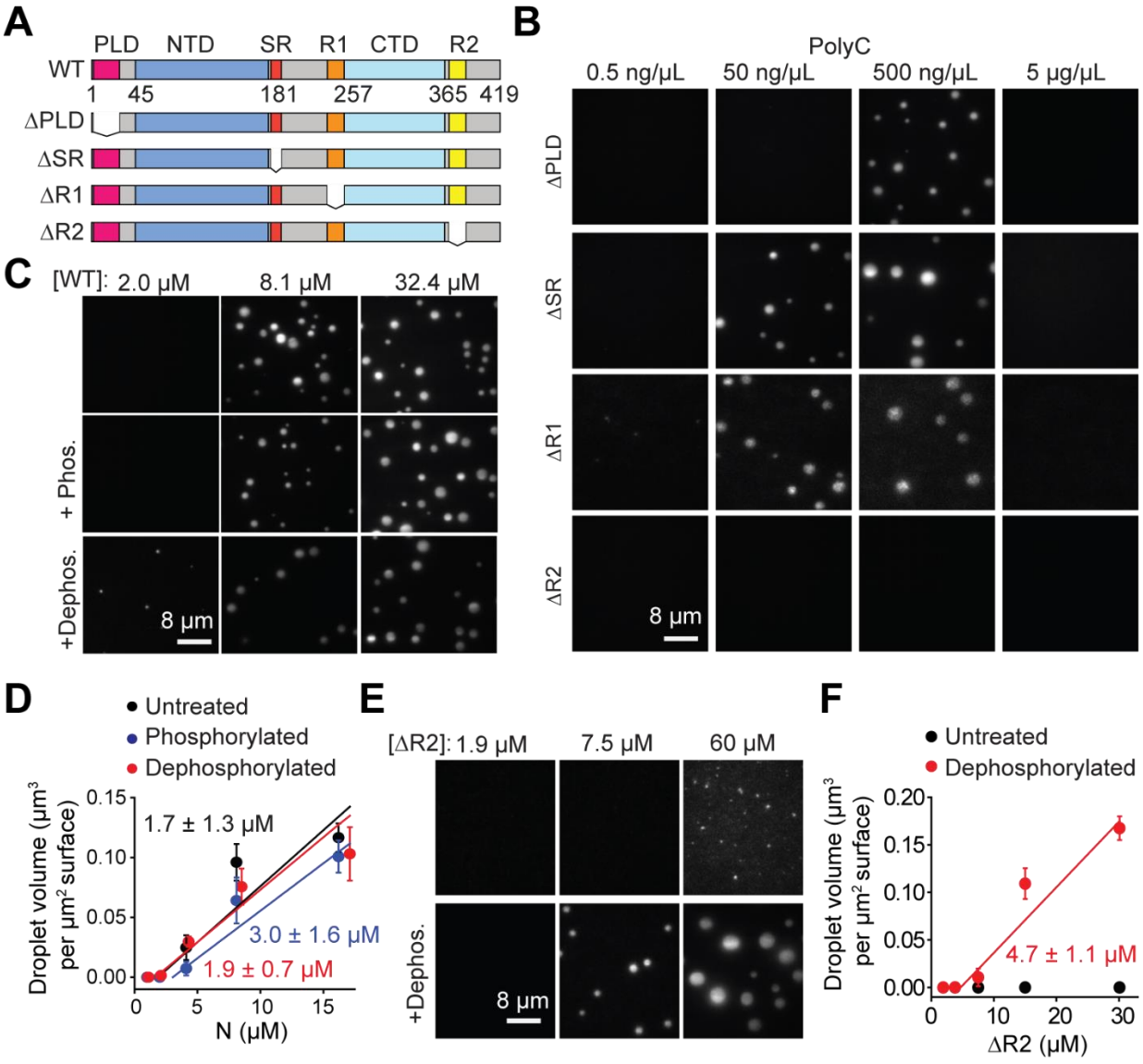

Figure 25. The effect of domain truncations and phosphorylation on phase separation of the N protein. A. Truncation mutants of the N protein. **B.** While ΔPLD, ΔSR, and ΔR1 phase separate, ΔR2 does not phase separate when mixed with polyC RNA. Protein concentration was set at 18 μM for all conditions. **C.** Images of condensates formed by untreated, phosphorylated, and dephosphorylated N protein in 50 ng/μL polyC RNA. **D.** The total volume of N-RNA condensates settled per micron squared area on the coverslip under different phosphorylation conditions (mean ± s.d., n = 20 with two technical replicates). Linear fits (solid lines) reveal c_{sat} (± s.e.). **E.** Images of condensates formed by untreated and dephosphorylated ΔR2 in 50 ng/μL polyC RNA. **F.** The total volume of ΔR2-RNA condensates settled per micron squared area on the coverslip as a function of ΔR2 concentration (mean ± s.d., n = 20 with two technical replicates). The linear fit (solid line) reveals c_{sat} (± s.e.).

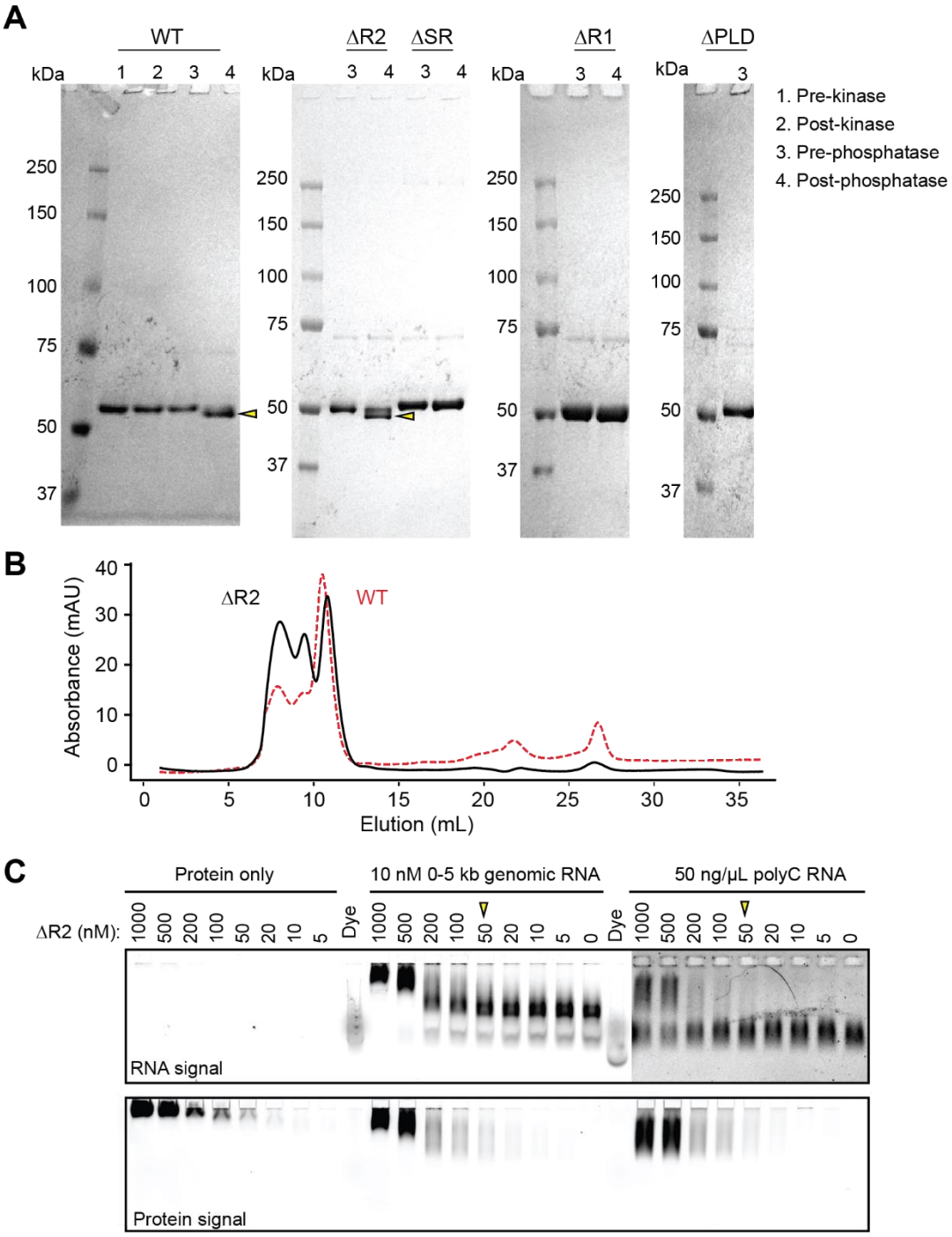

Figure 26. Purification and characterization of the deletion mutants. A. Denaturing gel pictures of purified WT and deletion mutants of N protein in the presence and absence of kinase and phosphatase treatment (see Methods). The gels were stained with Coomassie. Yellow arrowheads highlight a reduction in molecular weight upon treatment with λ phosphatase. **B.** UV absorbance of the ΔR2 mutant eluting from a gel filtration column. UV absorbance of WT N under the same experimental conditions is shown in a dashed red curve for comparison. **C.** EMSA gels using no RNA, 10 nM 0-5 kb viral RNA, or 50 ng/μL polyC RNA and decreasing

concentration of the ΔR2 mutant. The protein was labeled with LD655. RNA was labeled with Cy3. Arrows indicate the minimum protein concentration for each condition with a noticeable shift in the protein band.

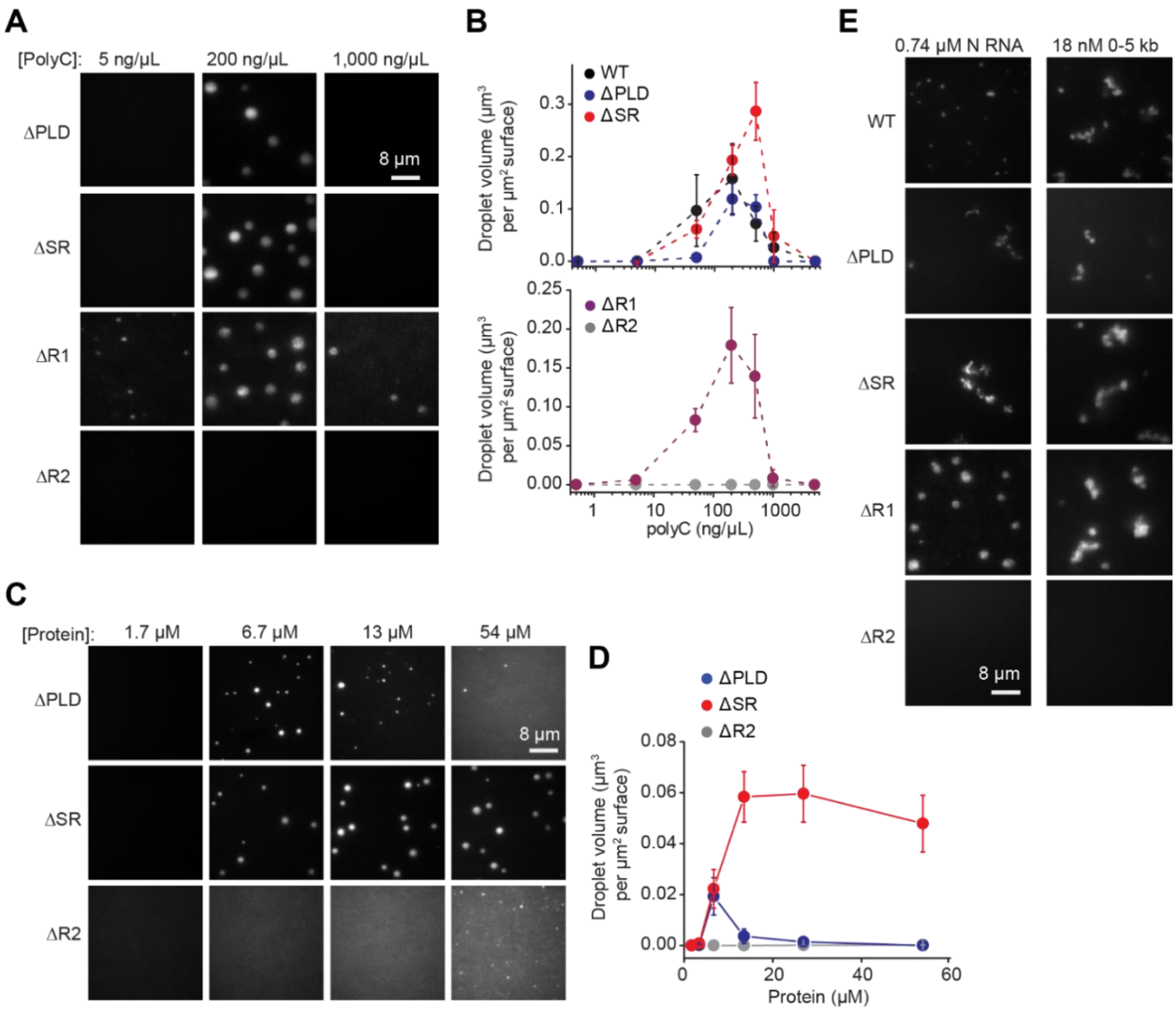

Figure 27. The in vitro **phase of the truncation mutants of the N protein. A.** Example pictures show that the N truncation mutants, except ΔR2, form spherical condensates with polyC RNA under different RNA concentrations. The N protein concentration was set to 18.5 μM. **B.** The total volume of N-RNA condensates settled per micron squared area on the coverslip (mean ± s.d.; n = 20, two technical replicates) exhibits a reentrant behavior under an increasing RNA concentration. **C.** Example pictures show that the N truncation mutants, except ΔR2, form spherical condensates with polyC RNA under different protein concentrations. The polyC RNA concentration was set to 50 ng/μL. **D.** The total volume of N-RNA condensates settled per micron squared area on the coverslip (mean ± s.d.; n = 20, two technical replicates) under an increasing protein concentration. **E.** Phase separation of truncated N protein constructs with 0.74 μM in vitro transcribed N RNA or 18 nM 0-5 kb viral RNA. The protein concentration was set to 18.5 μM.

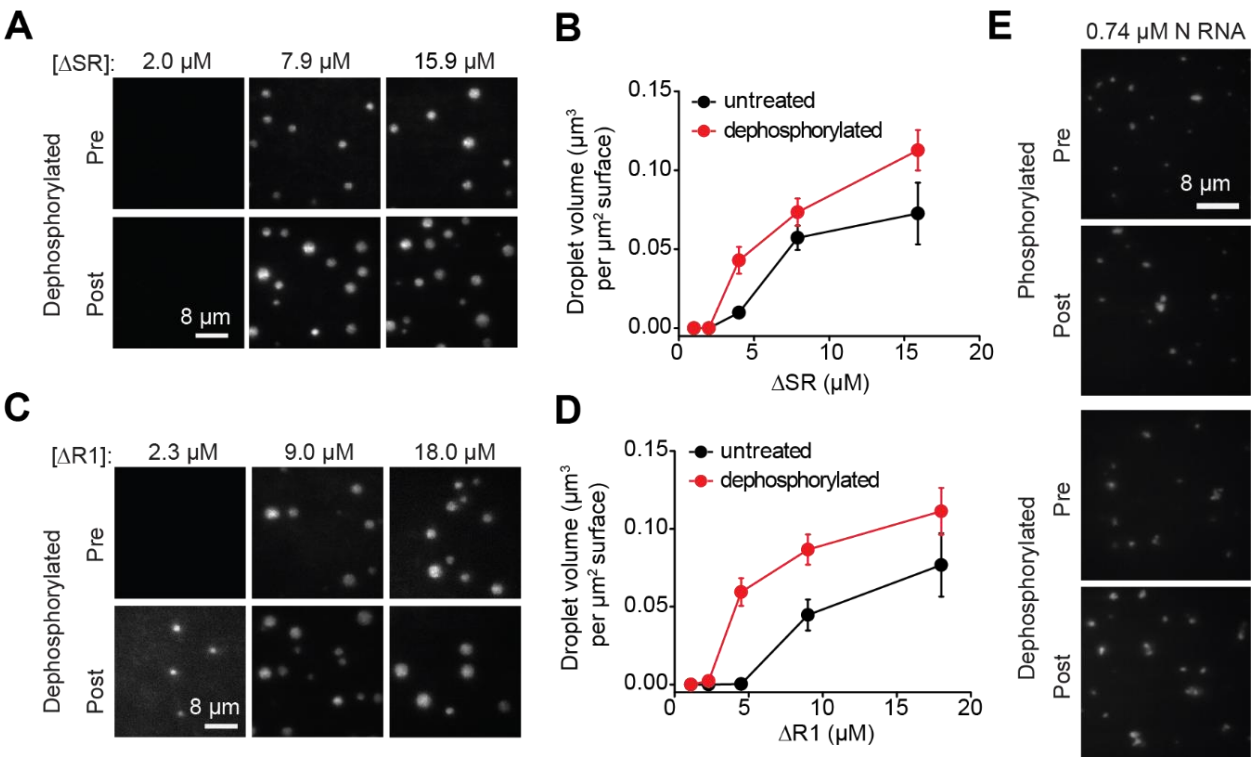

Figure 28. Phase separation of truncated N proteins under different phosphorylation conditions in vitro**. A.** Images of condensates formed by untreated and dephosphorylated ΔSR in the presence of 50 ng/μL polyC RNA. **B.** The total volume of the condensates settled per micron squared area on the coverslip as a function of ΔSR concentration (mean ± s.d., n = 20 with two technical replicates). **C.** Images of condensates formed by untreated and dephosphorylated ΔR1 in the presence of 50 ng/μL polyC RNA. **D.** The total volume of the condensates settled per micron squared area on the coverslip as a function of ΔR1 concentration (mean ± s.d., n = 20 with two technical replicates). **E.** Full-length N protein forms asymmetric condensates with 0.74 μM in vitro transcribed N RNA before and after phosphorylation and dephosphorylation. The protein concentration was set to 35 μM.

