Human hnRNPA1 reorganizes telomere-bound Replication Protein A

Sophie L. Granger[1], Richa Sharma[2], Vikas Kaushik[3], Mortezaali Razzaghi[1], Masayoshi Honda[1], Divya S. Bhat[1], Marcin Wlodarski[2], Edwin Antony[3], and Maria Spies[1,*]

[1]Department of Biochemistry and Molecular Biology, University of Iowa Carver College of Medicine, 51 Newton Road, Iowa City, IA 52242, USA.

[2]Department of Hematology, St Jude Children's Research Hospital, Memphis, TN 38105, USA.

[3]Department of Biochemistry and Molecular Biology, St. Louis University, School of Medicine, 1250 Carr Lane, St. Louis, MO 63104, USA.

*Correspondence to: maria-spies@uiowa.edu

One Sentence Summary:

At the single-stranded ends of human telomeres, the heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1) binds to and modulates conformational dynamics of the ssDNA binding protein RPA forming a ternary complex which is controlled by TERRA RNA.

Abstract:

Human replication protein A (RPA) is a heterotrimeric ssDNA binding protein responsible for many aspects of cellular DNA metabolism. The binding to and dissociation of the four individual DNA binding domains (DBDs) from DNA result in configurational dynamics of the RPA-DNA complexes which are essential for replacement of RPA by downstream proteins in various cellular metabolic pathways. RPA plays several important functions at telomeres where it binds to and melts telomeric G-quadruplexes, non-canonical DNA structures formed at the G-rich telomeric ssDNA overhangs. Here, we combine single-molecule total internal reflection fluorescence microscopy (smTIRFM), mass photometry (MP) with biophysical and biochemical analyses, of a gain-of-function RPA mutant to demonstrate that heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1) specifically remodels RPA bound to telomeric ssDNA by dampening the RPA configurational dynamics and forming a stable ternary complex. Uniquely among hnRNPA1 target RNAs, TERRA is capable of releasing hnRNPA1 from the RPA-telomeric DNA complex. We speculate that this telomere specific RPA-DNA-hnRNPA1 complex is an important structure in telomere protection.

Introduction

Replication protein A (RPA) coordinates a plethora of DNA metabolic events by binding to virtually all exposed single-strand (ss)DNA in the cell. RPA serves as an interaction hub that recruits over three dozen proteins onto ssDNA, melts secondary DNA structures, activates the DNA damage response, and hands off ssDNA to appropriate downstream proteins[1,2]. RPA functions as a stable heterotrimer composed of RPA1 (70 kDa), RPA2 (32 kDa) and RPA3 (14 kDa) subunits with modular oligosaccharide/oligonucleotide binding (OB) domains spread across RPA1 (OB-F, A, B and C), RPA2 (OB-D), and RPA3 (OB-E)[1]. Functionally, these OB-domains are further classified as DNA-binding domains (DBDs-A, B, C & D) and protein-interaction domains. Since these domains are connected by flexible linkers, the four DBDs allow for dynamic protein-ssDNA interactions, whereby a macroscopically bound RPA cycles between high affinity and low affinity binding modes, and thus can be displaced or remodeled by lower affinity DNA binding proteins acting downstream of RPA[1-6]. RPA binds ssDNA with high affinity (sub-nM Kd) and little sequence specificity, though it displays a preference for pyrimidines over purines[7], as well as for G-rich[8] and telomeric DNA[9].

Human telomeres are nucleoprotein structures made of G-rich (TTAGGG) repeats with single-strand G-rich overhangs and shelterin complex proteins [10,11]. Exposure of the telomeric repeats in single-stranded form results in enrichment of G-quadruplex structures at chromosomal ends causing replication fork stalling and telomere replication stress[12,13] which can contribute to telomere dysfunction, DNA damage response and accelerated aging[14,15]. The presence of RPA at telomeres alleviates the replication challenges, as RPA is capable of unfolding many types of non-canonical DNA structures including G-quadruplexes formed by the telomere G-rich strand[16-18]. RPA activity at telomeres helps to maintain replication of chromosomal ends and to avoid recognition of linear chromosome ends as DNA damage in need of DNA repair machinery[12]. In agreement with a delicate balance of RPA functions at telomeres, we have recently identified germline heterozygous *RPA1*^{E240K} as a dominant gain of function allele associated with telomere biology disorder that manifests in short telomeres and refractory cytopenia of childhood[9]. Purified RPA carrying the E240K mutation in RPA1 binds ssDNA with higher affinity than the wild type protein and displays enhanced capacity to melt telomeric G-quadruplexes.

While the role of RPA in sequestering and protecting ssDNA are well established, its function and regulation at telomeric ssDNA is poorly understood. RPA recruitment to the broken DNA with a single strand overhang ensures proper DNA repair. However, its recruitment to natural chromosome ends can result in deleterious end-to-end fusions. To protect chromosome ends, the

shelterin complex binds telomeres via the RPA-like shelterin protein, Protection of Telomeres 1 (POT1), thereby preventing RPA recruitment and subsequent ATR driven DNA damage repair at telomeres. RPA transiently binds telomere overhangs during DNA replication in S-phase but must be quickly replaced by POT1 to ensure end protection. This process is tightly regulated by heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1), an RNA binding protein that also binds telomeres, which acts in association with telomeric repeat-containing RNA (TERRA) to displace RPA and allow POT1 binding[31]. Specifically, Flynn and colleagues proposed that hnRNPA1 displaces RPA from single-stranded telomeric DNA thus allowing for POT1 binding and telomere protection, and that this activity is inhibited by TERRA RNA in early S-phase when RPA is needed for telomere replication[19].

To determine the molecular mechanism underlying the hnRNPA1/TERRA controlled RPA displacement we combined single-molecule total internal reflection fluorescence microscopy (smTIRFM), mass-photometry, and biochemical studies with a gain of function RPA mutant that causes short telomere syndrome[9]. Unexpectedly, hnRNPA1 was unable to directly compete with RPA for telomeric ssDNA binding. Instead, we observed formation of the ternary complex between RPA, telomeric ssDNA and hnRNPA1. In this complex the two proteins physically interact, and hnRNPA1 alters the contacts between DBDs of RPA and DNA and subdues the RPA's conformational dynamics. While TERRA was a poor competitor of hnRNPA1 pre-bound to telomeric ssDNA, it readily removed hnRNPA1 from the RPA-DNA-hnRNPA1 complex restoring the architecture of the RPA-telomeric DNA complex. The formation of the RPA-DNA-hnRNPA1 complex and its remodeling are specific to telomeres and TERRA. Notably, hnRNPA1 was unable to remodel RPA complexes on the G-quadruplex forming *BCL-2* promoter 1245 sequence, *cMYC* promoter Pu27 sequence, or poly-dT. Likewise, the HIV-1 Exon Splicing Silencer 3 Element RNA was unable to remove hnRNPA1 from the RPA-DNA-hnRNPA1 complex despite strong interaction between this RNA and hnRNPA1[20]. Collectively, our data point towards an intricate choreography of a dynamic complex between telomeric ssDNA, RPA and hnRNPA1 which allows for melting of telomeric G-quadruplexes, and reorganization of the protein-ssDNA complex.

Results:

Human RPA forms a dynamic complex on ssDNA. Gain of function E240K mutation alters configurational dynamics of the ssDNA bound human RPA.

Previously, we speculated that the E240K mutation in the DBD-A of RPA1 extends the DNA binding site explaining tighter binding of this RPA variant to ssDNA[9]. We expected this

extended interaction to alter the microscopic configurational dynamics of mutant RPA, specifically of its DBD-A. To visualize the dynamic interaction of human RPA with ssDNA we produced human RPA and RPA^{E240K} heterotrimers labeled with an environmentally sensitive fluorescent dye MB543 at DBD-A (RPA-DBD-A^{MB543}) or DBD-D (RPA-DBD-A^{MB543})[21]. **Supplemental Figure 1** shows purified labeled proteins. RPA-DNA interaction was monitored using smTIRFM experiments, in which biotinylated ssDNA ($(dT)_{100}$) was tethered to the TIRFM flow cell surface illuminated with 532 nm evanescent field. MB543-labeled protein was then flown in, resulting in appearance of the fluorescent spots reflecting the presence of RPA on surface-tethered ssDNA (**Figure 1A&B**). MB543 proximity to DNA enhances this dye's fluorescence yield[3,22], with different modes of RPA-ssDNA interaction resulting in different levels of fluorescence (**Figure 1**). Our previous study of yeast RPA revealed several conformational states of the RPA-ssDNA complex with highest fluorescence level attributed to the most engaged state of the labeled domain[3]. Similar to its yeast counterpart, human RPA displays configurational dynamics when macroscopically bound to ssDNA (**Figure 1C, Supplemental Figure 2A&B**): both yeast and human RPAs labeled at DBD-A and DBD-D were found in four distinct configurational states characterized by different fluorescence intensities. The microscopic dynamics of human RPA DBD-A and DBD-D, however, was slower than that previously observed for yeast RPA[3] with dwell times about 5-fold longer (**Figure 1, Supplemental Figure 2**). RPA^{E240K}-DBD-A^{MB543} displayed an unexpectedly complex fluorescence trend (**Figure 1E**). Its initial encounter with ssDNA resulted in a gradual increase in the MB543 fluorescence to levels much higher than those observed with wild type RPA-DBD-A^{MB543}, RPA-DBD-D^{MB543}, or mutant RPA^{E240K}-DBD-D^{MB543} (see **Supplemental Figure 2** for examples of representative trajectories). After a period of 30-60 seconds, the MB543 fluorescence in the RPA^{E240K}-DBD-A^{MB543} acquired step-like behavior similar to that observed for the wild type RPA-DBD-A^{MB543}, though much brighter. This enhanced fluorescence may reflect the difference in the electrostatic environment of the dye brought about by the mutation and/or difference in the contacts between DBD-A and ssDNA due to the positively charged lysine in place of the negatively charged glutamic acid.

Our previous work unambiguously showed that the stepwise changes in the yeast RPA-linked MB543 are due to configurational dynamics of the RPA-DNA complex and are not photophysical effects, and that the changes in the fluorescence can be attributed to a single RPA molecule[3]. Nevertheless, we only quantified the dwell times of the fluorescent states after removal of unbound RPA from the TIRFM flow cell. The dwell time distributions for all states were binned and fit with single exponential decay and are summarized in **Figure 1G**. The main difference between the two proteins was in the duration of the least engaged state of the DBD-A, which was

reduced in the RPA^{E240K}-DBD-A^{MB543} to ~0.12 seconds compared to ~5.5 seconds in the wild type RPA-DBD-A^{MB543}. Notably, both the initial gradual increase in the MB543 fluorescence and the dwell times in each state were independent of the RPA^{E240K}-DBD-A^{MB543} concentration, while the number of observed trajectories and individual events increased linearly with increasing protein concentration (**Supplemental Figure S3**) confirming that the observed phenomena and the quantified dwell times reflect the behavior of RPA-DNA complexes containing a single RPA^{E240K}-DBD-A^{MB543} protein. These data support our earlier model that the mutations drives an alternate configuration for DBD-A on ssDNA. In contrast to the DBD-A, the conformational dynamics of the DBD-D was similar for the wild type and mutant protein (**Figure 1 and Supplemental Figure S2**).