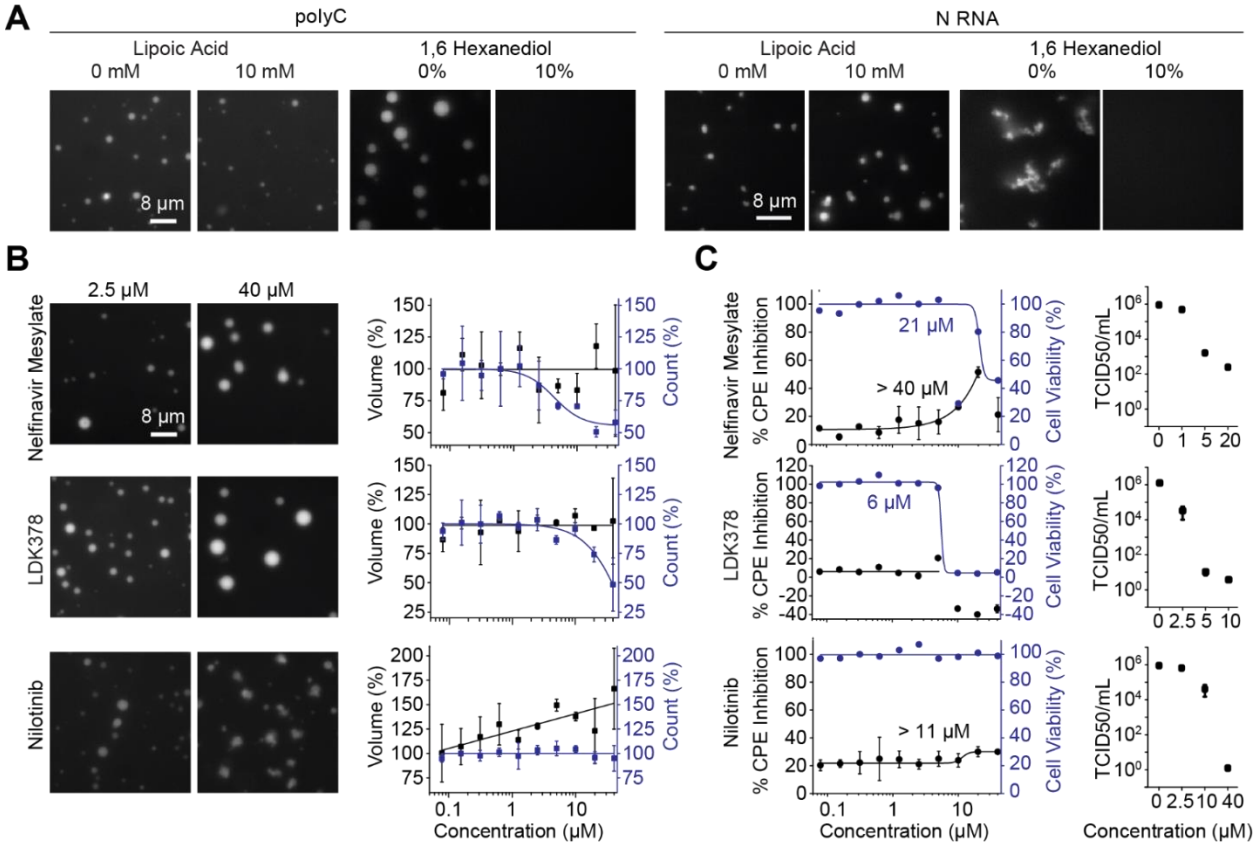

Figure 29. Identification of small molecules that alter phase separation of N in vitro **and** in vivo**. A.** Condensates formed in the presence of 7.8 µM N and either 50 ng/µL polyC or 18 nM 0-5 kb viral RNA were not affected by 10 mM lipoic acid but dissolved in the presence of 10% 1,6 hexanediol. **B.** (Left) Examples of drugs that affected N phase separation with polyC in vitro. The N protein and polyC RNA concentrations were set to 7.8 µM and 50 ng/µL, respectively. (Right) The percent change on the number (blue) and total volume (black) of N-polyC condensates settled per micron squared area on the coverslip under different drug concentrations (mean ± s.d., n = 8 with two technical replicates). Solid curves represent a fit to a dose-response equation (see Methods). **C.** (Left) Percent CPE inhibition (black, mean ± s.d., two technical replicates) and cell viability (blue) of SARS-CoV-2 infected Vero-E6 cells treated with serial dilutions of drugs. Solid curves represent a fit to a dose-response equation (see Methods) to determine a half-maximal response constant EC_{50} (Table 5). (Right) Viral titer in SARS-CoV-2 infected Vero-E6 cells treated with serial dilutions of drugs as measured by a $TCID_{50}$ assay (mean ± s.d., three technical replicates).

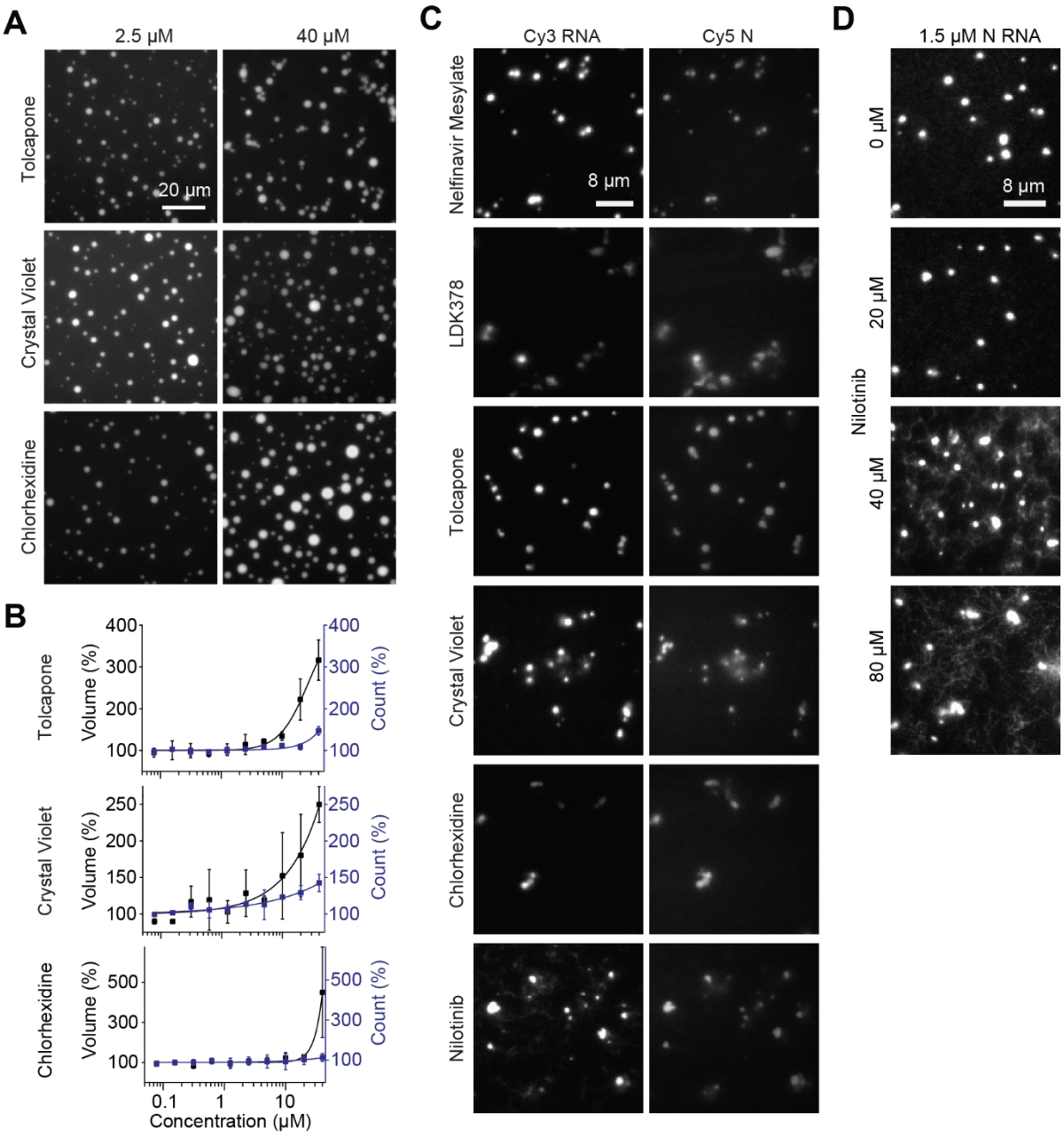

Figure 30. Phase separation of the N protein under drug treatment. A. Example pictures show phase separation of 7.8 μM N protein and 50 ng/μL polyC RNA at different drug concentrations. **B.** The percent change on the number (blue) and total volume (black) of N-polyC condensates settled per micron squared area on the coverslip under different drug concentrations (mean ± s.d., n = 8 with two technical replicates). **C.** Example pictures show phase separation of 57.6 μM LD655-labeled N protein and 1.5 μM Cy3-labeled in vitro transcribed N RNA in the presence of different drug treatments. The concentration of drugs was set to 40 μM. **D.** Condensates formed in the presence of 57.6 μM N protein and 1.5 μM N RNA form thread-like filaments at high nilotinib concentrations.

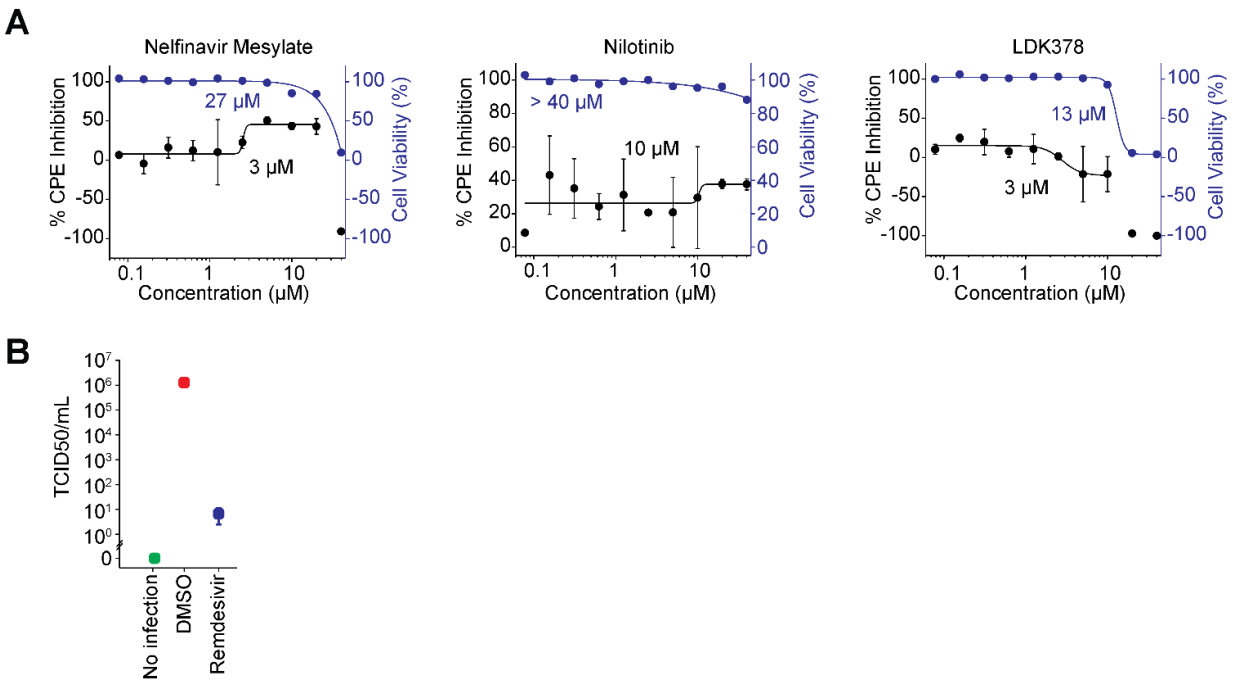

Figure 31. The viability of SARS-CoV-2 infected cells under drug treatment. A. Percent CPE inhibition (black, mean ± s.d., two technical replicates) and cell viability (blue) of SARS-CoV-2 infected Calu-3 cells treated with serial dilutions of drugs. Solid curves represent a fit to a dose-response equation to determine EC_{50}. **B.** Viral titer in uninfected and SARS-CoV-2 infected Vero-E6 cells as measured by a $TCID_{50}$ assay. Infected cells were treated with DMSO (negative control) and 50 μM remdesivir (positive control).

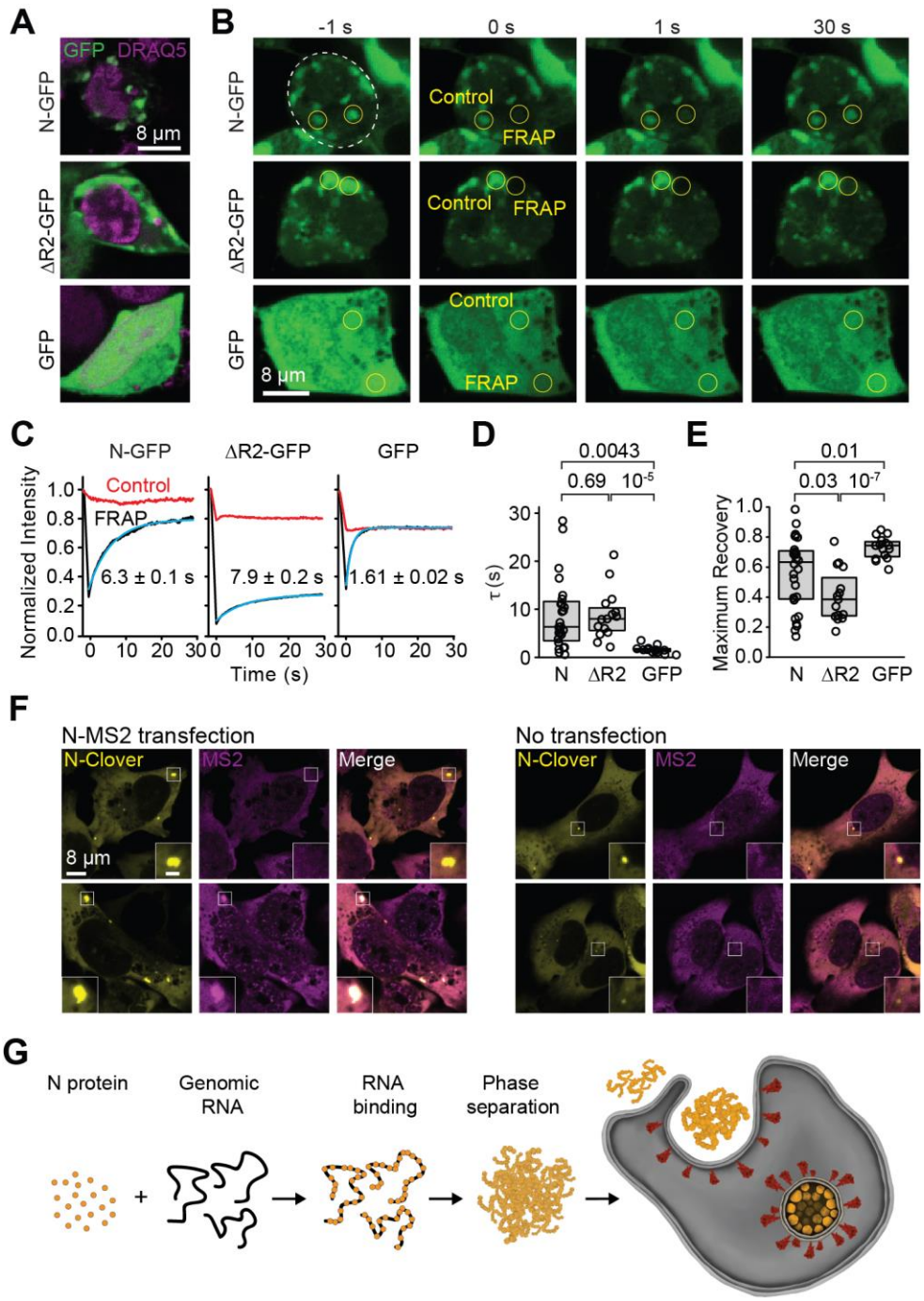

Figure 32. Dynamics of N condensates in vivo**. A.** Example images of cells expressing N-GFP, ΔR2-GFP, and GFP stained with DRAQ5. **B.** Representative FRAP imaging of cells exhibiting N-GFP or ΔR2-GFP puncta or high GFP expression. Circles show the photobleached area. **C.** Fluorescence recovery signals of the protein in the bleached versus the control regions. The solid curve represents a single exponential fit to reveal the recovery lifetime (τ, ±95% confidence interval). **D.** The distribution of fluorescence recovery lifetimes of cells expressing N-GFP (n = 28) or ΔR2-GFP (n = 15) exhibiting puncta and cells with high expression of GFP only (n = 15). The center and edges of the box represent the median with the first and third quartiles. The p

values were calculated from a two-tailed t-test. **E.** The maximum fractional recovery after photobleaching of cells expressing N-GFP (n = 28), ΔR2-GFP (n = 15) or GFP-only (n = 15). The center and edges of the box represent the median with the first and third quartiles. The p values were calculated from a two-tailed t-test. **F.** U2OS cells stably expressing N-Clover form condensates in the cytoplasm. (Left) Cells co-transfected with an N-MS2 expression plasmid exhibit colocalization of N-Clover and Cy3-MS2 FISH signal in N-condensates (N = 10 out of 55 cells, two technical replicates). (Right) The Cy3-MS2 FISH probe does not partition into N condensates in untransfected cells (N = 55 cells, two technical replicates). Inset scale bar is 2 μm. **G.** Model for remodeling of viral RNA genome by the SARS-CoV-2 N protein. The N protein packages viral genomic RNA through phase separation, which may facilitate efficient replication of the genomic RNA and the formation of the enveloped virus.

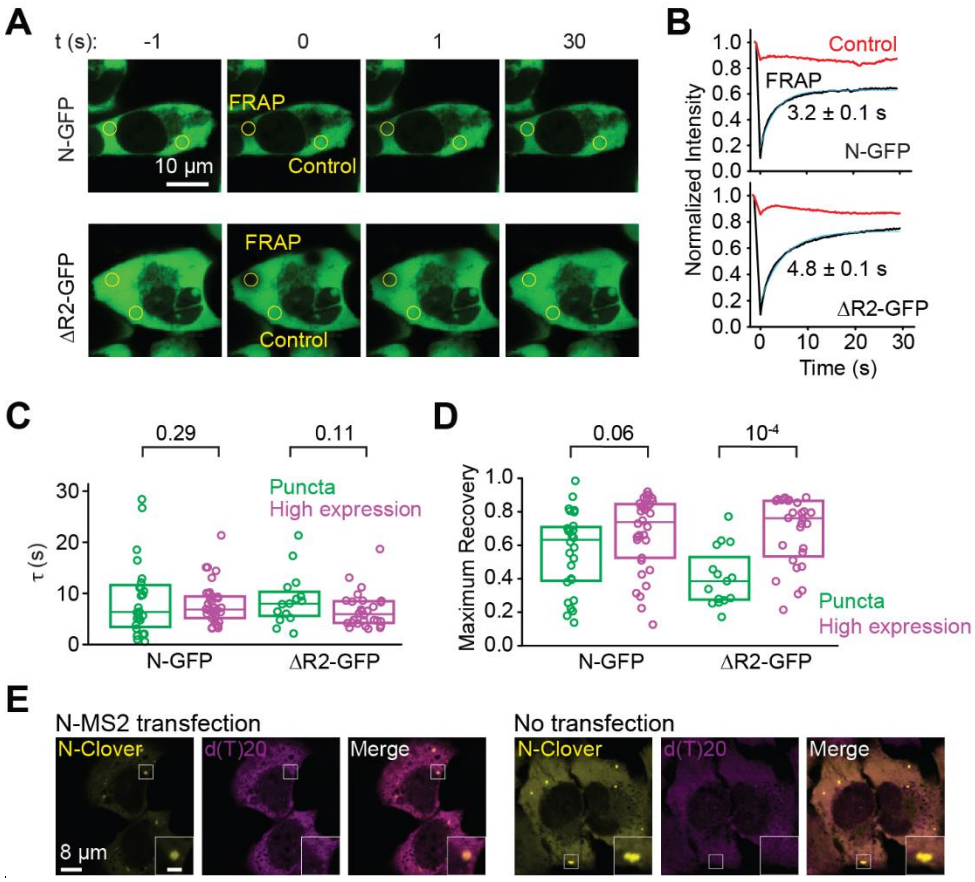

Figure 33. Dynamics of N protein in high-expressing cells. A. Representative FRAP imaging of a cell with high expression of N-GFP or ΔR2-GFP. Circles show the photobleached and control (not bleached) regions. **B.** Fluorescence recovery signals of the N protein in the bleached versus the control regions. The solid curve represents a single exponential fit to reveal the recovery lifetime (τ, ±95% confidence interval). **C.** The distribution of fluorescence recovery lifetimes of cells expressing N-GFP or ΔR2-GFP and exhibiting either puncta or high expression (from left to right, n=28, 34, 15, and 28). The center and edges of the box represent the median with the first and third quartiles. The p values were calculated from a two-tailed t-test. **D.** The maximum fractional recovery after photobleaching of cells expressing N-GFP or ΔR2-GFP and exhibiting either puncta or high expression (from left to right, n=28, 34, 15, and 28). The center and edges of the box represent the median with the first and third quartiles. The p values were calculated from a two-tailed t-test. **E.** U2OS cells stably expressing N-Clover form condensates in the cytoplasm. Cy3-d(T)20 FISH probe targeting the polyA tails of RNA transcripts is uniformly distributed in cells exhibiting N condensates with (left) or without (right) expression of N-MS2 (N = 40 cells, two technical replicates). Inset scale bar is 2 µm. **G.** Model for remodeling of viral RNA genome by the SARS-CoV-2 N protein. The N protein packages viral genomic RNA through phase separation, which may facilitate efficient replication of the genomic RNA and the formation of the enveloped virus.

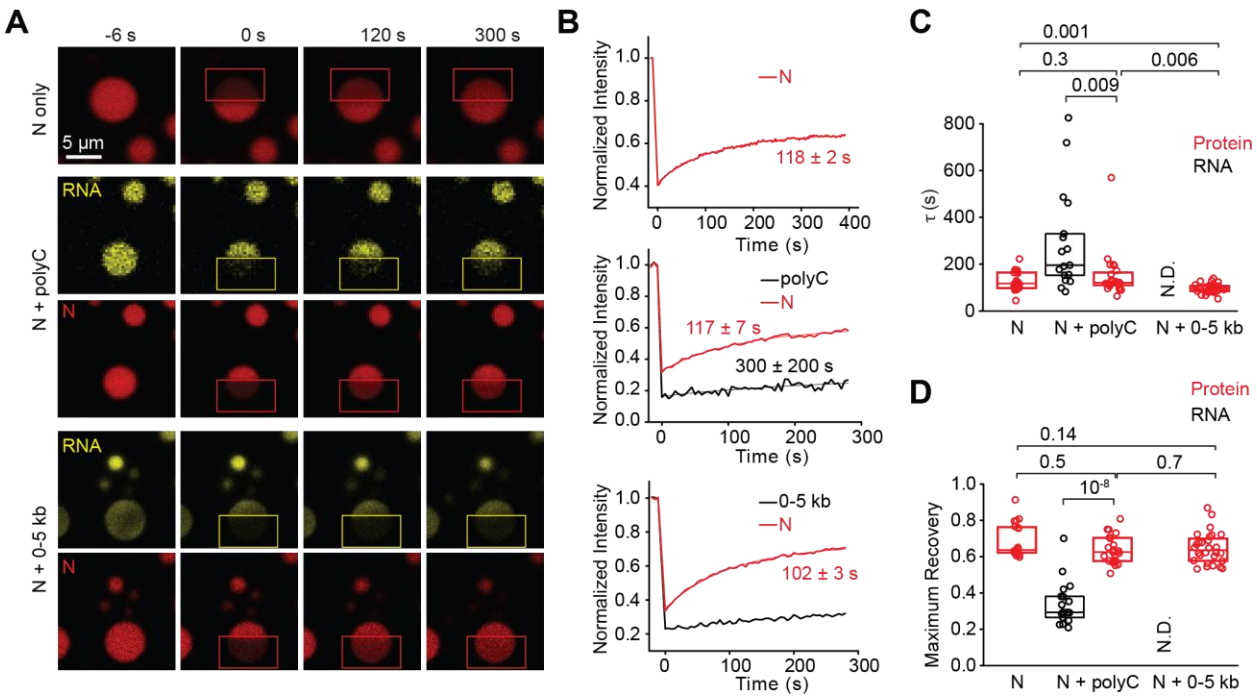

Figure 34. Dynamics of N protein in the presence of low salt (50 mM NaCl). A. Representative FRAP imaging of N protein only, or in the presence of polyC or 0-5 kb viral RNA. The concentrations of N, polyC, and 0-5 kb viral RNA were kept at 24 μM, 50 ng/μL, and 18 nM, respectively. Rectangles highlight the photobleached area. **B.** Fluorescence recovery signals of the N protein and RNA in the bleached versus the control regions. Solid curves represent a single exponential fit to reveal the recovery lifetime (τ, ±95% confidence interval). **C.** The distribution of fluorescence recovery lifetimes of the condensates (from left to right, n = 20, 17, 23, 28, and 28). The center and edges of the box represent the median with the first and third quartiles. The p values were calculated from a two-tailed t-test. **D.** The maximum fractional recovery after photobleaching (from left to right, n = 20, 17, 23, 28, and 28). The center and edges of the box represent the median with the first and third quartiles. The p values were calculated from a two-tailed t-test.

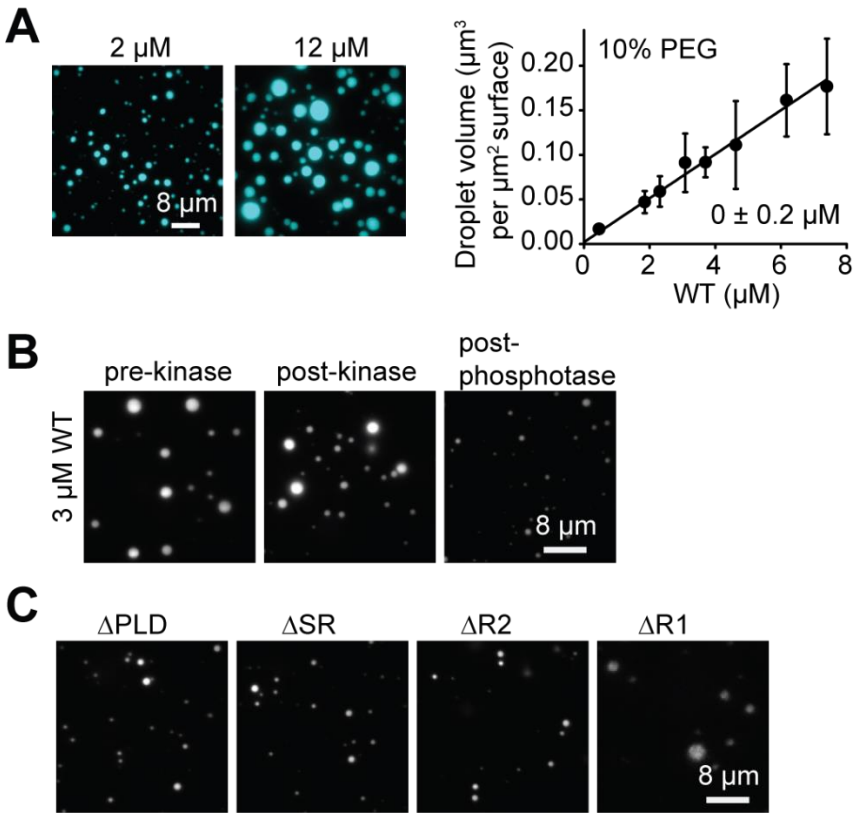

Figure 35. Phase separation of WT N and truncation mutants in the presence of a crowding agent. A. (Left) Images of the LD655-labeled N protein in the presence of 10% PEG and 150 mM NaCl. (Right) The total volume of N-RNA condensates settled per micron squared area on the coverslip (mean ± s.d., n = 20 with two technical replicates). A linear fit (solid line) reveals c_{sat} (± s.e.). **B.** Images of the condensates formed in the presence of LD655-labeled WT N protein before or after phosphatase and kinase treatments in the presence of 10% PEG and 150 mM NaCl. **C.** Images of the LD655-labeled truncated N protein mutants in the presence of 10% PEG and 150 mM NaCl. The protein concentration was set at 3 μM. Assays were performed in the absence of RNA.

Tables Related to Part 2

Fold change (log_2)	p-value	Xlink position 1	Xlink position 2
-2.441	1.05E-04	257	266
-2.226	5.88E-06	169	61
-2.026	8.84E-08	266	347
-1.697	1.23E-09	169	65
-1.679	1.32E-09	169 (phosphoS176)	65
-1.409	1.21E-05	38	61
-1.324	3.20E-05	169	266
-1.079	6.84E-04	169	38
-0.941	3.61E-12	266	299
-0.796	8.00E-07	102	61
-0.765	2.94E-06	266	266
-0.762	3.87E-04	342	38
-0.730	5.18E-07	233	266
-0.729	8.10E-05	266	38
-0.715	0.0034	355	375
-0.648	0.0086	102	266
-0.607	0.0013	342	355
-0.594	2.84E-05	347	38
-0.488	0.030	102	38
-0.450	8.76E-04	266	375
-0.448	1.53E-13	299	355
-0.415	3.30E-06	233	38
-0.382	0.048	38	65
-0.380	2.21E-04	237	266
-0.351	4.41E-04	266	388
-0.277	0.037	266	374
-0.264	2.72E-04	233	233
-0.249	0.0043	375	38
-0.222	0.015	233	248
-0.220	0.034	347	388
-0.185	0.042	342	388
0.103	0.038	233	249
0.271	0.0026	370	375
0.291	6.17E-06	233	237
0.347	6.53E-05	373	375
0.361	3.17E-06	237	388
0.378	5.90E-09	248	256
0.396	5.40E-05	256	342
0.398	1.41E-05	266	373
0.402	1.40E-06	257	342
0.402	9.55E-12	372	375

0.403	0.0014	237	249
0.408	1.05E-07	248	257
0.418	0.0024	370	374
0.428	3.83E-07	372	374
0.455	1.34E-06	266	372
0.522	1.41E-07	256	375
0.540	5.63E-06	256	38
0.578	3.49E-05	266	370
0.578	2.94E-05	257	374
0.580	4.27E-10	257	375
0.610	1.18E-07	237	257
0.641	1.58E-07	237	256
0.668	4.45E-04	248	388
0.670	4.93E-06	256	374
0.857	3.80E-12	237	248
0.902	1.31E-06	372	388
0.909	0.018	370	388
0.941	2.54E-05	248	249
0.976	2.58E-06	373	388

Table 2. Crosslink fold changes with p-values less than 0.05 upon phase separation. The crosslink (xlink) position indicates the amino acid number of the N protein sequence. 169 in Xlink position 1 represents the K169-K65 crosslink. 169 (phosphoS176) represents the phosphorylated form of this peptide, which is phosphorylated at S176.