Improved association of RPA^{E240K} with telomeric G-quadruplex.

Human RPA has a capacity to melt DNA secondary structures including G-quadruplexes and this activity is enhanced by the E240K mutation[9,16-18]. While the efficient telomeric G-quadruplex unfolding was observed with nearly stoichiometric ratio of RPA-G-quadruplex in solution studies using 10 nM DNA, subnanomolar concentrations were used in these smTIRFM experiments. Under such conditions RPA transiently binds telomeric G-quadruplex folded in the presence of potassium, and, when bound, RPA-DBD-A^{MB543} mostly spends time in the low fluorescence state (**Figure 2, Supplemental Figure S4**). Consistent with the proposed 5′-3′ G-quadruplex unfolding directionality[23], however, we observed high fluorescent states in the beginning of most of the individual trajectories (**Figure 2, Supplemental Figure S4**). Both the wild type RPA and RPA^{E240K} displayed dynamic changes when added to the reaction chamber with immobilized human telomeric G-quadruplex DNA (five ATTGGG repeats and a dsDNA region used for DNA tethering; **Figure 2A**). These changes in fluorescence may be attributed to configurational dynamics of individual RPAs or to simultaneous binding of multiple RPAs. Notably, virtually all trajectories for the wild type RPA displayed protein dissociation or transition into state 1 when excess of RPA was removed from the reaction chamber. In contrast, RPA^{E240K}-DBD-A^{MB543} continued displaying changes in fluorescence upon removal of unbound protein consistent with enhanced stability of its G-quadruplex binding.

RPA and hnRNPA1 physically interact and form ternary complexes on telomeric ssDNA *in vitro* and in cells.

Human hnRNPA1 was reported to act as a key mediator of the RPA to POT1 exchange[19]. To assess the effect of hnRNPA1 on the RPA-telomeric DNA complex we first carried out bulk FRET-based experiments that followed the geometry of the ssDNA in complex with RPA,

hnRNPA1 and both proteins (**Figure 3**). The unstructured ssDNA is extended to nearly contour length when bound by RPA, which can be visualized by incorporating FRET donor (Cy3) and FRET acceptor (Cy5) at the ends of the ssDNA and following change in FRET upon RPA binding (0.22 FRET)[24,25]. Similarly, G-quadruplex unfolding by RPA yields a FRET signal (0.59) consistent with partially extended telomeric ssDNA[9].

Structural[26] and single-molecule analyses[27] showed that hnRNPA1 binds human telomeric ssDNA reorganizing the quadruplex, and proposed that hnRNPA1 binding juxtaposes the individual telomeric repeats. Using FRET-based assays we confirmed that purified hnRNPA1 preferentially binds telomeric DNA ssDNA with an apparent K_d = 73.5±4.3 nM (**Figure 3A**). These assays utilized Cy3/Cy5-labeled, 15 nucleotide long ssDNA containing 2.5 telomeric repeats, and therefore were too short to fold into G-quadruplex structure. The high initial FRET (0.92) for 15-mer substrates was due to high compaction of the free ssDNA under our experimental conditions (150 mM K^+ and 5 mM Mg^{2+}). FRET values at the saturating hnRNPA1 concentrations (0.74) were consistent with the hnRNPA1 destabilizing telomeric G-quadruplex, but bridging the repeats. For reference, RPA binding to and extension of 15-mer telomeric DNA yields FRET of 0.6, and the RPA-mediated extension of dT15 ssDNA yields FRET of 0.42. hnRNPA1 bound telomeric G-quadruplex formed by folding the Cy3/Cy5-labeled ssDNA containing five telomeric repeats in K^+-containing buffer with K_d = 20.6±3.5 nM and FRET of 0.81 at saturating hnRNPA1 concentrations (**Figure 3B**). The five-repeat (30 nucleotide) long telomeric ssDNA was used in these experiments because it corresponds to the binding site of one RPA heterotrimer. When presented with a stoichiometric complex of RPA and telomeric ssDNA (**Figure 3B**, green curve), hnRNPA1 binds to this complex with affinity similar to that of the telomeric G-quadruplex (K_d = 27.4±2.7 nM). The saturating FRET value (0.76) is distinct from that of the telomeric G-quadruplex in complex with hnRNPA1 (0.81) suggesting that RPA is not fully displaced from the complex. Similar FRET values were observed when hnRNPA1 was titrated into nucleoprotein complex containing RPA^{E240K} (**Figure 3B**, orange curve). The affinity of hnRNPA1 for RPA^{E240K}-telomeric DNA complex was ~4-fold lower (K_d = 75.9±25.5 nM) than that for the complex containing wild type protein suggesting potential competition between the telomeric DNA engagement by the DBD-A of RPA and hnRNPA1. Formation of the ternary complex was further confirmed by EMSA experiments (**Supplemental Figure S5**). Complex between hnRNPA1 and h-telG4 DNA was readily detected with two to four molecules of hnRNPA1 being sufficient to shift all h-telG4 (**Supplemental Figure S5A**). When hnRNPA1 was added to the h-telG4 bound RPA, we observed a supershift indicative of the stable ternary complex formation (**Supplemental Figure S5B**). Notably, at high concentrations, hnRNPA1 was also able to bind unstructured ssDNA

(**Supplemental Figure S5A**), however, no formation of the ternary complex was detected (**Supplemental Figure S5B**).

Our FRET-based and EMSA-based analyses pointed towards coexistence of the hnRNPA1 and RPA on the same DNA molecule instead of previously suggested competition. To directly evaluate molecular species involved in the RPA-DNA-hnRNPA1 complex formation we used mass photometry (MP), a label-free single-molecule technique that applies interferometry approach to determine molecular mass of nucleoprotein complexes[28,29] (**Figure 4A-J**). Both, RPA and RPA^{E240K} produced single Gaussian peaks with molecular weighs corresponding to an intact RPA heterotrimer (**Figure 4A&B**), while hnRNPA1 existed in solution as a mixture of monomers and dimers (**Figure 4C**). Addition of the telomeric G-quadruplex DNA (five TTAGGG repeats) increased molecular weight of the RPA and RPA^{E240K} by 8 kDa (**Figure 4D&E**), while heterologous hnRNPA1-telomeric DNA complexes were observed (**Figure 4F**). To separate the effects of DNA binding and G-quadruplex unfolding, the G-quadruplex structure was destabilized by the presence of Li^+ in the buffer. When both, RPA (or RPA^{E240K}) and hnRNPA1 were mixed with the telomeric ssDNA, we observed an additional peak corresponding to the complex containing one DNA molecule, one RPA (or RPA^{E240K}) heterotrimer and several hnRNPA1 molecules (**Figure 4G&H**). In the absence of DNA, we observed some complex formation between RPA and hnRNPA1 (**Figure 4I**), while added together, RPA^{E240K} and hnRNPA1 yielded two peaks corresponding to the two proteins (**Figure 4J**). To confirm that the RPA-hnRNPA1 complex was formed in cells we carried out co-immunoprecipitation assays in induced pluripotent stem cell (iPSC) lines that have either wild-type RPA1 or RPA1 p.E240K. (**Figure 4K**) Notably, both wild type RPA and RPA^{E240K} were found in the hnRNPA1 pull-down. RPA and hnRNPA1 are highly abundant nuclear proteins. Our ability to detect RPA^{E240K} in the hnRNPA1 pull-down, but not in the MP measurements may reflect either difference in affinity, or DNA-mediated complex formation, as the RPA(RPA^{E240K})-hnRNPA1 complex may protect telomeric ssDNA from MNase digestion.

hnRNPA1 constrains the configurational dynamics of RPA on telomeric DNA.

When RPA-DBD-A^{MB543} or RPA^{E240K}-DBD-A^{MB543} was bound to individual DNA molecules containing 3′-ssDNA overhang with five telomeric repeats in G-quadruplex destabilizing buffer containing Li^+, we observed a dynamic behavior similar to that observed on the dT100 ssDNA (**Figure 5 and Supplemental Figure S6**). Addition of hnRNPA1 caused gradual decrease in fluorescence that stabilized at a lower fluorescence for the wild type RPA (**Figure 5C and S6A**), and at an intermediate fluorescence for RPA^{E240K} (**Figure 5D and S6B**), suggesting that first, the RPA dynamics is constrained in the RPA-DNA-hnRNPA1 complex, and second, that DBD-A is

not fully engaging the DNA in this complex. A rarer excursions from the low fluorescence state of RPA-DBD-A^{MB543} suggested that the RPA is still present in the complex. These observations were in an excellent agreement with the FRET and MP data discussed above.

TERRA RNA controls the RPA-ssDNA-hnRNPA1 complex.

We next set to determine how the presence of TERRA RNA affects the RPA-DNA-hnRNPA1 complex. **Figure 6A** shows FRET-based experiments where we pre-incubated the Cy3/Cy5-labeled telomeric G-quadruplex (five telomeric repeats) folded in the presence of K^+ with RPA, hnRNPA1 or both proteins. We then titrated unlabeled TERRA RNA (five UUAGGG repeats) into these complexes. TERRA had no effect on the RPA bound to telomeric DNA (**Figure 6A** filled green circles). Unexpectedly, TERRA also had no effect on the pre-formed complex between hnRNPA1 and telomeric DNA (**Figure 6A** open black circles). However, when hnRNPA1 was titrated into the mixture of 10 nM Cy3/Cy5-labeled telomeric DNA and 100 nM TERRA RNA, higher hnRNPA1 concentrations were needed to achieve G-quadruplex binding, confirming competition between TERRA and telomeric DNA (not shown). In contrast, when TERRA was titrated into the pre-formed RPA-DNA-hnRNPA1 complex we observed hnRNPA1 release and return of the FRET signal to the level corresponding to RPA-DNA complex with an apparent K_d = 12.8±0.3 nM.

To confirm that the same RPA molecule remains associated with telomeric ssDNA during hnRNPA1 binding and after the hnRNPA1 removal by TERRA, we carried out the following smTIRFM experiments (**Figure 6B** and **Supplemental Figure S7**). A biotinylated DNA construct containing five repeats of human telomeric DNA was tethered on the surface of the TIRFM flow cell. RPA-DBD-A^{MB543} or RPA^{E240K}-DBD-A^{MB543} (100 pM) was added, followed by the addition of 50 nM hnRNPA1, and 50 nM (molecules) TERRA RNA. After addition of each component, the flow cell was allowed to equilibrate for five minutes and five short movies (100 frames) were recorded in different regions of the flow cell. The flow cell was illuminated only during recording of the movies to reduce photobleaching. Representative frames shown in **Figure 6B** highlight the appearance of the fluorescent signal upon addition of RPA-DBD-A^{MB543} (dark spots). Fainter fluorescence spots are observed upon addition of hnRNPA1 to the RPA-DNA complex, but the signals recover after addition of TERRA indicating that the RPA-DBD-A^{MB543} remained associated with surface-tethered DNA after addition of hnRNPA1, but its DBD-A was constrained in a dark/less engaged state. The data from five replicates for each condition was quantified to reveal the number of observed trajectories per movie (**Supplemental Figure S7C**), and fluorescence intensity within the selected trajectories across 100 frames of each movie (**Supplemental Figure**

S7D). Addition of TERRA resulted in recovery of the trajectories that show fluorescence that can be distinguished form the background and overall fluorescence intensity in these trajectories. No fluorescence above background was observed in control experiments without MB543-labeled RPA. When the excess of hnRNPA1 was removed by addition of buffer instead of TERRA, we did not observe the RPA-DBD-A^{MB543} fluorescence recovery, confirming stability of the RPA-DNA-hnRNPA1 complex (**Supplemental Figure S7E-G**). Similar to the wild type protein, fluorescence of the RPA^{E240K}-DBD-A^{MB543} was reduced after addition of hnRNPA1 and then recovered upon addition of TERRA (**Supplemental Figure S7H-J**). Similar behavior was observed with RPA-DBD-D^{MB543} (**Figure 6C** and **S7K-M**) suggesting that hnRNPA1 remodels and constrains the whole RPA.

Formation of the RPA-DNA-hnRNPA1 complex is specific for telomeric ssDNA.

To probe whether the hnRNPA1 remodeling of RPA-ssDNA complex is telomere-specific, we preformed FRET-based and EMSA experiments on three additional DNA substrates decorated at termini similar to the human telomeric G-quadruplex with the Cy3 and Cy5 dyes, dT30, *BCL-2* promoter 1245 G-quadruplex[30], and *cMYC* promoter Pu27 G-quadruplex[31]. While RPA was able to bind to and form a stoichiometric 1:1 complex with each of these substrates (**Figure 7**), the FRET level at saturating protein concentration was different for the different DNAs suggesting a difference in configuration between these complexes (**Figure 7A**). In contrast to telomeric G-quadruplex, however, hnRNPA1 was unable to remodel non-telomeric ssDNA structures or their complexes with RPA. These FRET-based analyses were further confirmed in the orthogonal EMSA experiments (**Supplemental Figure S8**). Both RPA and hnRNPA1 were able to bind BCL-2 and Pu27 G-quadruplexes, albeit less robustly compared to telomeric G-quadruplex. Complexes containing both proteins were readily observed on telomeric G-quadruplex, but only detectable at very high hnRNPA1 concentrations on BCL-2 and Pu27 G-quadruplexes.

Telomere specific RPA-ssDNA-hnRNPA1 complex is not sensitive to non-telomeric hnRNPA1-interacting RNAs.

Human hnRNPA1 binds many different cellular and viral RNAs, affects gene expression, RNA splicing, and viral infection. Earlier SELEX experiments identified a preferred RNA sequence, which contained a telomeric-like signature[32]. One of the viral RNA sequences recognized by hnRNPA1 is the HIV viral splicing silencer (ESS3), which hnRNPA1 binds with similar affinity to TERRA RNA[20]. To test whether control of the RPA-DNA-hnRNPA1 complex by TERRA RNA is

unique to its sequence we have compared the effects of TERRA, SELEX-derived RNA and HIV ESS3 RNA on the conformation of the telomeric ssDNA bound by RPA and hnRNPA1 (**Figure 7B**). Similar to TERRA, SELEX-derived RNA was able to remove hnRNPA1 from the complex. HIV ESS3 RNA on the other hand had little effect on the configuration of the telomeric DNA bound by RPA and hnRNPA1 suggesting that in its presence, hnRNPA1 remains in the complex.

Discussion

Five to fifteen kilobases of repetitive telomeric DNA exists at the ends of chromosomes in a double stranded form with a 50- 500 nucleotide ssDNA G-strand overhang. Several molecular events at telomeres may provide opportunities for RPA binding to the G-strand: (1) discontinuous lagging strand synthesis during DNA replication through telomeric regions, (2) transcription of TERRA RNA (reviewed in[33]), and (3) the ssDNA overhang at the end of the telomere (**Figure 8**). While the two former events are transient, important during early S-phase and depend on the RPA ability to melt the telomeric G4 structures, the latter is the site of the competition between RPA and POT1. RPA retained at the telomeric ssDNA may elicit recognition of telomere as a DNA double strand break and DNA damage response through ATR signaling. Both POT1 and RPA have high affinity for telomeric ssDNA. The mechanisms by which RPA targets replicating DNA and the displaced strand of the transcription bubble, while POT1 protects the telomeric ssDNA overhang against RPA-mediated ATR signaling remains unclear. Here, we propose that the exchange of the RPA for POT1-TTP1 occurs via a ternary complex where telomeric ssDNA simultaneously is occupied by RPA and hnRNPA1. Previously it was suggested[19] that hnRNPA1 competes with RPA. Instead, we observe formation of a ternary complex, where both RPA and hnRNPA1 are bound to the same telomeric sequence. While the complex is stable persisting for many minutes, RPA configurational dynamics is constrained with in this complex, with both DBD-A and DBD-D being only partially engaged.

RPA was recently shown to form condensates on ssDNA that facilitate telomere maintenance[34]. In addition to its RNA-binding domains, hnRNPA1 contains a low complexity sequence domain at the C-terminus that endows hnRNPA1 with a propensity to undergo liquid-liquid phase separation[35]. It is tempting to speculate that both, phase separation and/or configuration of the ternary RPA-DNA-hnRNPA1 complex may restrict ATR access to the RPA -coated telomeric DNA and subsequent DNA damage signaling (**Figure 8**).

While unexpected, the ability of TERRA RNA to strip hnRNPA1 from the RPA-telomeric ssDNA-hnRNPA1 ternary complex but not from the binary hnRNPA1-telomeric ssDNA complex makes physiological sense. A broad specificity of hnRNPA1 for different RNA sequences[36,37]

underlies its many functions in cellular RNA metabolism which include regulation of alternative RNA splicing, mRNA transcription and translation, and RNA stability (reviewed in [38]). The preferred hnRNPA1 binding sequence identified by SELEX, UAGGG(A/U) resembles telomeric ssDNA, TERRA RNA, and consensus sequences of vertebrate splice sites[32]. Structurally, hnRNPA1 binding to and bridging RNA segments is likely to involve a similar set of contacts as its complex with telomeric ssDNA[26]. If TERRA were to disrupt all hnRNPA1-RNA/DNA contacts, it would have a deleterious effect on RNA metabolism. It is not surprising therefore, that the contacts between hnRNPA1 and telomeric ssDNA are different in the presence and absence of RPA.

It is notable that the ternary RPA-DNA-hnRNPA1 complex is specific to telomeric G-quadruplex DNA, as hnRNPA1 was unable to remodel RPA bound to non-structured ssDNA, or to G-quadruplexes formed in the promoters of *MYC* and *BCL-2* genes. These non-telomeric quadruplexes play important regulatory functions in gene expression and can also have pathological roles by interfering with DNA replication and repair. RPA presence at and melting of these G-quadruplexes is likely controlled by a range of G-quadruplex binding and unfolding factors. Formation of a stable RPA-hnRNPA1 complex would interfere with the function of these DNA structures. Similarly, hnRNPA1 removal from the ternary complex containing RPA and telomeric ssDNA by numerous hnRNPA1-binding cellular or viral RNAs would drastically reduce the chance of formation of the telomere specific RPA-DNA-hnRNPA1 complex.

Funding: This work was supported by grants from the National Institutes of Health R35GM131704 to MS, GM133967 and GM130746 to E.A., RS was supported by American Society of Hematology Research Training Awards for Fellows and K08 DK134873. MR was supported by a postdoctoral fellowship from the NIH NCI T32 in Free Radicals and Radiation Biology training program CA078586

Author Contributions: SLG: smTIRFM and FRET-based experiments and data analyses, conceptual design of the study; RS: cell-based studies, conceptual design of the study; VK: preparation and characterization of fluorescently-labeled RPA variants; MR: mass photometry; MH: EMSAs, smTIRFM data analysis; DB: FRET-based experiments; MW: conceptual design; EA: fluorescently-labeled RPA variants, data interpretation; MS: conceptual design, data analysis and interpretation. All authors contributed to manuscript preparation and editing.

Competing Financial Interests Statements:

The authors declare no competing financial interests.

Data and Materials Availability

Data, plasmids for protein expression, and code for single-molecule data analysis are available from the corresponding authors upon request.

References:

1 Caldwell, C. C. & Spies, M. Dynamic elements of replication protein A at the crossroads of DNA replication, recombination, and repair. *Crit Rev Biochem Mol Biol* **55**, 482-507, doi:10.1080/10409238.2020.1813070 (2020).

2 Chen, R. & Wold, M. S. Replication protein A: single-stranded DNA's first responder: dynamic DNA-interactions allow replication protein A to direct single-strand DNA intermediates into different pathways for synthesis or repair. *Bioessays* **36**, 1156-1161, doi:10.1002/bies.201400107 (2014).

3 Pokhrel, N. *et al.* Dynamics and selective remodeling of the DNA-binding domains of RPA. *Nat Struct Mol Biol* **26**, 129-136, doi:10.1038/s41594-018-0181-y (2019).

4 Arunkumar, A. I., Stauffer, M. E., Bochkareva, E., Bochkarev, A. & Chazin, W. J. Independent and coordinated functions of replication protein A tandem high affinity single-stranded DNA binding domains. *J Biol Chem* **278**, 41077-41082, doi:10.1074/jbc.M305871200 (2003).

5 Gibb, B. *et al.* Concentration-dependent exchange of replication protein A on single-stranded DNA revealed by single-molecule imaging. *PLoS One* **9**, e87922, doi:10.1371/journal.pone.0087922 (2014).

6 Chen, R., Subramanyam, S., Elcock, A. H., Spies, M. & Wold, M. S. Dynamic binding of replication protein a is required for DNA repair. *Nucleic Acids Res* **44**, 5758-5772, doi:10.1093/nar/gkw339 (2016).

7 Kim, C., Snyder, R. O. & Wold, M. S. Binding properties of replication protein A from human and yeast cells. *Mol Cell Biol* **12**, 3050-3059, doi:10.1128/mcb.12.7.3050-3059.1992 (1992).

8 Prakash, A., Natarajan, A., Marky, L. A., Ouellette, M. M. & Borgstahl, G. E. Identification of the DNA-Binding Domains of Human Replication Protein A That Recognize G-Quadruplex DNA. *J Nucleic Acids* **2011**, 896947, doi:10.4061/2011/896947 (2011).