A

MSDNGPQNQR NAPRITFGGP SDSTGSNQNG ERSGARSKQR RPQGLPNNTA
SWFTALTQHG KEDLKFPRGQ GVPINTNSSP DDQIGYYRRA TRRIRGGDGK
MKDLSPRWYF YYLGTGPEAG LPYGANKDGI IWVATEGALN TPKDHIGTRN
PANNAAIVLQ LPQGTTLPKG FYAEGSRGGS QASSRSSSRS RNSSRNSTPG
SSRGTSPARM AGNGGDAALA LLLLDRLNQL ESKMSGKGQQ QQGQTVTKKS
AAEASKKPRQ KRTATKAYNV TQAFGRRGPE QTQGNFGDQE LIRQGTDYKH
WPQIAQFAPS ASAFFGMSRI GMEVTPSGTW LTYTGAIKLD DKDPNFKDQV
ILLNKHIDAY KTFPPTEPKK DKKKADETQ ALPQRQKKQQ TVTLLPAADL
DDFSKQLQQS MSSADSTQAL EGGGGWSHPQ FEKGGGSGGG SGGGSWSHPQ
FEK

B

Phosphosite	Found in	Site Localization (this study)
S21	This study	Unambiguous
S23	This study,	Unambiguous
T24	This study,	Unambiguous
S26	This study, [172]	Ambiguous (S26|S33)
S33	This study	Ambiguous (S26|S33)
T76	[40, 172]	
S78	This study, [40]	Ambiguous (S78|S79)
S79	This study, [40, 172]	Unambiguous
S105	[40, 172]	
T141	[40]	
T166	[40]	
S176	This study, [40, 172]	Unambiguous
S180	[40, 172]	
S183	[40, 172]	
S184	[40, 172]	
S194	[40, 172]	
S197	[172]	
T198	[40, 172]	
S201	[40, 172]	
S202	[40, 172]	
T205	[40, 172]	
S206	[40, 172]	
T391	[40]	
T393	[173]	
S412	This study	Ambiguous (S412|S413)
S413	This study	Ambiguous (S412|S413)
S441	This study	Unambiguous

Table 3. A compilation of the SARS-CoV2 N protein phosphosites identified by mass spectrometry. A. The coverage map and phosphorylation sites for the N protein detected in

proteomics experiments. Text in red indicates peptide regions that were not detected; green highlight indicates an unambiguous phosphorylation site; yellow highlight indicates an ambiguous phosphorylation site. **B.** Each site identified by my study is characterized as being unambiguously or ambiguously localized based on manual inspection of the product ion series. See Methods for the full protein sequence and a link to supporting evidence. S176 is a phosphorylation site identified on a crosslinked peptide.

A

Gene	Unique Peptides	Percent Coverage	SAF	Species	UniProt Accession	Protein Name
N	57	70.0	104.6	SARS2	P0DTC9	Nucleoprotein
HSPA1A	54	59.3	27.5	HUMAN	P0DMV8	Heat shock 70 kDa protein 1A
G3BP2	38	42.5	19.5	HUMAN	Q9UN86	Ras GTPase-activating protein-binding protein 2
G3BP1	37	60.5	17.8	HUMAN	Q13283	Ras GTPase-activating protein-binding protein 1
HSPA8	27	37.8	9.8	HUMAN	P11142	Heat shock cognate 71 kDa protein
RPS12	2	17.4	8.3	HUMAN	P25398	40S ribosomal protein S12
NA	4	17.3	7.4	PIG*	P00761	Trypsin
TUBB	20	56.3	7.0	HUMAN	P07437	Tubulin beta chain
TUBA1B	19	54.3	6.7	HUMAN	P68363	Tubulin alpha-1B chain
TUBB4B	18	50.6	6.5	HUMAN	P68371	Tubulin beta-4B chain
KRT9	14	52.6	6.1	HUMAN	P35527	Keratin, type I cytoskeletal 9
A2M	27	20.7	5.8	BOVIN*	Q7SIH1	Alpha-2-macroglobulin
HSPA5	17	29.5	5.2	HUMAN	P11021	Endoplasmic reticulum chaperone BiP
NPM1	6	37.1	5.1	HUMAN	P06748	Nucleophosmin

B

Protein 1	Xlink position 1	Protein 2	Xlink position 2	SVM score
G3BP2_HUMAN	50	G3BP2_HUMAN	281	2.17
Nucleoprotein	375	G3BP1_HUMAN	453	1.63
G3BP2_HUMAN	50	Nucleoprotein	266	1.56
G3BP2_HUMAN	50	Nucleoprotein	102	1.26
G3BP2_HUMAN	370	G3BP2_HUMAN	407	1.04
G3BP2_HUMAN	365	G3BP2_HUMAN	407	0.88
G3BP2_HUMAN	50	Nucleoprotein	38	0.71
Nucleoprotein	388	G3BP1_HUMAN	453	0.67
Nucleoprotein	266	G3BP1_HUMAN	393	0.14
Nucleoprotein	374	G3BP1_HUMAN	453	0.00

Table 4. Proteins identified by mass spectrometry in the high salt sample of N protein from human cells. A. The proteins were ranked by spectral abundance factor (SAF). These protein sequences were included in the CLMS search. Asterisks indicate exogenous proteins. **B.** Crosslinks between N and the G3BP1 and G3BP2 stress granule proteins identified in the

qualitative CLMS experiment ranked by a support vector machine (SVM) score. Values are frequency of detection (two conditions with two replicates each). Xlink position indicates the amino acid number of the given protein sequence. SVM indicates the confidence that the crosslink is correctly identified.

Condition	**IC_{50} / EC_{50}**	***n***	**R^2**	**Fig**
N with NaCl, condensate volume	90 ± 1 mM	12.26 ± 0.06	1	1D
N + PolyC with NaCl, condensate volume	152 ± 6 mM	11 ± 9	0.97	1D
N + PolyC with nelfinavir mesylate, condensate number	4.3 ± 1.3 μM	1.9 ± 1.0	0.90	5B

Table 5. Quantitative analysis of N condensates. The parameters of fitting to a dose-response equation in the presence and absence of polyC RNA in response to increasing concentrations of salt or nelfinavir mesylate.

Methods Related to Part 2

N protein purification

A construct for expressing N protein with a C-terminal Strep-tag was obtained from the Krogan lab (UCSF). A C-terminal yBBr labeling site was added for labeling with a fluorescent dye. Truncation mutants, ΔPLD (amino acids 2-30: SDNGPQNQRNAPRITFGGPSDSTGSNQNG), ΔSR (amino acids 183-195, SSRSSSRSRNSSR), ΔR1 (amino acids 235-256, SGKGQQQQGQTVTKKSAAEASK) and ΔR2 (amino acids 369-390, KKDKKKKADETQALPQRQKKQQ) were generated using Gibson cloning (New England Biosciences). Sufficient amounts of DNA were obtained growing 1 L of transfected XL1 Blue *Escherichia coli* cells overnight and performing a gigaprep (Zymo). For protein expression, HEK293S GNTI- cells (RRID: CVCL_A785) were grown in suspension in Freestyle media (Gibco) supplemented with 2% fetal bovine serum (FBS, VWR Inc.) and 1% penicillin-streptomycin (PS, Gemini Bio-products) to 2 million cells/mL. Cells were spun down for 10 min at 1,200 *g* and resuspended in fresh, antibiotic-free media. For 250 mL cells, the transfection solution was created by mixing 1.8 mL polyethylenimine (1mg/mL, pH 7.0 in PBS) dissolved in 20 mL Freestyle media and 0.66 mg DNA dissolved in 20 mL Freestyle media. The mixture was incubated at room temperature for 15 min before being adding it to the cell culture. Transfected cell culture was grown for 72 h at 125 rpm at 37° with 5% CO_2 and 5% humidity.

Cells were then harvested at 4,000 *g* for 10 min and resuspended in 50 mL lysis buffer (50 mM HEPES pH 7.4, 1 M NaCl, 1 mM PMSF, 1 mM DTT, and 1 tablet of protease inhibitor (Sigma)). Lysis was performed using 15 loose and 15 tight plunges of a Wheaton glass dounce. The lysate was clarified using a 45 min, 360,000 *g* spin in a Ti70 rotor. The supernatant was incubated with 1 mL Streptactin sepharose beads (IBA Life Sciences) for 1 h. Beads were washed with 40 mL of lysis buffer followed by 30 mL labelling buffer (50 mM HEPES pH 7.4, 300 mM NaCl, 10 mM $MgCl_2$, 1 mM EGTA, 10% glycerol, 1 mM DTT). Beads were then collected and incubated with purified SFP protein and an LD655 dye functionalized with CoA (Lumidyne) at room temperature for 30 min. Beads were washed with 30 mL labeling buffer. If proteins were to be kinase or phosphatase treated, they were additionally washed with 30 mL kinase (20 mM HEPES pH 7.5, 300mM NaCl, 10 mM MgCl2, 200 μM ATP, 10% glycerol, 1 mM DTT) or phosphatase (20 mM HEPES pH 7.5, 300mM NaCl, 1mM MnSO4, 10% glycerol, 1 mM DTT) buffer. Protein was eluted in 1 mL fractions in its final buffer supplemented with 10 mM desthiobiotin and concentrated using Amicon Ultra 30K concentrators. For kinase and phosphatase treatment, 5 μL of CKII kinase (New England Biolabs) or 2.5 μL λ phosphatase (New England Biolabs) was added per 50 μL concentrated protein, respectively, and the samples were incubated at 30 ℃ for 1 h. Final protein concentration was measured using Bradford reagent, and aliquots were snap-frozen in liquid nitrogen.

In vitro Transcription and RNA Labeling

For in vitro transcription of long viral RNA, the region of interest was first PCR amplified from a plasmid (N plasmid was a generous gift from the Krogan lab [46], 5'UTR plasmid was a

generous gift from the Gladfelter lab [35], and SARS-CoV-2 cDNA plasmid was a generous gift of the Thiel lab [167]) using a forward primer with a T7 polymerase binding site (TAATACGACTCACTATAGGG). The amplified DNA was tested for purity on a 0.8% agarose gel and cleaned up using GlycoBlue and ethanol precipitation. RNA was generated using the HiScribe T7 Quick kit (New England Biolabs) and extracted using trizol and isopropanol precipitation. RNA was Cy3 labeled using a Label IT kit (Mirus Bio), and RNA purity was verified using a 0.8% agarose gel. All RNA structure predictions were done on the RNAfold server [168, 169].

In vitro FRAP

In vitro FRAP assays were performed with a Zeiss 880 Confocal Laser Scanning Microscope equipped with a 100x, 1.4 NA oil immersion objective. Samples were prepared as described above. All in vitro FRAP experiments used the 405, 488, 561, and 633 nm lasers at 100% for bleaching conditions. For N-polyC at 150 mM NaCl, 5 bleaching iterations were used, and images were taken every 2 s for 232 s. For N-SARS RNA at 150 mM NaCl, 100 bleaching iterations were used, and images were taken every 10 s for 8 min. For the experiments conducted at 50 mM NaCl, droplets containing polyC RNA or SARS RNA were bleached with 20 iterations, and images were taken every 5 s for 5 min. For droplets containing N protein only, 20 bleaching iterations were used, and images were collected every 2 s for 6.5 min. All data were acquired using Zeiss Zen 2.3 SP1 FP3 (black) (64bit) Version 14.0.20.201. Data were background corrected, converted to normalized intensity, and fit to an exponential decay function in Origin.

Mass Spectrometry

The crosslinking analysis was performed as described by McGilvray et al. [192] with deviations outlined here. 15 µL 90 µM N protein in the labeling buffer was diluted 1:3 v/v into either water or labeling buffer. The diluted protein formed condensates in 100 mM NaCl in the final buffer, but not in the labeling buffer that contains 300 mM NaCl. Isotopically coded light (H12) and heavy (D12) BS3 (bis(sulfosuccinimidyl)suberate) crosslinkers (Creative Molecules Inc.) were immediately added to the diluted solution. Final concentration of BS3 was 0.8 mM for 100 mM NaCl dilution and 2.5 mM for 300 mM NaCl dilution. Crosslinking was performed for 30 min at room temperature and then quenched with 1 M Tris pH 8.0 buffer. This experiment was performed in duplicate, once with D12-BS3 crosslinking N protein in condensates and H12-BS3 crosslinking soluble N protein and once with the labels reversed. The crosslinked proteins from both channels were pooled before acetone precipitation. Protein was precipitated with acetone overnight at -20 ℃. Protein was pelleted, the supernatant was removed, and the samples were air-dried for 10 min. The pellets were brought up in 8 M urea, reduced with TCEP, alkylated with iodoacetamide, diluted 4-fold, and digested with two rounds of trypsin. Peptides were desalted on a C18 MacroTrap column (Michrom Bioresources) and fractionated on a Superdex Peptide (GE Life Sciences) size-exclusion column. Fractions enriched in crosslinked peptides were dried and resuspended in 0.1% formic acid for mass spectrometry. Later eluting fractions were used for phosphorylation analysis. Each of the replicates was fractionated by size exclusion chromatography (SEC) and each fraction was injected twice, generating 16 mass spectrometry files for analysis.

Mass spectrometry was acquired on an Orbitrap Fusion Lumos coupled with an Easy-Spray nanoelectrospray ion source, a 15 cm x 75 µm PepMap C_{18} column (Thermo), and an M-class NanoAcuity UPLC system (Waters). Liquid chromatography mass spectrometry (LC-MS) runs were 90 min long. Precursor ions were measured in the Orbitrap at 120 k resolution. Selected precursor ions (triply charged and higher) were isolated, and the product ions were measured in the Orbitrap at 30 k resolution. Samples that were analyzed for phosphorylation were run similarly except only HCD product ions spectra were collected and doubly charge precursors were included.

Crosslinked spectra were identified with Protein Prospector 6.2.23 [193] using the combination of DSS/DSS:2H12 at uncleaved Lys residues and protein N-terminus as the crosslinking reagents. The corresponding light and heavy dead-end modifications, incorrect monoisotopic peak assignment (+1Da neutral loss), N-terminal pyroglutamate formation, methionine oxidation, and acetylation and loss of the protein N-terminal methionine were set as variable modifications with three variable modifications per peptide allowed. Trypsin with two missed cleavages was the digestion enzyme and the mass tolerances were 20 and 30 ppm for precursor and product ions. The search database comprised the 14 most abundant proteins (sorted by Spectral Abundance Factor) found in the linear peptide SEC fractions, alongside a decoy database that was ten times longer. Crosslink spectral matches were classified at a 1% false discovery rate (FDR) threshold.

Protein Prospector was used to search for phosphopeptides of the linear peptide fractions using the following parameters: tryptic specificity with 2 missed cleavages, 7 and 15 ppm precursor and product ion tolerances, and carbamidomethyl (C) fixed modification. Variable modifications were as above except that Phospho (STY) was included and only dead-end crosslink modifications were included. Peptides were reported with a maximum expectation value of 0.001, and a minimum prospector score of 15. A site localization in peptide (SLIP) score threshold of 6 was used to determine site-localization. All phosphopeptide spectra reported were manually inspected for evidence of correct site assignment. Where S or T residues are adjacent, the location of the phosphosite is ambiguous (see Table 3). All phosphorylation data can be accessed here (or on the MS-Viewer website with search key "jyovxenjny").

A supplemental search of the crosslinked peak lists was made for crosslinks that also contained a phosphorylation site. This search was identical to the previous crosslinking searches except that Phospho (ST) was included as a variable modification, and the search was limited to the N protein sequence. Search results containing a phosphorylation site were manually assessed. Only phosphorylation sites that had been discovered in the phospho search of the linear peptides were considered.

For quantitating the isotopically labeled crosslinks, peak areas were measured from the extracted precursor ion chromatograms (XICs) using the small molecule interface of Skyline (v20.1.0.155). A Skyline transition list was generated containing the elemental composition of each distinct peptide pair with both light and heavy BS3 modification and in each charge state detected in the Prospector search. Retention times were present in the transition list to help with peak detection. Peptide level measurements were imported into R and summarized at the level of unique residue pairs ("crosslinks"). For each precursor ion, the $\log_2$ ratios of heavy to light

peptides were calculated and mean normalized to 0 for each of the two biological replicates and then transformed into $\log_2$ (water/salt) ratios. Weighted t-tests were performed in R.

Drug Screening and Image Processing

For the FDA-approved drug screen, 75 mL of N protein at 16 µM was purified from HEK293S GNTI- cells. polyC RNA was obtained from Sigma. The FDA-Approved Drug Library (TargetMol) has 1,200 compounds of well-characterized biological activity. 10 mM compounds were stored in 100% dimethyl sulfoxide (DMSO) in 384 well plates. For screening plates, 20 µL imaging buffer (50 mM HEPES pH 7.4, 150 mM NaCl, 5 mM $MgCl_2$, 1 mM EGTA, 1 mM DTT, 1% pluronic) was aliquoted into 384 well, glass-bottomed plates (Greiner Bio-One) and incubated for 5 min. Buffer was removed, and 0.5 µL of 2 mM compounds were stamped into 384 well plates with an Analytik-Jena Cybio Well Vario liquid handler. The final concentration of 7.8 µM N protein and 50 ng/µL polyC RNA were added to each plate. Each well contained a 25 µL mixture and 40 µM compound in the primary screen.