9 Sharma, R. *et al.* Gain-of-function mutations in RPA1 cause a syndrome with short telomeres and somatic genetic rescue. *Blood* **139**, 1039-1051, doi:10.1182/blood.2021011980 (2022).

10 McElligott, R. & Wellinger, R. J. The terminal DNA structure of mammalian chromosomes. *The EMBO journal* **16**, 3705-3714 (1997).

11 Wright, W. E., Tesmer, V. M., Huffman, K. E., Levene, S. D. & Shay, J. W. Normal human chromosomes have long G-rich telomeric overhangs at one end. *Genes & development* **11**, 2801-2809 (1997).

12 Audry, J. *et al.* RPA prevents G-rich structure formation at lagging-strand telomeres to allow maintenance of chromosome ends. *The EMBO Journal* **34**, 1942-1958, doi:https://doi.org/10.15252/embj.201490773 (2015).

13 Biffi, G., Tannahill, D., McCafferty, J. & Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nature chemistry* **5**, 182-186 (2013).

14 Blackburn, E. H., Epel, E. S. & Lin, J. Human telomere biology: A contributory and interactive factor in aging, disease risks, and protection. *Science* **350**, 1193-1198, doi:doi:10.1126/science.aab3389 (2015).

15 Smogorzewska, A. & de Lange, T. Different telomere damage signaling pathways in human and mouse cells. *The EMBO journal* **21**, 4338-4348 (2002).

16 Ray, S. *et al.* RPA-mediated unfolding of systematically varying G-quadruplex structures. *Biophys J* **104**, 2235-2245, doi:10.1016/j.bpj.2013.04.004 (2013).

17 Fan, J.-H., Bochkareva, E., Bochkarev, A. & Gray, D. M. Circular Dichroism Spectra and Electrophoretic Mobility Shift Assays Show That Human Replication Protein A Binds and Melts Intramolecular G-Quadruplex Structures. *Biochemistry* **48**, 1099-1111, doi:10.1021/bi801538h (2009).

18 Qureshi, M. H., Ray, S., Sewell, A. L., Basu, S. & Balci, H. Replication Protein A Unfolds G-Quadruplex Structures with Varying Degrees of Efficiency. *The Journal of Physical Chemistry B* **116**, 5588-5594, doi:10.1021/jp300546u (2012).

19 Flynn, R. L. *et al.* TERRA and hnRNPA1 orchestrate an RPA-to-POT1 switch on telomeric single-stranded DNA. *Nature* **471**, 532-536, doi:10.1038/nature09772 (2011).

20 Rollins, C., Levengood, J. D., Rife, B. D., Salemi, M. & Tolbert, B. S. Thermodynamic and phylogenetic insights into hnRNP A1 recognition of the HIV-1 exon splicing silencer 3 element. *Biochemistry* **53**, 2172-2184, doi:10.1021/bi500180p (2014).

21 Kuppa, S., Pokhrel, N., Corless, E., Origanti, S. & Antony, E. Generation of Fluorescent Versions of Saccharomyces cerevisiae RPA to Study the Conformational Dynamics of Its ssDNA-Binding Domains. *Methods Mol Biol* **2281**, 151-168, doi:10.1007/978-1-0716-1290-3_9 (2021).

22 Pokhrel, N. *et al.* Monitoring Replication Protein A (RPA) dynamics in homologous recombination through site-specific incorporation of non-canonical amino acids. *Nucleic Acids Res* **45**, 9413-9426, doi:10.1093/nar/gkx598 (2017).

23 Safa, L. *et al.* Binding polarity of RPA to telomeric sequences and influence of G-quadruplex stability. *Biochimie* **103**, doi:10.1016/j.biochi.2014.04.006 (2014).

24 Grimme, J. M. & Spies, M. FRET-based assays to monitor DNA binding and annealing by Rad52 recombination mediator protein. *Methods Mol Biol* **745**, 463-483, doi:10.1007/978-1-61779-129-1_27 (2011).

25 Yates, L. A. *et al.* A structural and dynamic model for the assembly of Replication Protein A on single-stranded DNA. *Nat Commun* **9**, 5447, doi:10.1038/s41467-018-07883-7 (2018).

26 Ding, J. *et al.* Crystal structure of the two-RRM domain of hnRNP A1 (UP1) complexed with single-stranded telomeric DNA. *Genes Dev* **13**, 1102-1115, doi:10.1101/gad.13.9.1102 (1999).

27 Krüger, A. C. *et al.* Interaction of hnRNP A1 with telomere DNA G-quadruplex structures studied at the single molecule level. *Eur Biophys J* **39**, 1343-1350, doi:10.1007/s00249-010-0587-x (2010).

28 Asor, R. & Kukura, P. Characterising biomolecular interactions and dynamics with mass photometry. *Curr Opin Chem Biol* **68**, 102132, doi:10.1016/j.cbpa.2022.102132 (2022).

29 Young, G. & Kukura, P. Interferometric Scattering Microscopy. *Annu Rev Phys Chem* **70**, 301-322, doi:10.1146/annurev-physchem-050317-021247 (2019).

30 Agrawal, P., Lin, C., Mathad, R. I., Carver, M. & Yang, D. The major G-quadruplex formed in the human BCL-2 proximal promoter adopts a parallel structure with a 13-nt loop in K+ solution. *J Am Chem Soc* **136**, 1750-1753, doi:10.1021/ja4118945 (2014).

31 Siddiqui-Jain, A., Grand, C. L., Bearss, D. J. & Hurley, L. H. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc Natl Acad Sci U S A* **99**, 11593-11598, doi:10.1073/pnas.182256799 (2002).

32 Burd, C. G. & Dreyfuss, G. RNA binding specificity of hnRNP A1: significance of hnRNP A1 high-affinity binding sites in pre-mRNA splicing. *Embo j* **13**, 1197-1204, doi:10.1002/j.1460-2075.1994.tb06369.x (1994).

33 Luke, B. & Lingner, J. TERRA: telomeric repeat-containing RNA. *Embo j* **28**, 2503-2510, doi:10.1038/emboj.2009.166 (2009).

34 Spegg, V. *et al.* Phase separation properties of RPA combine high-affinity ssDNA binding with dynamic condensate functions at telomeres. *Nature Structural & Molecular Biology*, doi:10.1038/s41594-023-00932-w (2023).

35 Molliex, A. *et al.* Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **163**, 123-133, doi:10.1016/j.cell.2015.09.015 (2015).

36 Huelga, S. C. *et al.* Integrative genome-wide analysis reveals cooperative regulation of alternative splicing by hnRNP proteins. *Cell Rep* **1**, 167-178, doi:10.1016/j.celrep.2012.02.001 (2012).

37 Bruun, G. H. *et al.* Global identification of hnRNP A1 binding sites for SSO-based splicing modulation. *BMC Biol* **14**, 54, doi:10.1186/s12915-016-0279-9 (2016).

38 Feng, J., Zhou, J., Lin, Y. & Huang, W. hnRNP A1 in RNA metabolism regulation and as a potential therapeutic target. *Frontiers in Pharmacology* **13**, doi:10.3389/fphar.2022.986409 (2022).

39 Henricksen, L. A., Umbricht, C. B. & Wold, M. S. Recombinant replication protein A: expression, complex formation, and functional characterization. *J Biol Chem* **269**, 11121-11132 (1994).

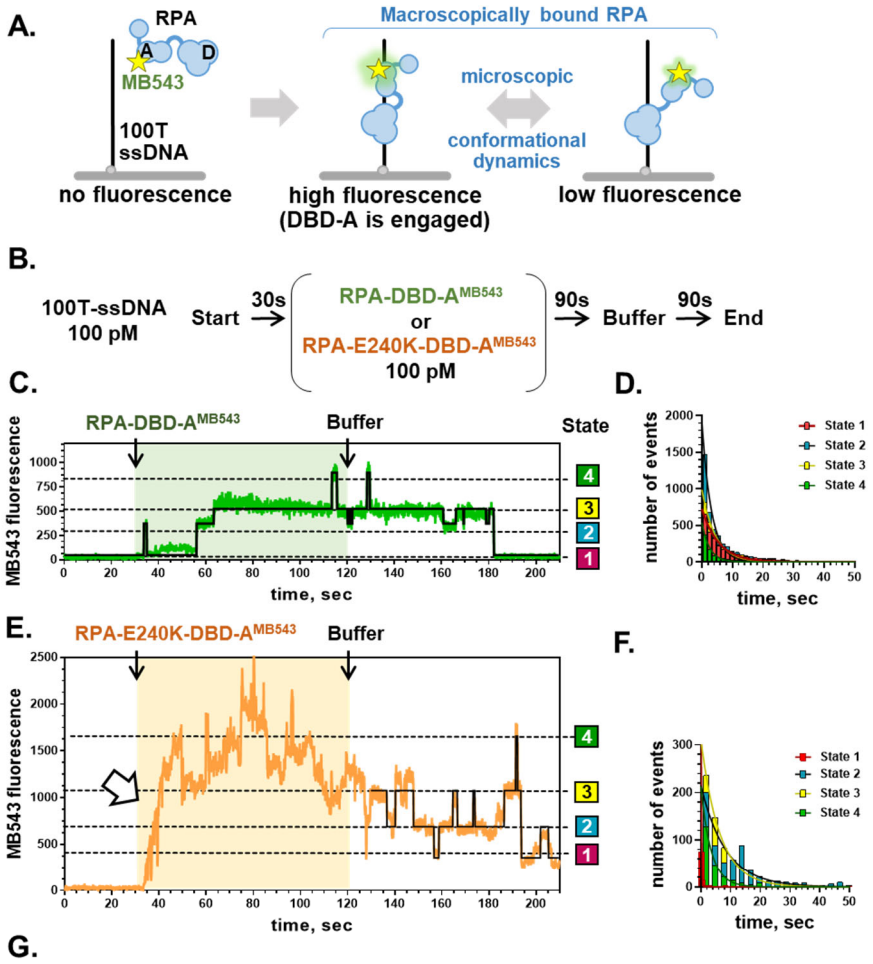
40 Fairlamb, M. S., Whitaker, A. M., Bain, F. E., Spies, M. & Freudenthal, B. D. Construction of a Three-Color Prism-Based TIRF Microscope to Study the Interactions and Dynamics of Macromolecules. *Biology (Basel)* **10**, doi:10.3390/biology10070571 (2021).

41 Bain, F. E., Fischer, L. A., Chen, R. & Wold, M. S. Single-Molecule Analysis of Replication Protein A-DNA Interactions. *Methods in enzymology* **600**, 439-461, doi:10.1016/bs.mie.2017.11.016 (2018).

42 Hon, J. & Gonzalez, R. L., Jr. Bayesian-Estimated Hierarchical HMMs Enable Robust Analysis of Single-Molecule Kinetic Heterogeneity. *Biophys J* **116**, 1790-1802, doi:10.1016/j.bpj.2019.02.031 (2019).

43 Tibbs, J. *et al.* KERA: analysis tool for multi-process, multi-state single-molecule data. *Nucleic Acids Res* **49**, e53, doi:10.1093/nar/gkab087 (2021).

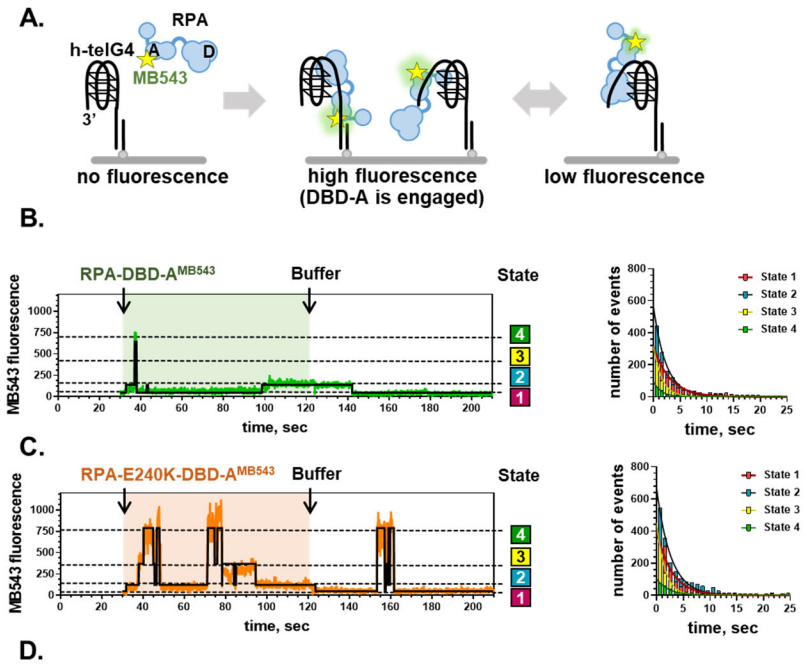
Figures and Figure Legends



protein \ state	1	2	3	4
RPA-DBD-A^{MB543}	5.5 ± 0.2 s 25.1%	3.1 ± 0.1 s 39.1%	4.4 ± 0.1 s 27.4%	2.5 ± 0.1 s 8.4%
RPA-DBD-D^{MB543}	5.6 ± 0.3 s 12%	3.9 ± 0.3 s 39.3%	2.8 ± 0.2 s 37.6%	6.4 ± 0.5 s 10.9%
RPA-E240K-DBD-A^{MB543}	0.12 ± 0.01 s 10.8%	8.1 ± 0.8 s 38%	6.4 ± 0.1 s 36.6%	3.4 ± 0.2 s 14.5%
RPA-E240K-DBD-D^{MB543}	3.0 ± 0.2 s 16.1%	2.4 ± 0.1 s 29.4%	2.9 ± 0.1 s 36.5%	8.4 ± 0.4 s 18%

Figure 1. DBD-A of RPA^{E240K} displays altered conformational dynamics. A. Single-molecule TIRFM assay for monitoring the conformational dynamics of RPA. Biotinylated ssDNA molecules (dT100) are tethered to the surface of the smTIRFM flow cells. Binding of the MB543-labeled RPA manifests in the appearance of the fluorescence signal, while conformational dynamics of the DNA-bound RPA manifests in changes in the fluorescence intensity. **B**. Experimental scheme. **C**. A representative fluorescence trajectory (time-based changes in the MB543 fluorescence in a specific location in the smTIRFM flow cell) for the wild type RPA labeled with MB543 at the DBD-

A (RPA-DBD-A^{MB543}; green) overlaid with an idealized trajectory (black) obtained by globally fitting all trajectories to a four-state model using hFRET. **D**. Dwell time distributions for the individual states of RPA-DBD-A^{MB543}. The dwell times were binned with a bin size of 2.4 seconds (bars). Solid lines represent exponential fits for each distribution. **E**. A representative fluorescence trajectory for the RPA^{E240K} mutant labeled with MB543 at the DBD-A (RPA^{E240K}-DBD-A^{MB543}; orange) overlaid with an idealized trajectory (black). **F**. Dwell time distributions for the individual states of RPA^{E240K}-DBD-A^{MB543}. The dwell times for states 2, 3 and 4 were binned with a bin size of 2.4 seconds (blue, yellow and green bars, respectively), while dwell times for state 1 were binned with a bin size 0.2 second (red bars). Solid lines represent exponential fits for each distribution. **G**. Summary of the dwell time analysis for all proteins in this study. The dwell times are shown as time constants from the exponential fitting of the dwell time distributions ± fitting error. Percentage of visitations to each state is shown below the respective dwell times.



D.

protein \ state	1	2	3	4
RPA-DBD-A^{MB543}	3.1 ± 0.1 s 55.7%	2.3 ± 0.1 s 45.1%	1.6 ± 0.1 s 15.4%	1.5 ± 0.1 s 3.9%
RPA-E240K-DBD-A^{MB543}	2.8 ± 0.1 s 30.2%	2.6 ± 0.1 s 48.6%	1.5 ± 0.1 s 16.2%	2.3 ± 0.1 s 4.9%

Figure 2. Binding and conformational dynamics of RPA-DBD-A^{MB543} and RPA^{E240K}-DBD-A^{MB543} on telomeric G-quadruplex DNA. A. Experimental setup. Biotinylated partial duplex DNA with 3'-ssDNA overhang containing five TTAGGG repeats was prepared under conditions enforcing formation of the G-quadruplex, and tethered to the TIRFM flow cell. At 30 seconds, the indicated fluorescently-labeled RPA was flown in, and at 120 seconds the unbound protein was removed by flowing in buffer. **B&C**. Representative MB543 trajectories and dwell time distributions for RPA-DBD-A^{MB543} (**B**) and RPA^{E240K}-DBD-A^{MB543} (**C**), respectively. All dwell times were binned with a bin size of 2.4 seconds (bars). Solid lines represent exponential fits for each distribution. **D**. Dwell time analysis for all states.

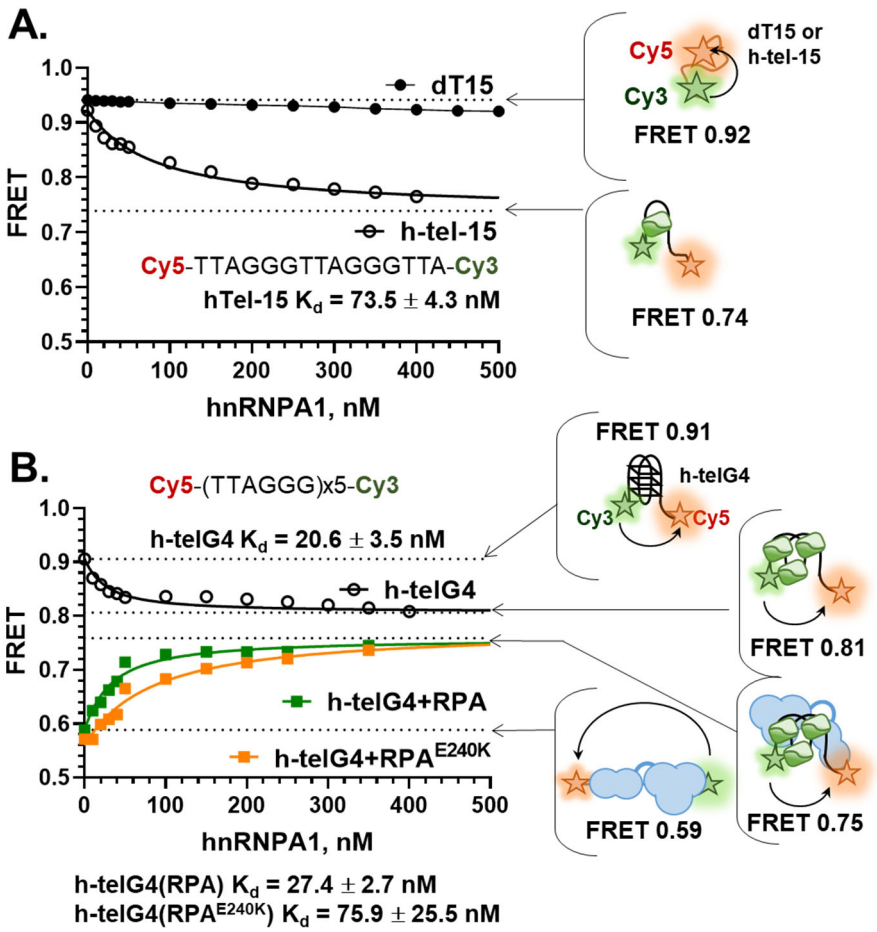


Figure 3. RPA and hnRNPA1 reorganize telomeric G-quadruplex DNA and form ternary complex. A. FRET-based analysis of the hnRNPA1 binding to 10 nM 15-mer poly(dT) (filled circles) and telomeric DNA (open circles) labeled with Cy3 (FRET donor) and Cy5 (FRET acceptor) at the two termini. The free and bound DNA substrates are schematically shown on the right with their respective FRET values. The experiments were carried out in triplicates and the values are plotted as average ± standard deviation for the three independent titrations. Note that the error bars for most measurements were smaller than the symbols' sizes. The binding curve to hnRNPA1 and telomeric DNA was fitted to a quadratic binding equation. The calculated K_d is shown with its fitting error. **B.** hnRNPA1 binding to 10 nM telomeric G-quadruplex folded in K^+ containing buffer (open black circles), and to telomeric G-quadruplex melted by RPA (green squares) or RPA^{E240K} (orange squares). The binding curves were fitted using quadratic binding equation, and the K_ds are shown with their respective fitting errors. Each complex with their respective FRET values are shown schematically on the right.