Wells were homogenized with a Bioshake 3000 ELM orbital shaker at 2,400 rpm for 45 s and condensates were allowed to settle for 1 h before imaging. 4 images (each 224 x 167 µm) were taken at 40x with a Molecular Devices ImageXpress Micro High Content Imaging System. Images were analyzed with Metamorph Imaging software calculating condensate count and area. The volume of the condensates was calculated as described above. Data was then uploaded to CCDvault for normalization to DMSO vehicle control wells. Wells that exhibit 3 standard deviations from the untreated sample and as well as others that produced qualitative morphological changes in N condensates were selected for dose-response screening. The same N protein and RNA were tested against candidate drugs using a 10-point serial dilution starting at 40 µM using the same procedure. Dose-response curves were generated by fitting to the equation:

$$y(C) = y_{min} + \frac{(y_{max} - y_{min})}{1 + \left(\frac{C}{EC_{50}}\right)^{n}}$$

where C is the concentration of the drug, n is the Hill coefficient, EC_{50} is the half-maximal response concentration, and y is either the volume of the N condensates or the number of N condensates settled onto per micron squared area on the coverslip.

Dose-Response with SARS-CoV-2 Infected Cells

Vero-E6 (ATCC, CRL-1586) and Calu-3 cells (ATCC HTB-55) were cultured in high glucose DMEM (Gibco) supplemented with 10% FBS (R&D Systems), 1X GlutaMAX (Gibco), and 1X PenStrep (Gibco) at 37°C and 5% CO_2. For screening selected drugs against infected cells, 2500 Vero-E6 (12 µL/well) or 10000 Calu-3 (12 µL/well) were seeded in 384-well white optical-bottom tissue culture plates (Nunc) with the Multidrop Combi liquid handling instrument (Thermo Fisher Scientific). Cells were allowed to recover for 24 h for Vero-E6 and 48 h for Calu-3 at 37°C and 5% CO_2. Dose responses were generated by diluting the compounds using a Cybio Well Vario liquid handler (Analytik Jena), leading to a final concentration of DMSO at 0.4% in the assay plate (v/v). Cells were incubated at 37°C and 5% CO_2 for 1 h before infection. The viral inoculum was prepared such that the final multiplicity of infection (MOI) = 0.05 upon addition of 6 µL/well viral inoculum. After complete CPE was observed in DMSO-treated, infected wells (72 h post-infection (hpi) for Vero-E6 and 96 hpi for Calu-3), the plates were

developed with the CellTiter-Glo 2.0 reagent (Promega) according to the manufacturer's instructions. For Vero-E6, the reagent was diluted 1:1 (v/v) in PBS. Luminescence of developed plates was read on a Spectramax L (Molecular Devices). Each plate contained 24 wells uninfected/DMSO treated cells (100% CPE inhibition) and 24 wells infected/DMSO treated cells (0% CPE inhibition). Average values from those wells were used to normalize data and determine % CPE inhibition for each compound well. To determine the cytotoxicity of the compounds, the same protocol was used but with 6 µL growth media added instead of viral inoculum. The data were plotted and analyzed in Origin and fit to the dose-response function above.

Dose-response determination by $TCID_{50}$ assay

Vero-E6 cells were plated at 10,000 cells per well in Vero growth medium as above in 96-well plates and allowed to recover for 24 h. Drug treatments were prepared in the same medium at 1.33x of final concentration and 0.53% v/v DMSO. Growth medium was removed from cells and replaced with 150 µL medium with drug treatment and incubated for 1 h. Virus in 50 µL medium was added to each well to produce a final MOI of 0.05, final drug concentrations as indicated, and 0.4% v/v DMSO per well. Vehicle-only treatment was used as a positive control for productive viral infection and cytopathic effect, and remdesivir at 50 µM was used as a positive control for viral restriction by drug treatment. Each condition was plated in triplicate. Virus titer following drug treatment was determined by $TCID_{50}$ assay in Vero-E6 cells. At 24 h post infection, 100 µL of supernatant from drug-treated infected cells was removed from each well and used to prepare a 10-fold dilution series in Vero growth medium. 50 µL of the dilution series was added to Vero cells in 96-well plates, prepared as above, in 100 µL medium. Cells were observed for cytopathic effect for 3 days. $TCID_{50}$/mL results were calculated using the Spearman and Kärber method [194].

HEK293T Cell Culture and Transfection

HEK293T (RRID: CVCL_0063) cells were cultured in phenol-negative DMEM media supplemented with 10% FBS and 1% PS at 37° with 5% CO_2. Prior to imaging, cells were transferred to glass-bottomed plates (Nunc Lab-Tek, 0.4 mL working volume) at ~25% confluence and allowed to recover for 24 h. Media was exchanged into phenol-negative DMEM media supplemented with 10% FBS, and cells were transfected with N-GFP, NΔR2-GFP, or GFP only constructs. For each well, 400 ng of DNA was added to 20 µL DMEM media. 1.2 µL FuGENE HD transfection reagent (Promega) was added, and the mixture was incubated at room temperature for 15 min before being added to the cell culture. Cells were allowed to express the constructs for 4 days. 1 h prior to imaging, DRAQ5 stain at 1:1000 was added to the media.

In vivo Imaging and FRAP Assays

In vivo imaging and FRAP measurements were performed with a Zeiss 880 Confocal Laser Scanning Microscope equipped with a 63x, 1.4 NA oil immersion objective. Samples were prepared as described above. All in vivo FRAP experiments used the 405, 488, 561, and 633 nm lasers at 100% for bleaching conditions. Cells expressing N-GFP, ΔR2-GFP, or GFP-only were bleached with 30 iterations, and images were collected continuously at 6.7 Hz for 30 s. All data were acquired using Zeiss Zen 2.3 SP1 FP3. Data were converted to the normalized intensity and fit to an exponential decay function in Origin.

U2OS Cell Culture, Transfection, and RNA fluorescence in situ hybridization

U2OS cells stably expressing SARS-CoV-2 N protein with a C-terminal Clover tag were a gift from the Cleveland lab at UCSD. Cells were cultured in phenol-negative DMEM media supplemented with 10% FBS and 1% PS at 37° with 5% CO_2. Prior to imaging, cells were transferred to glass-bottomed plates (Nunc Lab-Tek, 0.4 mL working volume) at ~25% confluence. Cells were transfected with a construct expressing N-MS2 RNA (the N cDNA sequence with six upstream stop codons and three C-terminal MS2 repeats) as described above. Cells were induced with doxycycline at a final concentration of 1000 ng/mL and allowed to express for 14 h. FISH was performed following the Stellaris RNA fluorescence in situ hybridization (FISH) protocol for adherent cells. Cells were fixed with 3.7% formaldehyde, permeabilized with 70% ethanol, and incubated in the dark at 37 °C overnight with hybridization buffer (Biosearch Technologies) supplemented with 10% deionized formamide and either 125 nM Cy3-(d)T20 oligonucleotides (Gene Link) or 1 µM Cy3-MS2 probe (IDT, sequence AGGCAATTAGGTACCTTAGG) [140]. Cells were washed with wash buffer A (Biosearch Technologies) for 30 min, then incubated with DRAQ5 1:1000 in wash buffer A for 30 min and exchanged into wash buffer B (Biosearch Technologies) before imaging.

Conclusion

The driving forces of LLPS vary across organelles; however, I have shown that both viral RNA and eukaryotic DNA are capable of acting as scaffolds to recruit high local concentrations of low-complexity proteins and form chemically distinct condensates. This ability of nucleic acids and their binding proteins to drive LLPS, thus sequestering molecules and controlling reaction rates, may be a general mechanism for biological organization [195, 196]. New examples of systems that form condensates in vitro are constantly being identified; however, establishing physiological relevance for this behavior remains a challenge [197].

While this work shows that telomeric components have characteristics of biological condensates both in vitro and in cells, a causative relationship between shelterin's ability to phase separate and its protective function still needs to be established. The TRF2 dominant-negative mutant (DNM) presents a possible route for future study. This mutant consists of only the TRFH and Hinge domains, and it causes telomere deprotection when expressed in cells, even when wild-type (WT) TRF2 is still present [117, 198]. Previous studies have proposed that the TRF2 DNM forms heterodimers with WT TRF2, interfering with its recruitment to the telomere [198, 199]. Preliminary data shows that both TRF2 DNM homodimer and DNM-WT heterodimer are phase separation deficient in vitro, suggesting that TRF2 DNM expression may destabilize the telomeric condensate and provide a link between loss of condensation and DDR accessibility. The optogenetic system of on-demand telomere fusion described in this work could be used to measure changes in telomere condensate dynamics. Measuring telomere coalescence frequency and rate in the absence or presence of the TRF2 DNM could reveal changes in condensate behavior induced by the expression of this well-studied mutant, which would strengthen the hypothesis that phase separation contributes to telomere protection.

Additionally, performing in vitro experiments at the single telomere level could help us understand how telomeric condensates are formed and maintained. This work studied in vitro shelterin phase separation using telomeric substrates that are much shorter than physiological length telomeres; shelterin compaction of multi-kilobase telomeres would face steric limitations. To study the dynamics of a single telomeric condensate, a future study could purify genomic telomeres [22, 200], biotinylate them, and fluorescently image DNA and associated proteins [7]. This experiment could test shelterin-mediated compaction of the telomere in relation to both the length of the telomeric substrate and the concentration of individual shelterin components. It would also be able to test whether and how DDR proteins can access a shelterin condensate formed around a single physiological telomere. These experiments would help us understand what biological role phase separation plays in telomere protection and maintenance.

In comparison to telomeres, the link between viral protein phase separation and function is stronger. Recent work showed that a peptide capable of disrupting nucleocapsid phase separation in vitro also inhibits SARS-CoV-2 infection in a mouse model [51]. This peptide decreased viral titer, suggesting a loss of viral replication efficiency, and restored the innate immune response, perhaps by reducing the nucleocapsid protein's inhibition of native stress granules. As both of these mechanisms are related to phase separation, this study provides strong evidence linking LLPS and biological function. Further research on this and other nucleocapsid proteins could reveal phase separation to be a general viral mechanism, which would suggest new directions for

drug development away from the more typical strategies of specifically blocking receptors or binding pockets [136, 137, 201]. Because condensate formation tends to depend on broad motifs more than specific sequences [202, 203], the principles that drive phase separation in one virus may be generalizable across a virus family, even when the coding sequence is not fully conserved. Therefore, experimental approaches that target the phase separation of one viral protein could help prepare us to combat different variants or even different viruses.

Future studies combating phase-separation-driven infections could operate similarly to this one: First, the system of relevant proteins and nucleic acids would be reconstituted in vitro, and phase separation would be used as a readout in high-throughput screening to identify small molecules or peptides that modulate their condensation. Next, live cell imaging could be used to determine whether candidate molecules change the dynamics of viral replication compartments, and animal models would provide measurements of the molecules' effect on viral replication and immune response. Finally, a more in-depth study of the functional groups or binding sites of the molecules using mass spectrometry or crystallography would allow for derivitization of the candidates, and the cycle would repeat. This strategy could be both rapid and adaptive to provide treatments for novel diseases.

Since the discovery of the nucleus in 1833, scientists of all disciplines have tried to explain how cells organize themselves. From characterizing membrane-bound organelles, to studying how the cytoskeleton directs intracellular traffic, to understanding how the multivalent interactions of intrinsically disordered proteins can form chemical compartments, we aim to build a more complete model of cellular structure. While phase separation in biology is a relatively new field, it has already recontextualized the way we think about numerous cellular processes, and I look forward to seeing how this model of cellular organization continues to develop.

References

1. Shin Y, Brangwynne CP. Liquid phase condensation in cell physiology and disease. Science. 2017;357(6357). Epub 2017/09/25. doi: 10.1126/science.aaf4382.

2. Banani SF, Lee HO, Hyman AA, Rosen MK. Biomolecular condensates: organizers of cellular biochemistry. Nat Rev Mol Cell Biol. 2017;18(5):285-98. Epub 2017/02/23. doi: 10.1038/nrm.2017.7.

3. Altmeyer M, Neelsen KJ, Teloni F, Pozdnyakova I, Pellegrino S, Grofte M, et al. Liquid demixing of intrinsically disordered proteins is seeded by poly(ADP-ribose). Nat Commun. 2015;6:8088. Epub 2015/08/20. doi: 10.1038/ncomms9088.

4. Riback JA, Zhu L, Ferrolino MC, Tolbert M, Mitrea DM, Sanders DW, et al. Composition-dependent thermodynamics of intracellular phase separation. Nature. 2020;581(7807):209-14. Epub 2020/05/15. doi: 10.1038/s41586-020-2256-2.

5. Alberti S, Gladfelter A, Mittag T. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. Cell. 2019;176(3):419-34. Epub 2019/01/27. doi: 10.1016/j.cell.2018.12.035.

6. Strom AR, Emelyanov AV, Mir M, Fyodorov DV, Darzacq X, Karpen GH. Phase separation drives heterochromatin domain formation. Nature. 2017;547(7662):241-5. Epub 2017/06/22. doi: 10.1038/nature22989.

7. Larson AG, Elnatan D, Keenen MM, Trnka MJ, Johnston JB, Burlingame AL, et al. Liquid droplet formation by HP1alpha suggests a role for phase separation in heterochromatin. Nature. 2017;547(7662):236-40. Epub 2017/06/22. doi: 10.1038/nature22822.

8. Gibson BA, Doolittle LK, Schneider MWG, Jensen LE, Gamarra N, Henry L, et al. Organization of Chromatin by Intrinsic and Regulated Phase Separation. Cell. 2019;179(2):470-84.e21. Epub 2019/09/24. doi: 10.1016/j.cell.2019.08.037.

9. Palm W, de Lange T. How shelterin protects mammalian telomeres. Annu Rev Genet. 2008;42:301-34. Epub 2008/08/06. doi: 10.1146/annurev.genet.41.110306.130350.

10. Galati A, Micheli E, Cacchione S. Chromatin structure in telomere dynamics. Front Oncol. 2013;3:46. Epub 2013/03/09. doi: 10.3389/fonc.2013.00046.

11. O'Connor MS, Safari A, Liu D, Qin J, Songyang Z. The human Rap1 protein complex and modulation of telomere length. J Biol Chem. 2004;279(27):28585-91. Epub 2004/04/22. doi: 10.1074/jbc.M312913200.

12. Janoušková E, Nečasová I, Pavloušková J, Zimmermann M, Hluchý M, Marini V, et al. Human Rap1 modulates TRF2 attraction to telomeric DNA. Nucleic Acids Res. 2015;43(5):2691-700. Epub 2015/02/14. doi: 10.1093/nar/gkv097.

13. Bernal A, Tusell L. Telomeres: Implications for Cancer Development. Int J Mol Sci. 2018;19(1). Epub 2018/01/20. doi: 10.3390/ijms19010294.

14. Herbig U, Jobling WA, Chen BP, Chen DJ, Sedivy JM. Telomere shortening triggers senescence of human cells through a pathway involving ATM, p53, and p21(CIP1), but not p16(INK4a). Mol Cell. 2004;14(4):501-13. Epub 2004/05/20. doi: 10.1016/s1097-2765(04)00256-4.

15. Karlseder J, Smogorzewska A, de Lange T. Senescence induced by altered telomere state, not telomere loss. Science. 2002;295(5564):2446-9. doi: 10.1126/science.1069523.

16. Bernardes de Jesus B, Blasco MA. Telomerase at the intersection of cancer and aging. Trends Genet. 2013;29(9):513-20. Epub 20130719. doi: 10.1016/j.tig.2013.06.007.

17. Martínez P, Blasco MA. Telomeric and extra-telomeric roles for telomerase and the telomere-binding proteins. Nat Rev Cancer. 2011;11(3):161-76. doi: 10.1038/nrc3025.

18. de Lange T. Shelterin-Mediated Telomere Protection. Annu Rev Genet. 2018;52:223-47. Epub 2018/09/13. doi: 10.1146/annurev-genet-032918-021921.

19. Wang Y, Ghosh G, Hendrickson EA. Ku86 represses lethal telomere deletion events in human somatic cells. Proc Natl Acad Sci U S A. 2009;106(30):12430-5. Epub 20090706. doi: 10.1073/pnas.0903362106.

20. Okamoto K, Bartocci C, Ouzounov I, Diedrich JK, Yates JR, 3rd, Denchi EL. A two-step mechanism for TRF2-mediated chromosome-end protection. Nature. 2013;494(7438):502-5. doi: 10.1038/nature11873.

21. Doksani Y, Wu JY, de Lange T, Zhuang X. Super-Resolution Fluorescence Imaging of Telomeres Reveals TRF2-Dependent T-loop Formation. Cell. 2013;155(2):345-56. Epub 2013/10/15. doi: 10.1016/j.cell.2013.09.048.

22. Griffith JD, Comeau L, Rosenfield S, Stansel RM, Bianchi A, Moss H, et al. Mammalian telomeres end in a large duplex loop. Cell. 1999;97(4):503-14.

23. Bower BD, Griffith JD. TRF1 and TRF2 differentially modulate Rad51-mediated telomeric and nontelomeric displacement loop formation in vitro. Biochemistry. 2014;53(34):5485-95. Epub 2014/08/15. doi: 10.1021/bi5006249.

24. de Lange T. Shelterin: the protein complex that shapes and safeguards human telomeres. Genes Dev. 2005;19(18):2100-10. doi: 10.1101/gad.1346005.