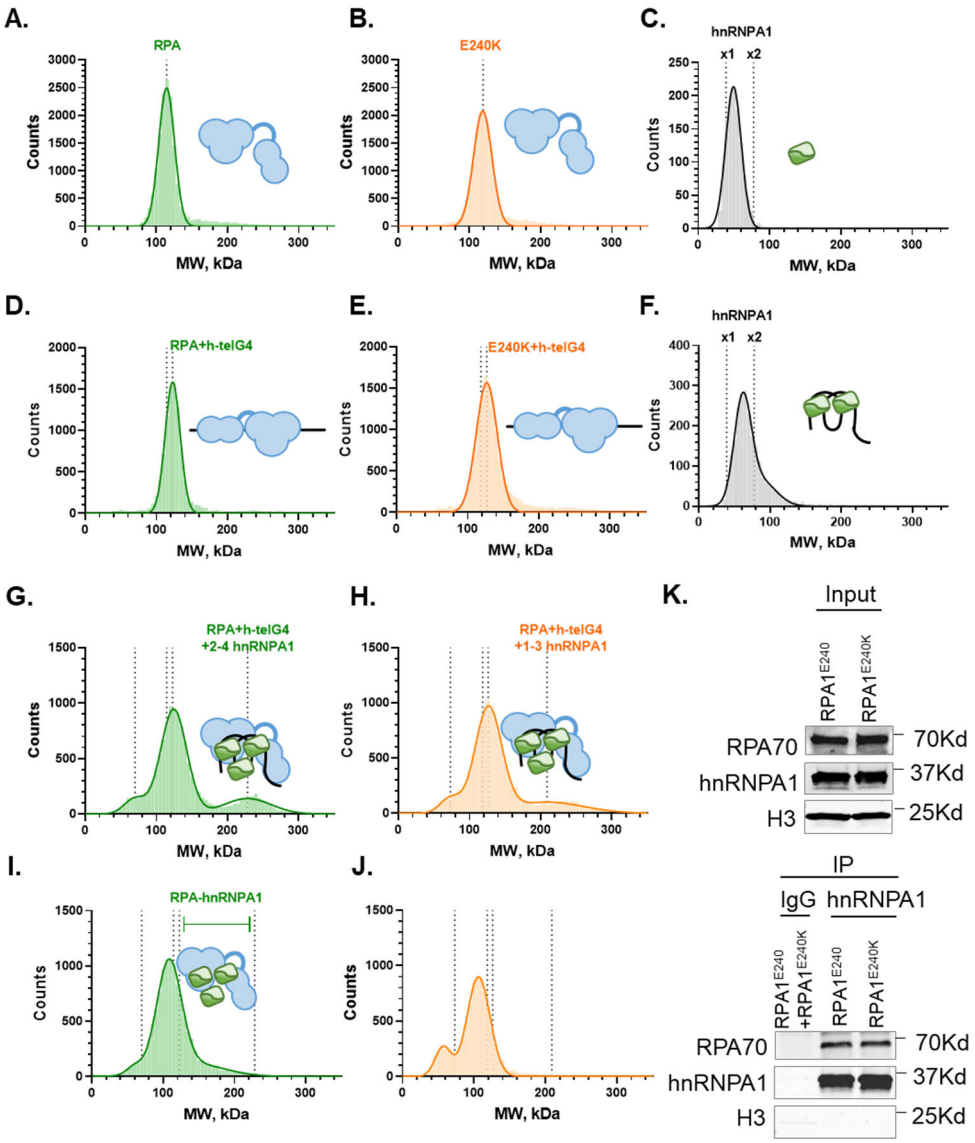


Figure 4. RPA and hnRNPA1 form ternary complex on telomeric ssDNA. A-J. The MP analyses of the 50 nM RPA (green), 50 nM RPA^{E240K} (orange) and 300 nM hnRNPA1 (grey), and their complexes with 50 nM telomeric ssDNA (five TTAGGG repeats) and each other. Recorded mass values were binned in 3 kDa bind, plotted and fitted with one, two or three Gaussians. **A-C**. Individual proteins. **D-F.** Complexes of individual proteins and telomeric DNA. **G&H**. Molecular weights of RPA (**G**) or RPA^{E240K} (**H**) mixed with telomeric DNA and hnRNPA1. The new Gaussian peak which we attribute to the ternary RPA-DNA-hnRNPA1 complex is marked with a cartoon representation of the complex. **I&J**. Molecular weights of RPA (**I**) or RPA^{E240K} (**J**) mixed with hnRNPA1. **K.** hnRNPA1 physically associates with RPA in cell lysates. Cell lysates of PCS201 iPSCs with wild type RPA1 and RPA1 p.E240K were treated with micrococcal nuclease prior to immunoprecipitation with anti-hnRNPA1. $N = 3$ biologically independent experiments.

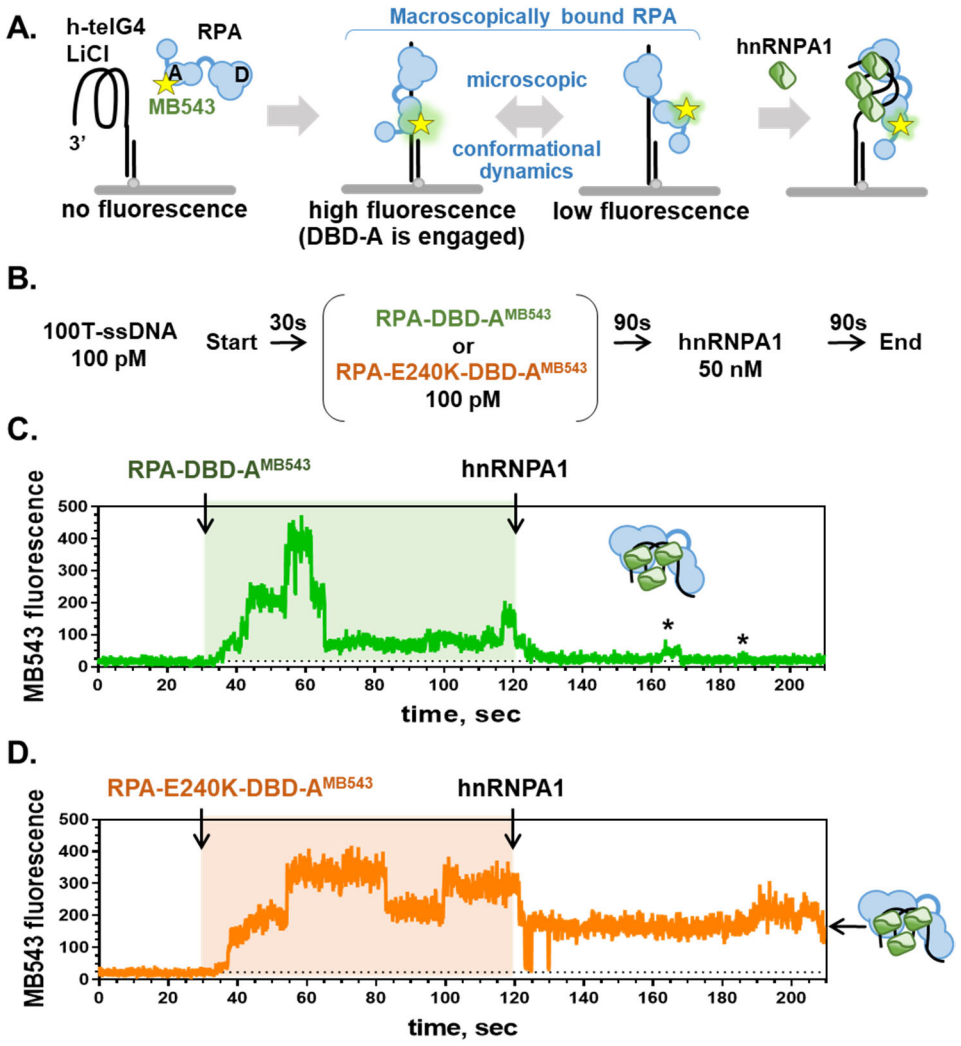


Figure 5. hnRNPA1 constrains conformational dynamics of the DBD-A of RPA bound to telomeric ssDNA. A. Experimental setup of the smTIRFM experiment. **B.** Experimental scheme. Biotinylated partial duplex DNA with 3'-ssDNA overhang containing five TTAGGG repeats was prepared under conditions refractive to G-quadruplex formation (Li^+), and tethered to the TIRFM flow cell. At 30 seconds, the indicated fluorescently-labeled RPA was flown in, and at 120 seconds the unbound protein was replaced with 50 nM unlabeled hnRNPA1. **C**. Representative MB543 trajectory for the wild type RPA-DBD-A^{MB543}. Asterisks mark deviation of the fluorescence signal from the lowest state suggestive of the presence of RPA. **D.** Representative MB543 trajectory for the RPA^{E240K}-DBD-A^{MB543}. Signal level corresponding to the RPA^{E240K}-hnRNPA1-DNA ternary complex is marked by an arrow and a carton representation of the complex.

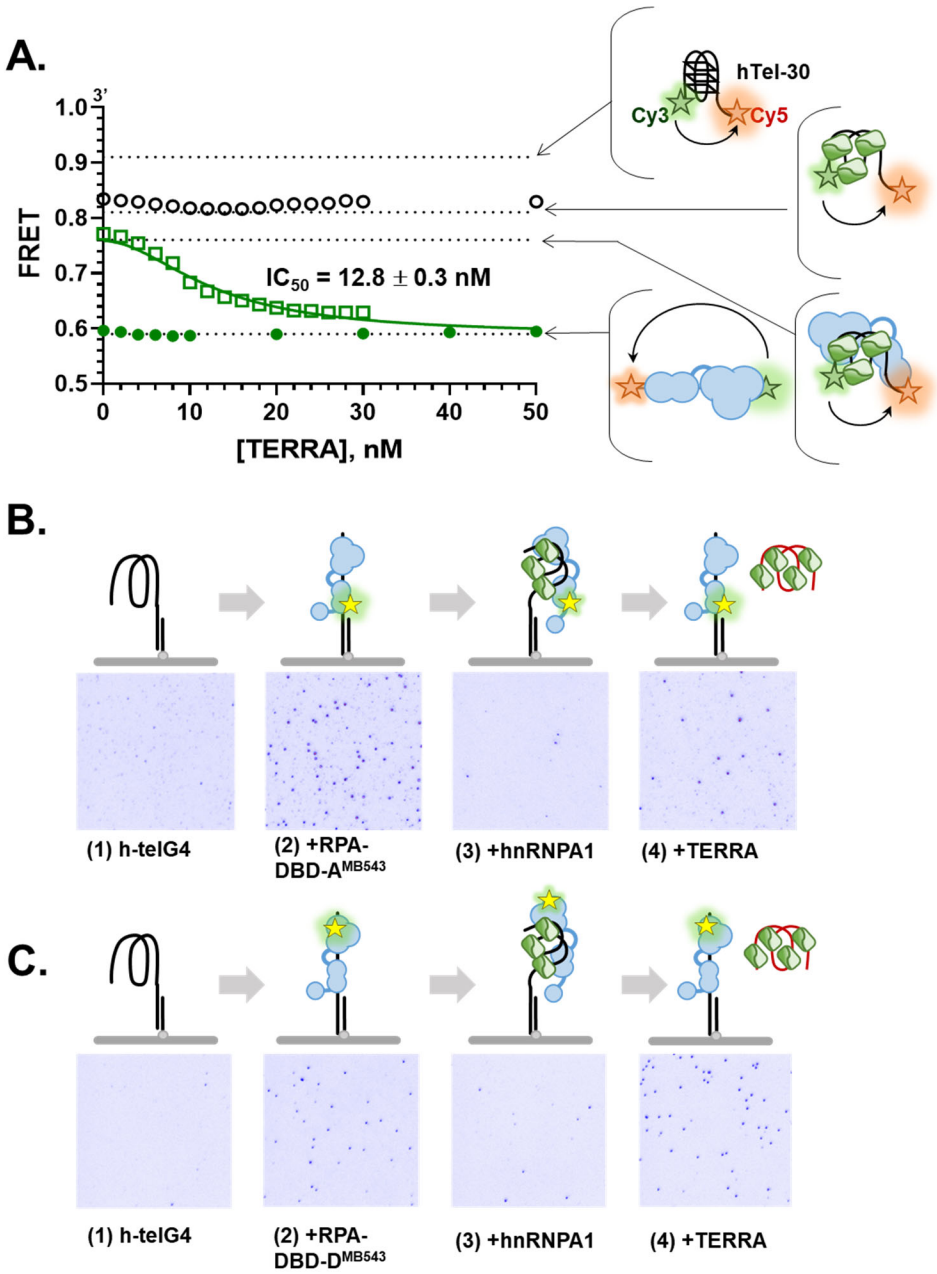


Figure 6. RPA-hnRNPA1 complex on telomeric ssDNA is remodeled by TERRA RNA. A. Bulk FRET experiment. RPA-DNA (filled green circles), hnRNPA1-DNA (open black circles) and RPA-DNA-hnRNPA1 (open green squares) were preassembled on 10 nM of telomeric G-quadruplex folded in the presence of K^+ and challenged by addition of TERRA RNA at indicated concentrations. The experiments were carried out in triplicates and the values are plotted as average ± standard deviation for the three independent titrations. Note that the error bars for most measurements were smaller than the symbols' sizes. The curve for hnRNPA1 displacement from the ternary complex was fitted to an inhibition dose response equation. The apparent IC_{50} value is shown with its fitting error. **B.&C.** Representative frames depicting portions of fields of view in the smTIRFM experiment after addition of 100 pM RPA-DBD-A^{MB543} (**B**) or RPA-DBD-D^{MB543} (**C**), 50 nM hnRNPA1, and 50 nM TERRA. The images we color-inverted in ImageJ and cropped to 1/4 of the field of view. See Supplemental Figure S7 for quantification and statistical analysis.

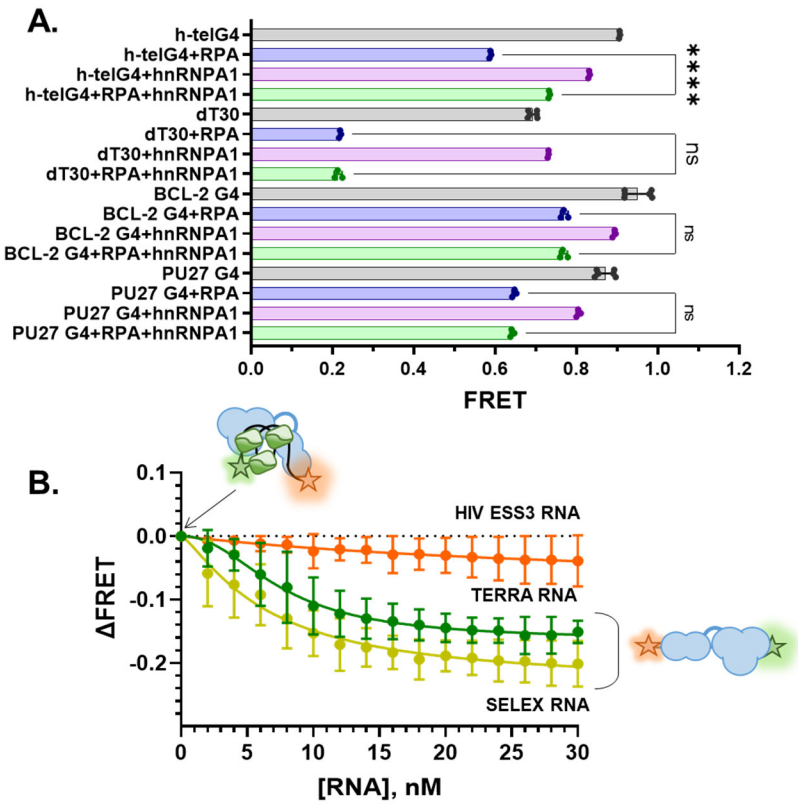


Figure 7. hnRNPA1-mediated RPA remodeling is specific to telomeric G-quadruplex and TERRA RNA. A. Bulk FRET experiment of 10 nM RPA (blue bars), 200 nM hnRNPA1 (purple bars) and 10 nM RPA + 200 nM hnRNPA1 (green bars) binding to 10 nM of telomeric G-quadruplex folded in the presence of K^+ (hTelG4), dT30 ssDNA (in Na^+), BCL-2, and PU27 G-quadruplexes folded in the presence of K^+. Statistical analysis: ordinary one-way ANOVA; **** $P<0.001$, n.s. not significant. **B.** RPA-DNA-hnRNPA1 complex was preassembled on 10 nM of telomeric G-quadruplex folded in the presence of K^+ and challenged by addition of TERRA RNA (green circles), SELEX RNA (gold circles) or HIV ESS3 RNA (orange circles) at indicated concentrations (see Supplemental Table 1 for sequences of DNA and RNA oligos). The data are plotted as change in FRET relative to the RPA-DNA-hnRNPA1 complex. The experiments were carried out at least in triplicates and the values are plotted as average ± standard deviation for the three or more independent titrations.

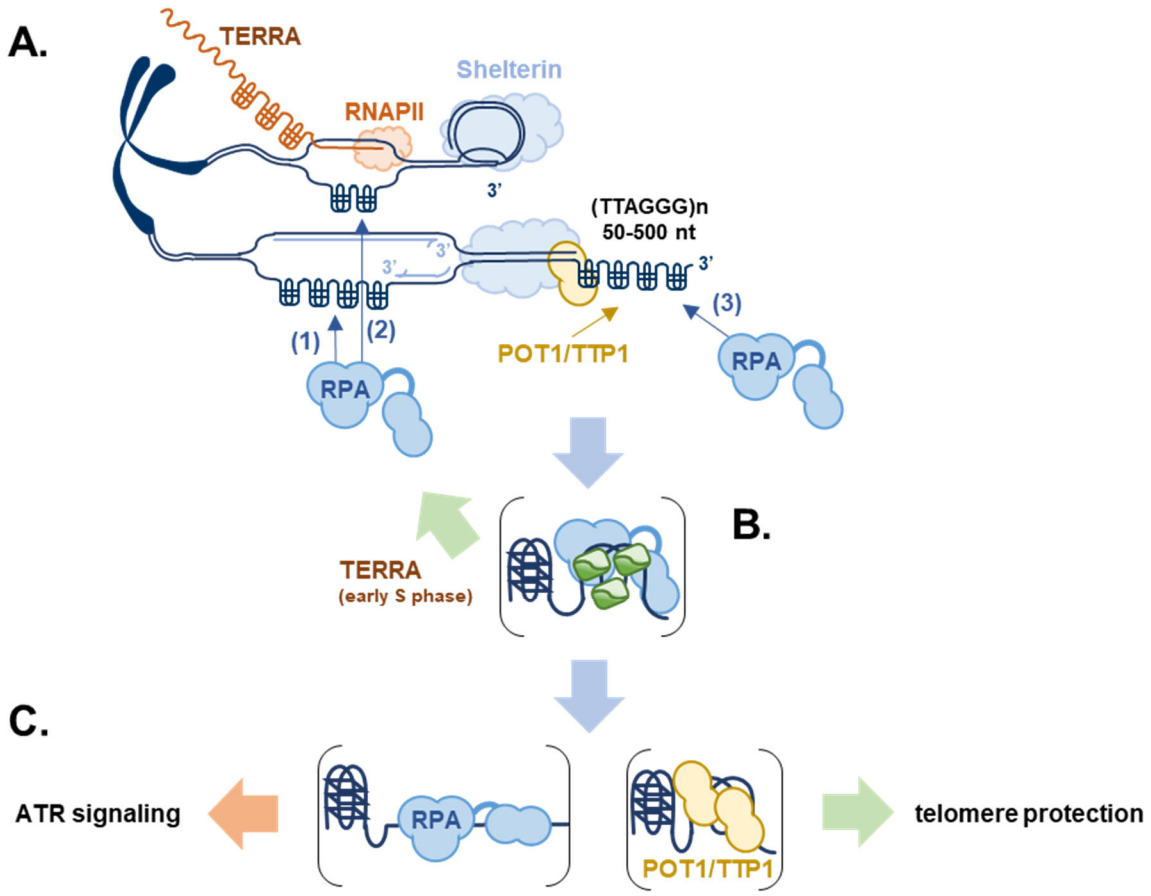


Figure 8. Telomere-specific RPA-DNA-hnRNPA1 complex. A. Several distinct events at human telomeres may lead to formation of G-quadruplexes that require RPA presence: (1) replication of the telomeric DNA and (2) transcription of TERRA RNA expose the G-strand, while (3) the end of human telomeres is a 50-500 nt 3' ssDNA overhang. **B**. Formation of the RPA-DNA-hnRNPA1 complex is specific to telomeric G-rich DNA and may play a protective role until RPA is replaced at the ssDNA overhangs with telomere-specific POT1/TTP1. Stable RPA-hnRNAPA1 complex with constrained conformational dynamics may present a nucleoprotein structure refractory to DNA damage signaling (**C**).

Methods

Chemicals and reagents

All chemicals were reagent grade (Sigma-Aldrich, St. Louis, MO). All fluorophores used to generate fluorescently labeled proteins were purchased from Click Chemistry Tools. Cy3-labeled, Cy-5 labeled, biotinylated and unmodified oligonucleotides were purchased from Integrated DNA Technologies. Sequences of all DNA oligonucleotides are listed in Supplemental Table 1. Twelve mM Trolox (6-hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid, Sigma-Aldrich; 238813-

1G) solution was prepared as described previously[3] by adding 60 mg of Trolox powder (238813-5G, Sigma-Aldrich) to 10 mL of water with 60 µL of 2 M NaOH, mixing for 3 days, filtering, and storing at 4 °C. Oxygen scavenging system, Gloxy was prepared as a mixture of 4 mg/mL catalase (C40-500MG, Sigma-Aldrich) and 100 mg/mL glucose oxidase (G2133-50KU, Sigma-Aldrich) in K50 buffer.