25. Chiba K, Lorbeer FK, Shain AH, McSwiggen DT, Schruf E, Oh A, et al. Mutations in the promoter of the telomerase gene TERT contribute to tumorigenesis by a two-step mechanism. Science. 2017;357(6358):1416-20. Epub 2017/08/19. doi: 10.1126/science.aao0535.

26. Smogorzewska A, van Steensel B, Bianchi A, Oelmann S, Schaefer MR, Schnapp G, et al. Control of human telomere length by TRF1 and TRF2. Mol Cell Biol. 2000;20(5):1659-68. doi: 10.1128/mcb.20.5.1659-1668.2000.

27. Cech TR. Beginning to understand the end of the chromosome. Cell. 2004;116(2):273-9. doi: 10.1016/s0092-8674(04)00038-8.

28. Cong Y, Ulasli M, Schepers H, Mauthe M, V'Kovski P, Kriegenburg F, et al. Nucleocapsid Protein Recruitment to Replication-Transcription Complexes Plays a Crucial Role in Coronaviral Life Cycle. J Virol. 2020;94(4). Epub 2019/11/30. doi: 10.1128/jvi.01925-19.

29. Snijder EJ, Limpens R, de Wilde AH, de Jong AWM, Zevenhoven-Dobbe JC, Maier HJ, et al. A unifying structural and functional model of the coronavirus replication organelle: Tracking down RNA synthesis. PLoS Biol. 2020;18(6):e3000715. Epub 2020/06/09. doi: 10.1371/journal.pbio.3000715.

30. Carlson CR, Asfaha JB, Ghent CM, Howard CJ, Hartooni N, Safari M, et al. Phosphoregulation of Phase Separation by the SARS-CoV-2 N Protein Suggests a Biophysical Basis for its Dual Functions. Mol Cell. 2020;80(6):1092-103.e4. Epub 2020/11/29. doi: 10.1016/j.molcel.2020.11.025.

31. McBride R, van Zyl M, Fielding BC. The coronavirus nucleocapsid is a multifunctional protein. Viruses. 2014;6(8):2991-3018. Epub 2014/08/12. doi: 10.3390/v6082991.

32. Wolff G, Limpens R, Zevenhoven-Dobbe JC, Laugks U, Zheng S, de Jong AWM, et al. A molecular pore spans the double membrane of the coronavirus replication organelle. Science. 2020;369(6509):1395-8. Epub 2020/08/09. doi: 10.1126/science.abd3629.

33. Yao H, Song Y, Chen Y, Wu N, Xu J, Sun C, et al. Molecular Architecture of the SARS-CoV-2 Virus. Cell. 2020;183(3):730-8.e13. Epub 2020/09/28. doi: 10.1016/j.cell.2020.09.018.

34. Klein S, Cortese M, Winter SL, Wachsmuth-Melm M, Neufeldt CJ, Cerikan B, et al. SARS-CoV-2 structure and replication characterized by in situ cryo-electron tomography. Nature communications. 2020;11(1):5885. doi: 10.1038/s41467-020-19619-7.

35. Iserman C, Roden CA, Boerneke MA, Sealfon RSG, McLaughlin GA, Jungreis I, et al. Genomic RNA Elements Drive Phase Separation of the SARS-CoV-2 Nucleocapsid. Mol Cell. 2020;80(6):1078-91.e6. Epub 2020/12/09. doi: 10.1016/j.molcel.2020.11.041.

36. Roden CA, Dai Y, Seim I, Lee M, Sealfon R, McLaughlin GA, et al. Double-stranded RNA drives SARS-CoV-2 nucleocapsid protein to undergo phase separation at specific temperatures. bioRxiv. 2021:2021.06.14.448452. doi: 10.1101/2021.06.14.448452.

37. Kim D, Lee JY, Yang JS, Kim JW, Kim VN, Chang H. The Architecture of SARS-CoV-2 Transcriptome. Cell. 2020;181(4):914-21.e10. Epub 20200423. doi: 10.1016/j.cell.2020.04.011.

38. Wu CH, Chen PJ, Yeh SH. Nucleocapsid phosphorylation and RNA helicase DDX1 recruitment enables coronavirus transition from discontinuous to continuous transcription. Cell Host Microbe. 2014;16(4):462-72. Epub 2014/10/10. doi: 10.1016/j.chom.2014.09.009.

39. Caruso IP, dos Santos Almeida V, Juliani do Amaral M, de Andrade GC, de Araújo GR, de Araújo TS, et al. Structure insights, thermodynamic profiles, dsDNA melting activity, and liquid-liquid phase separation of the SARS-CoV-2 nucleocapsid N-terminal domain binding to DNA. bioRxiv. 2021:2021.07.21.453232. doi: 10.1101/2021.07.21.453232.

40. Bouhaddou M, Memon D, Meyer B, White KM, Rezelj VV, Correa Marrero M, et al. The Global Phosphorylation Landscape of SARS-CoV-2 Infection. Cell. 2020;182(3):685-712.e19. Epub 2020/07/10. doi: 10.1016/j.cell.2020.06.034.

41. Guseva S, Milles S, Jensen MR, Salvi N, Kleman JP, Maurin D, et al. Measles virus nucleo- and phosphoproteins form liquid-like phase-separated compartments that promote nucleocapsid assembly. Sci Adv. 2020;6(14):eaaz7095. Epub 2020/04/10. doi: 10.1126/sciadv.aaz7095.

42. Heinrich BS, Maliga Z, Stein DA, Hyman AA, Whelan SPJ. Phase Transitions Drive the Formation of Vesicular Stomatitis Virus Replication Compartments. mBio. 2018;9(5). Epub 2018/09/06. doi: 10.1128/mBio.02290-17.

43. Monette A, Niu M, Chen L, Rao S, Gorelick RJ, Mouland AJ. Pan-retroviral Nucleocapsid-Mediated Phase Separation Regulates Genomic RNA Positioning and Trafficking. Cell Rep. 2020;31(3):107520. Epub 2020/04/23. doi: 10.1016/j.celrep.2020.03.084.

44. Chen H, Cui Y, Han X, Hu W, Sun M, Zhang Y, et al. Liquid-liquid phase separation by SARS-CoV-2 nucleocapsid protein and RNA. Cell Res. 2020;30. doi: 10.1038/s41422-020-00408-2.

45. Zhao D, Xu W, Zhang X, Wang X, Ge Y, Yuan E, et al. Understanding the phase separation characteristics of nucleocapsid protein provides a new therapeutic opportunity against SARS-CoV-2. Protein Cell. 122021. p. 734-40.

46. Gordon DE, Jang GM, Bouhaddou M, Xu J, Obernier K, White KM, et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature. 2020;583(7816):459-68. Epub 2020/05/01. doi: 10.1038/s41586-020-2286-9.

47. Savastano A, Ibáñez de Opakua A, Rankovic M, Zweckstetter M. Nucleocapsid protein of SARS-CoV-2 phase separates into RNA-rich polymerase-containing condensates. Nat Commun. 2020;11(1):6041. Epub 2020/11/29. doi: 10.1038/s41467-020-19843-1.

48. Wang J, Shi C, Xu Q, Yin H. SARS-CoV-2 nucleocapsid protein undergoes liquid-liquid phase separation into stress granules through its N-terminal intrinsically disordered region. Cell Discov. 72021. p. 5.

49. Nabeel-Shah S, Lee H, Ahmed N, Marcon E, Farhangmehr S, Pu S, et al. SARS-CoV-2 Nucleocapsid protein attenuates stress granule formation and alters gene expression via direct interaction with host mRNAs. bioRxiv. 2020:2020.10.23.342113. doi: 10.1101/2020.10.23.342113.

50. Samavarchi-Tehrani P, Abdouni H, Knight JDR, Astori A, Samson R, Lin Z-Y, et al. A SARS-CoV-2 – host proximity interactome. bioRxiv. 2020:2020.09.03.282103. doi: 10.1101/2020.09.03.282103.

51. Wang S, Dai T, Qin Z, Pan T, Chu F, Lou L, et al. Targeting liquid-liquid phase separation of SARS-CoV-2 nucleocapsid protein promotes innate antiviral immunity by elevating MAVS activity. Nature cell biology. 2021;23(7):718-32. doi: 10.1038/s41556-021-00710-0.

52. Maciejowski J, de Lange T. Telomeres in cancer: tumour suppression and genome instability. Nat Rev Mol Cell Biol. 2017;18(3):175-86. Epub 2017/01/18. doi: 10.1038/nrm.2016.171.

53. d'Adda di Fagagna F, Reaper PM, Clay-Farrace L, Fiegler H, Carr P, Von Zglinicki T, et al. A DNA damage checkpoint response in telomere-initiated senescence. Nature. 2003;426(6963):194-8. Epub 2003/11/11. doi: 10.1038/nature02118.

54. Wellinger RJ, Zakian VA. Everything you ever wanted to know about Saccharomyces cerevisiae telomeres: beginning to end. Genetics. 2012;191(4):1073-105. doi: 10.1534/genetics.111.137851.

55. Harley CB, Futcher AB, Greider CW. Telomeres shorten during ageing of human fibroblasts. Nature. 1990;345(6274):458-60. Epub 1990/05/31. doi: 10.1038/345458a0.

56. Blackburn EH. Structure and function of telomeres. Nature. 1991;350(6319):569-73. Epub 1991/04/18. doi: 10.1038/350569a0.

57. Baird DM, Rowson J, Wynford-Thomas D, Kipling D. Extensive allelic variation and ultrashort telomeres in senescent human cells. Nat Genet. 2003;33(2):203-7. Epub 2003/01/23. doi: 10.1038/ng1084.

58. Nandakumar J, Cech TR. Finding the end: recruitment of telomerase to telomeres. Nature reviews Molecular cell biology. 2013;14(2):69-82. Epub 2013/01/10. doi: 10.1038/nrm3505.

59. Ray S, Bandaria JN, Qureshi MH, Yildiz A, Balci H. G-quadruplex formation in telomeres enhances POT1/TPP1 protection against RPA binding. Proceedings of the National Academy of Sciences of the United States of America. 2014;111(8):2990-5. doi: 10.1073/pnas.1321436111.

60. Maestroni L, Matmati S, Coulon S. Solving the Telomere Replication Problem. Genes (Basel). 2017;8(2). Epub 2017/02/02. doi: 10.3390/genes8020055.

61. Stansel RM, de Lange T, Griffith JD. T-loop assembly in vitro involves binding of TRF2 near the 3' telomeric overhang. Embo j. 2001;20(19):5532-40. Epub 2001/09/28. doi: 10.1093/emboj/20.19.5532.

62. Kar A, Willcox S, Griffith JD. Transcription of telomeric DNA leads to high levels of homologous recombination and t-loops. Nucleic Acids Res. 2016;44(19):9369-80. Epub 20160907. doi: 10.1093/nar/gkw779.

63. Bandaria JN, Qin P, Berk V, Chu S, Yildiz A. Shelterin Protects Chromosome Ends by Compacting Telomeric Chromatin. Cell. 2016;164(4):735-46. doi: 10.1016/j.cell.2016.01.036.

64. Janissen R, Arens MMA, Vtyurina NN, Rivai Z, Sunday ND, Eslami-Mossallam B, et al. Global DNA Compaction in Stationary-Phase Bacteria Does Not Affect Transcription. Cell. 2018;174(5):1188-99.e14. Epub 2018/07/31. doi: 10.1016/j.cell.2018.06.049.

65. Timashev LA, Babcock H, Zhuang X, de Lange T. The DDR at telomeres lacking intact shelterin does not require substantial chromatin decompaction. Genes Dev. 2017;31(6):578-89. Epub 2017/04/07. doi: 10.1101/gad.294108.116.

66. Shin Y, Chang YC, Lee DSW, Berry J, Sanders DW, Ronceray P, et al. Liquid Nuclear Condensates Mechanically Sense and Restructure the Genome. Cell. 2018;175(6):1481-91.e13. Epub 2018/12/01. doi: 10.1016/j.cell.2018.10.057.

67. Mattern KA, Swiggers SJ, Nigg AL, Lowenberg B, Houtsmuller AB, Zijlmans JM. Dynamics of protein binding to telomeres in living cells: implications for telomere structure and function. Mol Cell Biol. 2004;24(12):5587-94. Epub 2004/06/01. doi: 10.1128/mcb.24.12.5587-5594.2004.

68. Taylor NO, Wei MT, Stone HA, Brangwynne CP. Quantifying Dynamics in Phase-Separated Condensates Using Fluorescence Recovery after Photobleaching. Biophys J. 2019;117(7):1285-300. Epub 20190830. doi: 10.1016/j.bpj.2019.08.030.

69. Alshareedah I, Kaur T, Banerjee PR. Methods for characterizing the material properties of biomolecular condensates. Methods Enzymol. 2021;646:143-83. Epub 20200722. doi: 10.1016/bs.mie.2020.06.009.

70. Wang X, Kam Z, Carlton PM, Xu L, Sedat JW, Blackburn EH. Rapid telomere motions in live human cells analyzed by highly time-resolved microscopy. Epigenetics Chromatin. 2008;1(1):4. Epub 2008/11/19. doi: 10.1186/1756-8935-1-4.

71. Molenaar C, Wiesmeijer K, Verwoerd NP, Khazen S, Eils R, Tanke HJ, et al. Visualizing telomere dynamics in living mammalian cells using PNA probes. Embo j. 2003;22(24):6631-41. Epub 2003/12/06. doi: 10.1093/emboj/cdg633.

72. Bronshtein I, Kepten E, Kanter I, Berezin S, Lindner M, Redwood AB, et al. Loss of lamin A function increases chromatin dynamics in the nuclear interior. Nat Commun. 2015;6:8044. Epub 2015/08/25. doi: 10.1038/ncomms9044.

73. Feric M, Brangwynne CP. A nuclear F-actin scaffold stabilizes ribonucleoprotein droplets against gravity in large cells. Nat Cell Biol. 2013;15(10):1253-9. Epub 2013/09/03. doi: 10.1038/ncb2830.

74. Lee DSW, Wingreen NS, Brangwynne CP. Chromatin mechanics dictates subdiffusion and coarsening dynamics of embedded condensates. Nat Phys. 2021;17:531–8.

75. Bracha D, Walls MT, Wei MT, Zhu L, Kurian M, Avalos JL, et al. Mapping Local and Global Liquid Phase Behavior in Living Cells Using Photo-Oligomerizable Seeds. Cell. 2018;175(6):1467-80.e13. Epub 2018/12/01. doi: 10.1016/j.cell.2018.10.048.

76. Potts PR, Yu H. The SMC5/6 complex maintains telomere length in ALT cancer cells through SUMOylation of telomere-binding proteins. Nat Struct Mol Biol. 2007;14(7):581-90. Epub 2007/06/26. doi: 10.1038/nsmb1259.

77. Zhang H, Zhao R, Tones J, Liu M, Dilley R, Chenoweth DM, et al. Nuclear body phase separation drives telomere clustering in ALT cancer cells. Mol Biol Cell. 2020:mbcE19100589. Epub 2020/06/25. doi: 10.1091/mbc.E19-10-0589.

78. Min J, Wright WE, Shay JW. Clustered telomeres in phase-separated nuclear condensates engage mitotic DNA synthesis through BLM and RAD52. Genes Dev. 2019;33(13-14):814-27. Epub 2019/06/07. doi: 10.1101/gad.324905.119.

79. Grobelny JV, Godwin AK, Broccoli D. ALT-associated PML bodies are present in viable cells and are enriched in cells in the G(2)/M phase of the cell cycle. J Cell Sci. 2000;113 Pt 24:4577-85.

80. Takai KK, Hooper S, Blackwood S, Gandhi R, de Lange T. In vivo stoichiometry of shelterin components. J Biol Chem. 2010;285(2):1457-67. Epub 2009/10/30. doi: 10.1074/jbc.M109.038026.

81. Sanders DW, Kedersha N, Lee DSW, Strom AR, Drake V, Riback JA, et al. Competing Protein-RNA Interaction Networks Control Multiphase Intracellular Organization. Cell. 2020;181(2):306-24.e28. Epub 2020/04/18. doi: 10.1016/j.cell.2020.03.050.

82. Banerjee PR, Milin AN, Moosa MM, Onuchic PL, Deniz AA. Reentrant Phase Transition Drives Dynamic Substructure Formation in Ribonucleoprotein Droplets. Angew Chem Int Ed Engl. 2017;56(38):11354-9. Epub 2017/05/31. doi: 10.1002/anie.201703191.

83. Soranno A, Incicco JJ, De Bona P, Tomko EJ, Galburt EA, Holehouse AS, et al. Shelterin components modulate nucleic acids condensation and phase separation. bioRxiv. 2021:2021.04.30.442189. doi: 10.1101/2021.04.30.442189.