Protein Expression, labeling and generation of fluorescently labeled RPA variants

Both wild type RPA and mutant proteins were expressed and purified as previously described[9,39]. A pET28a(+) plasmid containing an open reading frame for human hnRNPA1 with an N-terminal 6xHis tag was synthesized by GeneScript. The 6xHis-hnRNPA1 was expressed in *E. coli* Rosetta cells using IPTG (0.5 mM) induction for 3 h at 37 °C. The cells were then harvested and lysed by sonication in lysis buffer (400 mM NaCl, 20mM imidazole, 1.72M sucrose, 50mM KPi buffer pH 7.4, 5mM BME, 0.5% Triton X-100, and 0.05% w/v lysozyme). The protein was purified using metal-affinity chromatography (HisTrap HP). Fractions containing hnRNPA1 were loaded onto a Heparin Sepharose column (HP) and eluted with a 30 mL gradient of 100 mM to 1 M NaCl in Heparin Buffer C containing 25 mM HEPES (pH 7.8), 0.1 mM EDTA, 1 mM DTT and 5% Glycerol. Peak fractions were then dialyzed against Heparin Buffer C containing 100 mM NaCl. RPA (wild type and mutant) and hnRNPA1 protein concentrations were determined by measuring absorbance at 280 nm using extinction coefficients of 88,830 M^{-1} cm^{-1} and 23,380 $^{-1}$ cm^{-1}, respectively.

4AZP-incorporated RPA proteins were purified and labeled with MB543 as previously described for yeast counterparts[3,22]. For human RPA, 4AZP was positioned at either Ser-217 (DBD-A; RPA1) or Trp-107 (DBD-D; RPA2), respectively. Briefly, ~3 ml of RPA4-AZP (10 µM) was incubated on a rocker with a 1.5-fold molar excess (15 µM) of dibenzocyclooctyne-amine fluorophore (DBCO-MB543, Click Chemistry Tools Inc.) for 2 h at 4 °C. Labeled RPA variants were separated from excess dye using a Biogel-P4 gel filtration column (Bio-Rad Laboratories; 65 ml bed volume) using storage buffer (30 mM HEPES, pH 7.8, 100 mM KCl and 10% (v/v) glycerol). Fractions containing labeled RPA were pooled, concentrated using a 30-kDa cut-off spin concentrator, and flash-frozen using liquid nitrogen. Fluorescent RPA was stored at −80 °C. Labeling efficiency was calculated using the respective extinction coefficients (ε) $\varepsilon 280 = 87{,}410\ M^{-1}\ cm^{-1}$ for RPA and $\varepsilon 550 = 105{,}000\ M^{-1}\ cm^{-1}$ for DBCO-MB543. We obtained 45 ± 17% and 40 ± 25% labeling efficiencies for the RPA–DBD-A^{MB543} and RPA–DBD-D^{MB543}, respectively.

Single-Molecule TIRFM

All single-molecule TIRFM studies were performed using a prism-based microscope custom-built around an Olympus IX71 microscope frame[3,40]. Videos were recorded using an electron-multiplying charge-coupled device camera (Andor; DU-897-E-CSO-#BV) at 100-ms time resolution. Background was set to 400, correction was set to 1,200 and gain was set to 250 for all videos recorded. Quartz slides (25 mm × 75 mm × 1 mm #1x3x1MM, G. Frinkenbeiner, Inc.) and cover glass (24 mm × 60 mm-1.5, Fisherbrand) were washed, coated and flow cells were assembled as described previously[3,41]. Assembled flow cells were mounted onto the microscope stage and rinsed with T50 buffer (10 mM Tris-HCl, pH 7.5, and 50 mM NaCl), K50 buffer (10 mM Tris-HCl, pH 7.5, and 100 mM KCl), or L50 buffer (10 mM Tris-HCl, pH 7.5, and 100 mM LiCl), then incubated with 0.2 mg ml–1 NeutrAvidin (Thermo Fisher) for 3 min and rinsed with buffer again.

DNA tethering and imaging was carried out in the following buffers: Poly-dT ssDNA binding experiments were carried out in reaction buffer containing 50 mM Tris-HCl, pH 7.5, 5 mM MgCl, 100 mM NaCl, 1 mM DTT, 0.1 mg ml^{-1} BSA, 0.8% w/v D-glucose, 12 μM glucose oxidase, 0.04 mg/ml catalase in Trolox solution. Telomeric G-quadruplex binding experiments were carried out in reaction buffer containing 50 mM Tris-HCl, pH 7.5, 5 mM MgCl, 100 mM KCl or LiCl, 1 mM DTT, 0.1 mg ml–1 BSA and 0.8% w/v D-glucose, 12 μM glucose oxidase, 0.04 mg/ml catalase in Trolox solution. To tether DNA to the slide surface, the flow cell was incubated for 3 min with 100 pM of biotinylated d100T or 1pM of biotinylated G-quadruplex (hTelG4-comp annealed to biotin base and pre-folded in either KCl or LiCl-containing buffer; see **Supplemental Table 1**) in the respective reaction buffer, then rinsed with reaction buffer to remove untethered DNA. Movies were recorded for 210 seconds at a frame rate of 100 ms. Wild type RPA or RPA^{E240K} labeled with MB543 in DBD-A or DBD-D was added at an indicated concentration after the first 30 seconds. At 120 seconds, free RPA was removed with reaction buffer or replaced with 50 nM hnRNPA1.

The effect of TERRA RNA was monitored by collecting 100 frame movies at a frame rate of 100ms after each substrate addition (100 pM DNA, 100 pM RPA, 50 nM hnRNPA1, and 50 nM TERRA RNA) within 5 different locations on the slide surface. Representative frames depicting portions of fields of view were color-inverted in ImageJ and cropped to 1/4 of the field of view. The number of trajectories with a fluorescence signal above background were quantified for 5 movies collected under the same condition. The fluorescence intensity of each point in these trajectories was calculated and plotted. Statistical analysis was performed using ordinary one-way ANOVA within GraphPad Prism.

Single-Molecule analysis of RPA conformational dynamics

An IDL script was used to extract fluorescence intensity trajectories from each video as previously described[3]. Trajectories were viewed and selected for analysis using a Matlab script. Only those trajectories that showed the appearance of the fluorescence signal between 30 s and 120 s, and not during the first 30 s, and had a signal-to-noise ratio > 4 (raw signal) were selected for analysis. The selected trajectories were then saved individually and globally processed for analysis by hFRET[42]. Idealized trajectories were imported from hFRET to KERA MATLAB script[43], which was used to optimize state assignment to avoid overfitting and remove transient events (1 or 2 frames in duration). Dwell times for each state were binned, plotted as frequency distributions and fitted to exponential functions using GraphPad Prism.

Bulk FRET measurements

After the baseline of buffer only was recorded, 10 nM of dT15, h-tel-15, dT30, h-telG4, BCL-2G4 1245, or PU27G4 (**Supplementary Table 1)** was added to the cuvette followed by addition of indicated concentrations of the wild type or mutant RPA, hnRNPA1, and/or TERRA RNA, SELEX RNA, or HIV ESS3 RNA (**Supplementary Table 1**)**.** FRET efficiency was calculated as FRET = $\frac{4.2 \cdot ICy5}{4.2 \cdot ICy5 + 1.7 \cdot ICy3}$, where *ICy5* is the averaged acceptor intensity and *Icy3* is the averaged donor intensity after subtracting the background fluorescence. The calculated FRET efficiency was plotted against protein or RNA concentrations and analyzed in GraphPad Prism. Equilibrium binding curves were fitted using quadratic binding equation to determine the Kds.

Mass photometry of the RPA (100 nM), hnRNPA1 (600 nM), and their complexes with each other and telomeric DNA (100 nM) was performed using the Refeyn TwoMP mass photometry instrument (Refeyn Ltd. Oxford, UK) in buffer containing 20mM Tris (pH 7.4), 100mM KCl, and 1mM DTT. Cover slides and silicon buffer gaskets were washed twice with miliQ water, 100% isopropanol, and again with miliQ water, and dried under an air stream at room temperature. Dried silicon gaskets were attached to the glass slide by applying a mild pressure and mounted. Molecular weight calibrations were performed using two protein oligomer solutions, β-amylase (56, 112 and 224 kDa) and Thyroglobulin (670 kDa). In each experiment, the indicated proteins and DNA were mixed to form 4x solution at room temperature, and then diluted 4-fold into the buffer-filled gasket. Individual molecular weights collected from 3000 frames (59.9 seconds) were binned in 3kDa bins and plotted as frequency histograms. GraphPad Prism was used to fit the molecular weight distributions to multiple Gaussians.

Electrophoretic Mobility Shift Assays (EMSAs)

DNA substrates (h-telG4, BCL-2 G4, or Pu27G4) were allowed for fold into G-quadruplexes by heating 1μM DNA solution at 95°C for 5 minutes followed by slow equilibration of the heating block to room temperature. Folded G-quadruplexes (30 nM molecules) or dT_{30} ssDNA were incubated with indicated amounts of RPA, hnRPA1 and RPA+hnRPA1 at 4°C for 30 min in 20 μl of standard reaction buffer, containing 20 mM Tris-Acetate (pH 7.5), 100 mM KCl, 1 mM EDTA, 5mM $MgCl_2$, 10% glycerol, 0.1 mg/ml BSA and 5 mM DTT. Two μl of Loading dye solution (20 mM Tris-Acetate (pH 8.0), 1 mM EDTA, 10% glycerol and 0.25 % Orange-G (w/v)) was added for each reaction. The reaction products were separated by electrophoresis on the 4.8% (29:1) native polyacrylamide gel in 20 mM Tris-Acetate (pH 8.0), 1 mM EDTA and 50mM KCl buffer. The gels were visualized separately in Cy3 and Cy5 channels using the ChemiDoc MP imaging system (BIO-RAD).

Co-immunoprecipitation and immunoblot

Previously established induced pluripotent stem cells (iPSCs) cell lines with wild type RPA1 ($RPA1^{E240}$) and RPA1 p.E240K ($RPA1^{E240K}$)[9] were lysed in RIPA lysis buffer (Cat# 20-188) supplemented with protease and phosphatase inhibitor cocktail (Thermo Cat# A32961 on ice for 20 minutes followed by centrifugation (14,000 × g, 10 min, 4 °C). The supernatant was removed and stored on ice while the remaining cell pellets were digested with micrococcal nuclease (BioLabs Cat no. M0247S) per protocol and to extract chromatin bound nuclear complexes and centrifuged (14,000 × g, 10 min, 37 °C). Supernatants from RIPA and micrococcal nuclease digestion were combined, pre cleared for 20 minutes with agarose IgG beads (Sigma Cat# A0919) and incubated with 1ug of anti-hnRNPA1 antibody (Santa Cruz Biotech, sc-32301) overnight at 4 °C with constant rotation. Protein A/G agarose beads (Thermo, Cat no. 88802) were used for pulldown according to the manufacturer's protocol. Precipitates were washed four times with cold lysis buffer, then resuspended in lysis buffer with SDS sample loading buffer, boiled for 10 min, and immediately subjected to SDS-PAGE for immunoblotting. Images were obtained using ChemiDoc Imaging System (BioRad). Anti-flag (Sigma, F1804, 1:1000) antibody was used for IgG control. Membranes were immunoblotted with anti-RPA1 (ThermoFisher, MA5-36226, 1:1000), anti-hnRNPA1 antibody (Santa Cruz Biotech, sc-32301) and anti-H3 (Cell Signaling, Cat no. 9717, 1:1000) antibodies.