84. Fairall L, Chapman L, Moss H, de Lange T, Rhodes D. Structure of the TRFH dimerization domain of the human telomeric proteins TRF1 and TRF2. Mol Cell. 2001;8(2):351-61. Epub 2001/09/08. doi: 10.1016/s1097-2765(01)00321-5.

85. Court R, Chapman L, Fairall L, Rhodes D. How the human telomeric proteins TRF1 and TRF2 recognize telomeric DNA: a view from high-resolution crystal structures. Embo Rep. 2005;6(1):39-45. Epub 2004/12/21. doi: 7400314 [pii]

86. Poulet A, Buisson R, Faivre-Moskalenko C, Koelblen M, Amiard S, Montel F, et al. TRF2 promotes, remodels and protects telomeric Holliday junctions. Embo j. 2009;28(6):641-51. Epub 2009/02/07. doi: 10.1038/emboj.2009.11.

87. Biffi G, Tannahill D, Balasubramanian S. An intramolecular G-quadruplex structure is required for binding of telomeric repeat-containing RNA to the telomeric protein TRF2. J Am Chem Soc. 2012;134(29):11974-6. Epub 2012/07/12. doi: 10.1021/ja305734x.

88. Amiard S, Doudeau M, Pinte S, Poulet A, Lenain C, Faivre-Moskalenko C, et al. A topological mechanism for TRF2-enhanced strand invasion. Nature structural & molecular biology. 2007;14(2):147-54. Epub 2007/01/16. doi: nsmb1192 [pii]

89. Liu D, Safari A, O'Connor MS, Chan DW, Laegeler A, Qin J, et al. PTOP interacts with POT1 and regulates its localization to telomeres. Nat Cell Biol. 2004;6(7):673-80. Epub 2004/06/08. doi: 10.1038/ncb1142.

90. Lim CJ, Zaug AJ, Kim HJ, Cech TR. Reconstitution of human shelterin complexes reveals unexpected stoichiometry and dual pathways to enhance telomerase processivity. Nat Commun. 2017;8(1):1075. Epub 2017/10/24. doi: 10.1038/s41467-017-01313-w.

91. Schultz LB, Chehab NH, Malikzay A, Halazonetis TD. p53 binding protein 1 (53BP1) is an early participant in the cellular response to DNA double-strand breaks. J Cell Biol. 2000;151(7):1381-90. doi: 10.1083/jcb.151.7.1381.

92. Takai H, Smogorzewska A, de Lange T. DNA damage foci at dysfunctional telomeres. Curr Biol. 2003;13(17):1549-56. doi: 10.1016/s0960-9822(03)00542-6.

93. Rai R, Chen Y, Lei M, Chang S. TRF2-RAP1 is required to protect telomeres from engaging in homologous recombination-mediated deletions and fusions. Nat Commun. 2016;7:10881. Epub 2016/03/05. doi: 10.1038/ncomms10881.

94. Chu HP, Cifuentes-Rojas C, Kesner B, Aeby E, Lee HG, Wei C, et al. TERRA RNA Antagonizes ATRX and Protects Telomeres. Cell. 2017;170(1):86-101.e16. Epub 2017/07/01. doi: 10.1016/j.cell.2017.06.017.

95. Deng Z, Norseen J, Wiedmer A, Riethman H, Lieberman PM. TERRA RNA binding to TRF2 facilitates heterochromatin formation and ORC recruitment at telomeres. Mol Cell. 2009;35(4):403-13. Epub 2009/09/01. doi: 10.1016/j.molcel.2009.06.025.

96. Wold MS. Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. Annual review of biochemistry. 1997;66:61-92. Epub 1997/01/01. doi: 10.1146/annurev.biochem.66.1.61.

97. Denchi EL, de Lange T. Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. Nature. 2007;448(7157):1068-71. Epub 2007/08/10. doi: 10.1038/nature06065.

98. Gong Y, de Lange T. A Shld1-controlled POT1a provides support for repression of ATR signaling at telomeres through RPA exclusion. Mol Cell. 2010;40(3):377-87. Epub 2010/11/13. doi: 10.1016/j.molcel.2010.10.016.

99. Takai KK, Kibe T, Donigian JR, Frescas D, de Lange T. Telomere protection by TPP1/POT1 requires tethering to TIN2. Mol Cell. 2011;44(4):647-59. Epub 2011/11/22. doi: 10.1016/j.molcel.2011.08.043.

100. Lamarche BJ, Orazio NI, Weitzman MD. The MRN complex in double-strand break repair and telomere maintenance. FEBS Lett. 2010;584(17):3682-95. Epub 2010/07/27. doi: 10.1016/j.febslet.2010.07.029.

101. Myler LR, Gallardo IF, Soniat MM, Deshpande RA, Gonzalez XB, Kim Y, et al. Single-Molecule Imaging Reveals How Mre11-Rad50-Nbs1 Initiates DNA Break Repair. Mol Cell. 2017;67(5):891-8.e4. Epub 2017/09/05. doi: 10.1016/j.molcel.2017.08.002.

102. Frey S, Gorlich D. A saturated FG-repeat hydrogel can reproduce the permeability properties of nuclear pore complexes. Cell. 2007;130(3):512-23. Epub 2007/08/19. doi: 10.1016/j.cell.2007.06.024.

103. Wei MT, Elbaum-Garfinkle S, Holehouse AS, Chen CC, Feric M, Arnold CB, et al. Phase behaviour of disordered proteins underlying low density and high permeability of liquid organelles. Nat Chem. 2017;9(11):1118-25. Epub 2017/10/25. doi: 10.1038/nchem.2803.

104. Söding J, Zwicker D, Sohrabi-Jahromi S, Boehning M, Kirschbaum J. Mechanisms for Active Regulation of Biomolecular Condensates. Trends Cell Biol. 2020;30(1):4-14. Epub 20191118. doi: 10.1016/j.tcb.2019.10.006.

105. Adam N, Degelman E, Briggs S, Wazen RM, Colarusso P, Riabowol K, et al. Telomere analysis using 3D fluorescence microscopy suggests mammalian telomere clustering in hTERT-immortalized Hs68 fibroblasts. Commun Biol. 2019;2:451. Epub 20191204. doi: 10.1038/s42003-019-0692-z.

106. Draskovic I, Arnoult N, Steiner V, Bacchetti S, Lomonte P, Londoño-Vallejo A. Probing PML body function in ALT cells reveals spatiotemporal requirements for telomere recombination. Proc Natl Acad Sci U S A. 2009;106(37):15726-31. Epub 20090826. doi: 10.1073/pnas.0907689106.

107. Nott TJ, Petsalaki E, Farber P, Jervis D, Fussner E, Plochowietz A, et al. Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. Mol Cell. 2015;57(5):936-47. doi: 10.1016/j.molcel.2015.01.013.

108. Wang J, Choi JM, Holehouse AS, Lee HO, Zhang X, Jahnel M, et al. A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. Cell. 2018;174(3):688-99.e16. Epub 2018/07/03. doi: 10.1016/j.cell.2018.06.006.

109. Banani SF, Rice AM, Peeples WB, Lin Y, Jain S, Parker R, et al. Compositional Control of Phase-Separated Cellular Bodies. Cell. 2016;166(3):651-63. Epub 20160630. doi: 10.1016/j.cell.2016.06.010.

110. Ye JZ, Donigian JR, van Overbeek M, Loayza D, Luo Y, Krutchinsky AN, et al. TIN2 binds TRF1 and TRF2 simultaneously and stabilizes the TRF2 complex on telomeres. The Journal of biological chemistry. 2004;279(45):47264-71. Epub 2004/08/19. doi: 10.1074/jbc.M409047200

111. Celli GB, de Lange T. DNA processing is not required for ATM-mediated telomere damage response after TRF2 deletion. Nat Cell Biol. 2005;7(7):712-8. Epub 2005/06/22. doi: 10.1038/ncb1275.

112. Sfeir A, de Lange T. Removal of Shelterin Reveals the Telomere End-Protection Problem. Science. 2012;336(6081):593-7. doi: DOI 10.1126/science.1218498.

113. Shay JW, Wright WE. Senescence and immortalization: role of telomeres and telomerase. Carcinogenesis. 2005;26(5):867-74. Epub 2004/10/09. doi: 10.1093/carcin/bgh296.

114. Cesare AJ, Hayashi MT, Crabbe L, Karlseder J. The telomere deprotection response is functionally distinct from the genomic DNA damage response. Mol Cell. 2013;51(2):141-55. Epub 2013/07/16. doi: 10.1016/j.molcel.2013.06.006.

115. Orun O, Tiber PM, Serakinci N. Partial knockdown of TRF2 increase radiosensitivity of human mesenchymal stem cells. Int J Biol Macromol. 2016;90:53-8. Epub 2015/11/26. doi: 10.1016/j.ijbiomac.2015.10.072.

116. Lackner DH, Durocher D, Karlseder J. A siRNA-based screen for genes involved in chromosome end protection. PLoS One. 2011;6(6):e21407. Epub 2011/07/16. doi: 10.1371/journal.pone.0021407.

117. van Steensel B, Smogorzewska A, de Lange T. TRF2 protects human telomeres from end-to-end fusions. Cell. 1998;92(3):401-13. doi: 10.1016/s0092-8674(00)80932-0.

118. Lazzerini-Denchi E, Sfeir A. Stop pulling my strings - what telomeres taught us about the DNA damage response. Nat Rev Mol Cell Biol. 2016;17(6):364-78. Epub 2016/05/12. doi: 10.1038/nrm.2016.43.

119. Ferro LS, Can S, Turner MA, ElShenawy MM, Yildiz A. Kinesin and dynein use distinct mechanisms to bypass obstacles. eLife. 2019;8. doi: 10.7554/eLife.48629.

120. Schaub JM, Zhang H, Soniat MM, Finkelstein IJ. Assessing Protein Dynamics on Low-Complexity Single-Stranded DNA Curtains. Langmuir. 2018;34(49):14882-90. Epub 2018/07/26. doi: 10.1021/acs.langmuir.8b01812.

121. Caron MC, Sharma AK, O'Sullivan J, Myler LR, Ferreira MT, Rodrigue A, et al. Poly(ADP-ribose) polymerase-1 antagonizes DNA resection at double-strand breaks. Nat Commun. 2019;10(1):2954. Epub 2019/07/06. doi: 10.1038/s41467-019-10741-9.

122. Langelier MF, Planck JL, Servent KM, Pascal JM. Purification of human PARP-1 and PARP-1 domains from Escherichia coli for structural and biochemical analysis. Methods Mol Biol. 2011;780:209-26. Epub 2011/08/27. doi: 10.1007/978-1-61779-270-0_13.

123. Luger K, Rechsteiner TJ, Richmond TJ. Preparation of nucleosome core particle from recombinant histones. Methods Enzymol. 1999;304:3-19. Epub 1999/06/18. doi: 10.1016/s0076-6879(99)04003-3.

124. Dyer PN, Edayathumangalam RS, White CL, Bao Y, Chakravarthy S, Muthurajan UM, et al. Reconstitution of nucleosome core particles from recombinant histones and DNA. Methods Enzymol. 2004;375:23-44. Epub 2004/02/12. doi: 10.1016/s0076-6879(03)75002-2.

125. Lowary PT, Widom J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. J Mol Biol. 1998;276(1):19-42. Epub 1998/03/26. doi: 10.1006/jmbi.1997.1494.

126. Chua EYD, Vogirala VK, Inian O, Wong ASW, Nordenskiöld L, Plitzko JM, et al. 3.9 Å structure of the nucleosome core particle determined by phase-plate cryo-EM. Nucleic Acids Research. 2016;44(17):8013-9. doi: 10.1093/nar/gkw708.

127. Day CA, Kraft LJ, Kang M, Kenworthy AK. Analysis of protein and lipid dynamics using confocal fluorescence recovery after photobleaching (FRAP). Curr Protoc Cytom. 2012;Chapter 2:Unit2.19. Epub 2012/10/09. doi: 10.1002/0471142956.cy0219s62.

128. Yang X, Li Z, Yang L, Lei H, Yu H, Liao Z, et al. Knockdown of telomeric repeat binding factor 2 enhances tumor radiosensitivity regardless of telomerase status. J Cancer Res Clin Oncol. 2015;141(9):1545-52. Epub 20150122. doi: 10.1007/s00432-015-1911-8.

129. Bolte S, Cordelières FP. A guided tour into subcellular colocalization analysis in light microscopy. J Microsc. 2006;224(Pt 3):213-32. doi: 10.1111/j.1365-2818.2006.01706.x.

130. Ollion J, Cochennec J, Loll F, Escudé C, Boudier T. TANGO: a generic tool for high-throughput 3D image analysis for studying nuclear organization. Bioinformatics. 2013;29(14):1840-1. Epub 20130516. doi: 10.1093/bioinformatics/btt276.

131. Jeynes JCG, Geraki K, Jeynes C, Zhaohong M, Bettiol AA, Latorre E, et al. Nanoscale Properties of Human Telomeres Measured with a Dual Purpose X-ray Fluorescence and

Super Resolution Microscopy Gold Nanoparticle Probe. ACS Nano. 2017;11(12):12632-40. Epub 2017/11/02. doi: 10.1021/acsnano.7b07064.

132. Berry J, Weber SC, Vaidya N, Haataja M, Brangwynne CP. RNA transcription modulates phase transition-driven nuclear body assembly. Proc Natl Acad Sci U S A. 2015;112(38):E5237-45. Epub 2015/09/10. doi: 10.1073/pnas.1509317112.

133. Shang J, Ye G, Shi K, Wan Y, Luo C, Aihara H, et al. Structural basis of receptor recognition by SARS-CoV-2. Nature. 2020;581(7807):221-4. Epub 2020/04/01. doi: 10.1038/s41586-020-2179-y. P

134. Chan KK, Dorosky D, Sharma P, Abbasi SA, Dye JM, Kranz DM, et al. Engineering human ACE2 to optimize binding to the spike protein of SARS coronavirus 2. Science. 2020:eabc0870. doi: 10.1126/science.abc0870.

135. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell. 2020;181(2):271-80.e8. Epub 2020/03/07. doi: 10.1016/j.cell.2020.02.052.

136. Mulligan MJ, Lyke KE, Kitchin N, Absalon J, Gurtman A, Lockhart S, et al. Phase I/II study of COVID-19 RNA vaccine BNT162b1 in adults. Nature. 2020;586(7830):589-93. Epub 2020/08/14. doi: 10.1038/s41586-020-2639-4.

137. Sahin U, Muik A, Derhovanessian E, Vogler I, Kranz LM, Vormehr M, et al. COVID-19 vaccine BNT162b1 elicits human antibody and T(H)1 T cell responses. Nature. 2020;586(7830):594-9. Epub 2020/10/01. doi: 10.1038/s41586-020-2814-7.

138. Laude H, Masters PS. The Coronavirus Nucleocapsid Protein. In: Siddell SG, editor. The Coronaviridae. Boston, MA: Springer US; 1995. p. 141-63.

139. Hyman AA, Weber CA, Julicher F. Liquid-liquid phase separation in biology. Annual review of cell and developmental biology. 2014;30:39-58. doi: 10.1146/annurev-cellbio-100913-013325.

140. Jain A, Vale RD. RNA phase transitions in repeat expansion disorders. Nature. 2017;546(7657):243-7. Epub 2017/06/01. doi: 10.1038/nature22386.

141. DW S, N K, DSW L, AR S, V D, JA R, et al. Competing Protein-RNA Interaction Networks Control Multiphase Intracellular Organization. Cell. 2020;181(2). doi: 10.1016/j.cell.2020.03.050.

142. Siu YL, Teoh KT, Lo J, Chan CM, Kien F, Escriou N, et al. The M, E, and N structural proteins of the severe acute respiratory syndrome coronavirus are required for efficient assembly, trafficking, and release of virus-like particles. J Virol. 2008;82(22):11318-30. Epub 2008/08/30. doi: 10.1128/jvi.01052-08.

143. Nikolic J, Le Bars R, Lama Z, Scrima N, Lagaudrière-Gesbert C, Gaudin Y, et al. Negri bodies are viral factories with properties of liquid organelles. Nat Commun. 2017;8(1):58. Epub 2017/07/07. doi: 10.1038/s41467-017-00102-9.

144. Ye Q, West AMV, Silletti S, Corbett KD. Architecture and self-assembly of the SARS-CoV-2 nucleocapsid protein. Protein Sci. 2020;29(9):1890-901. Epub 2020/07/13. doi: 10.1002/pro.3909.

145. Chen CY, Chang CK, Chang YW, Sue SC, Bai HI, Riang L, et al. Structure of the SARS coronavirus nucleocapsid protein RNA-binding dimerization domain suggests a mechanism for helical packaging of viral RNA. J Mol Biol. 2007;368(4):1075-86. Epub 2007/03/24. doi: 10.1016/j.jmb.2007.02.069.

146. Chang CK, Chen CM, Chiang MH, Hsu YL, Huang TH. Transient oligomerization of the SARS-CoV N protein--implication for virus ribonucleoprotein packaging. PLoS One. 2013;8(5):e65045. Epub 2013/05/30. doi: 10.1371/journal.pone.0065045.

147. Boeynaems S, Holehouse AS, Weinhardt V, Kovacs D, Van Lindt J, Larabell C, et al. Spontaneous driving forces give rise to protein-RNA condensates with coexisting phases and complex material properties. Proc Natl Acad Sci U S A. 2019;116(16):7889-98. Epub 2019/03/31. doi: 10.1073/pnas.1821038116.

148. Perlmutter JD, Qiao C, Hagan MF. Viral genome structures are optimal for capsid assembly. Elife. 2013;2:e00632. Epub 2013/06/26. doi: 10.7554/eLife.00632.

149. Cubuk J, Alston JJ, Incicco JJ, Singh S, Stuchell-Brereton MD, Ward MD, et al. The SARS-CoV-2 nucleocapsid protein is dynamic, disordered, and phase separates with RNA. Nature Communications. 2021;12.

150. Ukmar-Godec T, Hutten S, Grieshop MP, Rezaei-Ghaleh N, Cima-Omori MS, Biernat J, et al. Lysine/RNA-interactions drive and regulate biomolecular condensation. Nat Commun. 2019;10(1):2909. Epub 2019/07/04. doi: 10.1038/s41467-019-10792-y.

151. King OD, Gitler AD, Shorter J. The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. Brain Res. 2012;1462:61-80. Epub 2012/03/27. doi: 10.1016/j.brainres.2012.01.016.

152. Hurst KR, Kuo L, Koetzner CA, Ye R, Hsue B, Masters PS. A major determinant for membrane protein interaction localizes to the carboxy-terminal domain of the mouse coronavirus nucleocapsid protein. J Virol. 2005;79(21):13285-97. Epub 2005/10/18. doi: 10.1128/jvi.79.21.13285-13297.2005.

153. Kuo L, Hurst-Hess KR, Koetzner CA, Masters PS. Analyses of Coronavirus Assembly Interactions with Interspecies Membrane and Nucleocapsid Protein Chimeras. J Virol. 2016;90(9):4357-68. Epub 2016/02/19. doi: 10.1128/jvi.03212-15.

154. Chang CK, Hsu YL, Chang YH, Chao FA, Wu MC, Huang YS, et al. Multiple nucleic acid binding sites and intrinsic disorder of severe acute respiratory syndrome coronavirus

nucleocapsid protein: implications for ribonucleocapsid protein packaging. J Virol. 2009;83(5):2255-64. Epub 2008/12/05. doi: 10.1128/jvi.02001-08.

155. Lu S, Ye Q, Singh D, Cao Y, Diedrich JK, Yates JR, et al. The SARS-CoV-2 nucleocapsid phosphoprotein forms mutually exclusive condensates with RNA and the membrane-associated M protein. Nature Communications. 2021;12.

156. Kang S, Yang M, Hong Z, Zhang L, Huang Z, Chen X, et al. Crystal structure of SARS-CoV-2 nucleocapsid protein RNA binding domain reveals potential unique drug targeting sites. Acta Pharm Sin B. 2020;10(7):1228-38. Epub 2020/05/05. doi: 10.1016/j.apsb.2020.04.009.

157. Dinesh DC, Chalupska D, Silhan J, Koutna E, Nencka R, Veverka V, et al. Structural basis of RNA recognition by the SARS-CoV-2 nucleocapsid phosphoprotein. PLoS Pathog. 2020;16(12):e1009100. Epub 2020/12/03. doi: 10.1371/journal.ppat.1009100.

158. Bar-On YM, Flamholz A, Phillips R, Milo R. SARS-CoV-2 (COVID-19) by the numbers. Elife. 2020;9. Epub 2020/04/02. doi: 10.7554/eLife.57309.

159. Jawerth LM, Ijavi M, Ruer M, Saha S, Jahnel M, Hyman AA, et al. Salt-Dependent Rheology and Surface Tension of Protein Condensates Using Optical Traps. Phys Rev Lett. 2018;121(25):258101. doi: 10.1103/PhysRevLett.121.258101.

160. Boyko S, Qi X, Chen TH, Surewicz K, Surewicz WK. Liquid-liquid phase separation of tau protein: The crucial role of electrostatic interactions. J Biol Chem. 2019;294(29):11054-9. Epub 20190516. doi: 10.1074/jbc.AC119.009198.

161. Alberti S, Saha S, Woodruff JB, Franzmann TM, Wang J, Hyman AA. A User's Guide for Phase Separation Assays with Purified Proteins. J Mol Biol. 2018;430(23):4806-20. Epub 2018/06/27. doi: 10.1016/j.jmb.2018.06.038.

162. Van Treeck B, Protter DSW, Matheny T, Khong A, Link CD, Parker R. RNA self-assembly contributes to stress granule formation and defining the stress granule transcriptome. Proc Natl Acad Sci U S A. 2018;115(11):2734-9. Epub 20180226. doi: 10.1073/pnas.1800038115.

163. Conforti R, Ma Y, Morel Y, Paturel C, Terme M, Viaud S, et al. Opposing effects of toll-like receptor (TLR3) signaling in tumors can be therapeutically uncoupled to optimize the anticancer efficacy of TLR3 ligands. Cancer Res. 2010;70(2):490-500. Epub 20100112. doi: 10.1158/0008-5472.can-09-1890.

164. Ma W, Zhen G, Xie W, Mayr C. Unstructured mRNAs form multivalent RNA-RNA interactions to generate TIS granule networks. bioRxiv. 2020:2020.02.14.949503. doi: 10.1101/2020.02.14.949503.

165. Ma W, Mayr C. A Membraneless Organelle Associated with the Endoplasmic Reticulum Enables 3'UTR-Mediated Protein-Protein Interactions. Cell. 2018;175(6):1492-506.e19. Epub 2018/11/20. doi: 10.1016/j.cell.2018.10.007.

166. Perdikari TM, Murthy AC, Ryan VH, Watters S, Naik MT, Fawzi NL. SARS-CoV-2 nucleocapsid protein phase-separates with RNA and with human hnRNPs. EMBO J. 2020;39(24):e106478. Epub 2020/11/18. doi: 10.15252/embj.2020106478.

167. Thi Nhu Thao T, Labroussaa F, Ebert N, V'Kovski P, Stalder H, Portmann J, et al. Rapid reconstruction of SARS-CoV-2 using a synthetic genomics platform. Nature. 2020;582(7813):561-5. Epub 2020/05/05. doi: 10.1038/s41586-020-2294-9.

168. Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL. The Vienna RNA Websuite. Nucleic Acids Research. 2008;36(suppl_2):W70-W4. doi: 10.1093/nar/gkn188.

169. Lorenz R, Bernhart SH, Höner Zu Siederdissen C, Tafer H, Flamm C, Stadler PF, et al. ViennaRNA Package 2.0. Algorithms Mol Biol. 2011;6:26. Epub 2011/11/26. doi: 10.1186/1748-7188-6-26.

170. Chu F, Thornton DT, Nguyen HT. Chemical cross-linking in the structural analysis of protein assemblies. Methods. 2018;144:53-63. Epub 2018/06/02. doi: 10.1016/j.ymeth.2018.05.023.

171. Chen ZA, Rappsilber J. Protein Dynamics in Solution by Quantitative Crosslinking/Mass Spectrometry. Trends Biochem Sci. 2018;43(11):908-20. Epub 2018/10/16. doi: 10.1016/j.tibs.2018.09.003.

172. Davidson AD, Williamson MK, Lewis S, Shoemark D, Carroll MW, Heesom KJ, et al. Characterisation of the transcriptome and proteome of SARS-CoV-2 reveals a cell passage induced in-frame deletion of the furin-like cleavage site from the spike glycoprotein. Genome Medicine. 2020;12(68).

173. Supekar NT, Shajahan A, Gleinich AS, Rouhani D, Heiss C, Azadi P. SARS-CoV-2 Nucleocapsid protein is decorated with multiple N- and O-glycans. bioRxiv. 2020:2020.08.26.269043. doi: 10.1101/2020.08.26.269043.

174. Wu CH, Yeh SH, Tsay YG, Shieh YH, Kao CL, Chen YS, et al. Glycogen synthase kinase-3 regulates the phosphorylation of severe acute respiratory syndrome coronavirus nucleocapsid protein and viral replication. J Biol Chem. 2009;284(8):5229-39. Epub 2008/12/25. doi: 10.1074/jbc.M805747200.

175. Wheeler RJ, Lee HO, Poser I, Pal A, Doeleman T, Kishigami S, et al. Small molecules for modulating protein driven liquid-liquid phase separation in treating neurodegenerative disease. bioRxiv. 2019:721001. doi: 10.1101/721001.

176. Patel A, Lee HO, Jawerth L, Maharana S, Jahnel M, Hein MY, et al. A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. Cell. 2015;162(5):1066-77. Epub 2015/09/01. doi: 10.1016/j.cell.2015.07.047.

177. Nguyenla X, Wehri E, Van Dis E, Biering SB, Yamashiro LH, Stroumza J, et al. Discovery of SARS-CoV-2 antiviral synergy between remdesivir and approved drugs in human lung cells. bioRxiv. 2020:2020.09.18.302398. doi: 10.1101/2020.09.18.302398.

178. Pruijssers AJ, George AS, Schäfer A, Leist SR, Gralinksi LE, Dinnon KH, 3rd, et al. Remdesivir Inhibits SARS-CoV-2 in Human Lung Cells and Chimeric SARS-CoV Expressing the SARS-CoV-2 RNA Polymerase in Mice. Cell Rep. 2020;32(3):107940. Epub 20200707. doi: 10.1016/j.celrep.2020.107940.

179. Zhao H, Wu D, Nguyen A, Li Y, Adão RC, Valkov E, et al. Energetic and structural features of SARS-CoV-2 N-protein co-assemblies with nucleic acids. iScience. 2021;24(6):102523. Epub 20210507. doi: 10.1016/j.isci.2021.102523.

180. Tayeb-Fligelman E, Cheng X, Tai C, Bowler JT, Griner S, Sawaya MR, et al. Inhibition of amyloid formation of the Nucleoprotein of SARS-CoV-2. bioRxiv. 2021. Epub 20210318. doi: 10.1101/2021.03.05.434000.

181. Langdon EM, Qiu Y, Ghanbari Niaki A, McLaughlin GA, Weidmann CA, Gerbich TM, et al. mRNA structure determines specificity of a polyQ-driven phase separation. Science. 2018;360(6391):922-7. Epub 2018/04/14. doi: 10.1126/science.aar7432.

182. He R, Dobie F, Ballantine M, Leeson A, Li Y, Bastien N, et al. Analysis of multimerization of the SARS coronavirus nucleocapsid protein. Biochem Biophys Res Commun. 2004;316(2):476-83. Epub 2004/03/17. doi: 10.1016/j.bbrc.2004.02.074.

183. Peng TY, Lee KR, Tarn WY. Phosphorylation of the arginine/serine dipeptide-rich motif of the severe acute respiratory syndrome coronavirus nucleocapsid protein modulates its multimerization, translation inhibitory activity and cellular localization. Febs j. 2008;275(16):4152-63. Epub 2008/07/18. doi: 10.1111/j.1742-4658.2008.06564.x.

184. Huang W, Ju X, Tian M, Li X, Yu Y, Sun Q, et al. Molecular determinants for regulation of G3BP1/2 phase separation by the SARS-CoV-2 nucleocapsid protein. Cell Discov. 2021;7(1):69. doi: 10.1038/s41421-021-00306-w.

185. Farag A, Wang P, Ahmed M, Sadek H. Identification of FDA Approved Drugs Targeting COVID-19 Virus by Structure-Based Drug Repositioning. ChemRxiv. 2020. doi: 10.26434/chemrxiv.12003930.v1.

186. Kumar D, Chandel V, Raj S, Rathi B. In silico identification of potent FDA approved drugs against Coronavirus COVID-19 main protease: A drug repurposing approach. Chemical Biology Letters. 2020;7(3):10. Epub 2020-03-25.

187. Li Y, Zhang J, Wang N, Li H, Shi Y, Guo G, et al. Therapeutic Drugs Targeting 2019-nCoV Main Protease by High-Throughput Screening. bioRxiv. 2020:2020.01.28.922922. doi: 10.1101/2020.01.28.922922.

188. Mittal L, Kumari A, Srivastava M, Singh M, Asthana S. Identification of potential molecules against COVID-19 main protease through structure-guided virtual screening

approach. J Biomol Struct Dyn. 2020:1-19. Epub 2020/05/13. doi: 10.1080/07391102.2020.1768151.

189. Yamamoto N, Yang R, Yoshinaka Y, Amari S, Nakano T, Cinatl J, et al. HIV protease inhibitor nelfinavir inhibits replication of SARS-associated coronavirus. Biochem Biophys Res Commun. 2004;318(3):719-25. Epub 2004/05/18. doi: 10.1016/j.bbrc.2004.04.083.

190. Dyall J, Coleman CM, Hart BJ, Venkataraman T, Holbrook MR, Kindrachuk J, et al. Repurposing of clinically developed drugs for treatment of Middle East respiratory syndrome coronavirus infection. Antimicrob Agents Chemother. 2014;58(8):4885-93. Epub 2014/05/21. doi: 10.1128/aac.03036-14.

191. Minasov G, Shuvalova, L. , Wiersum, G. , Satchell, K.J.F. 2.05 Angstrom Resolution Crystal Structure of C-terminal Dimerization Domain of Nucleocapsid Phosphoprotein from SARS-CoV-2. 2020;(Center for Structural Genomics of Infectious Diseases (CSGID)). doi: http://doi.org/10.2210/pdb6WJI/pdb.

192. McGilvray PT, Anghel SA, Sundaram A, Zhong F, Trnka MJ, Fuller JR, et al. An ER translocon for multi-pass membrane protein biogenesis. Elife. 2020;9. Epub 2020/08/22. doi: 10.7554/eLife.56889.

193. Trnka MJ, Baker PR, Robinson PJ, Burlingame AL, Chalkley RJ. Matching cross-linked peptide spectra: only as good as the worse identification. Mol Cell Proteomics. 2014;13(2):420-34. Epub 2013/12/18. doi: 10.1074/mcp.M113.034009.

194. Lei C, Yang J, Hu J, Sun X. On the Calculation of TCID(50) for Quantitation of Virus Infectivity. Virol Sin. 362021. p. 141-4.

195. Ditlev JA, Case LB, Rosen MK. Who's In and Who's Out-Compositional Control of Biomolecular Condensates. J Mol Biol. 2018;430(23):4666-84. Epub 20180809. doi: 10.1016/j.jmb.2018.08.003.

196. Boeynaems S, Alberti S, Fawzi NL, Mittag T, Polymenidou M, Rousseau F, et al. Protein Phase Separation: A New Phase in Cell Biology. Trends Cell Biol. 2018;28(6):420-35. Epub 20180327. doi: 10.1016/j.tcb.2018.02.004.

197. McSwiggen DT, Mir M, Darzacq X, Tjian R. Evaluating phase separation in live cells: diagnosis, caveats, and functional consequences. Genes Dev. 2019;33(23-24):1619-34. Epub 20191008. doi: 10.1101/gad.331520.119.

198. Karlseder J, Broccoli D, Dai Y, Hardy S, de Lange T. p53- and ATM-dependent apoptosis induced by telomeres lacking TRF2. Science. 1999;283(5406):1321-5. doi: 10.1126/science.283.5406.1321.

199. Zhang P, Furukawa K, Opresko PL, Xu X, Bohr VA, Mattson MP. TRF2 dysfunction elicits DNA damage responses associated with senescence in proliferating neural cells

and differentiation of neurons. J Neurochem. 2006;97(2):567-81. Epub 20060315. doi: 10.1111/j.1471-4159.2006.03779.x.

200. Jenkins FJ, Kerr CM, Fouquerel E, Bovbjerg DH, Opresko PL. Modified Terminal Restriction Fragment Analysis for Quantifying Telomere Length Using In-gel Hybridization. J Vis Exp. 2017;(125). Epub 20170710. doi: 10.3791/56001.

201. Wang M, Cao R, Zhang L, Yang X, Liu J, Xu M, et al. Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. Cell Res. 2020;30(3):269-71. doi: 10.1038/s41422-020-0282-0.

202. Martin EW, Holehouse AS, Peran I, Farag M, Incicco JJ, Bremer A, et al. Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. Science. 2020;367(6478):694-9. doi: 10.1126/science.aaw8653.

203. Holehouse ASaGGMaGDaBE. Clustering of Aromatic Residues in Prion-like Domains Can Tune the Formation, State, and Organization of Biomolecular Condensates. Biochemistry. 2021;60(47):3566-81. doi: 10.1021/acs.biochem.1c00465 , note = PMID: 34784177.