RPA-Mediated Unfolding of Systematically Varying G-Quadruplex Structures

Sujay Ray,[†] Mohammad H. Qureshi,[‡] Dominic W. Malcolm,[†] Jagat B. Budhathoki,[†] Uğur Çelik,[§] and Hamza Balci[†*]

[†]Physics Department and [‡]Biological Sciences Department, Kent State University, Kent, Ohio; and [§]Department of Genetics and Bioengineering, Fatih University, Istanbul, Turkey

ABSTRACT G-quadruplex (GQ) is a noncanonical nucleic acid structure that is formed by guanine rich sequences. Unless it is destabilized by proteins such as replication protein A (RPA), GQ could interfere with DNA metabolic functions, such as replication or repair. We studied RPA-mediated GQ unfolding using single-molecule FRET on two groups of GQ structures that have different loop lengths and different numbers of G-tetrad layers. We observed a linear increase in the steady-state stability of the GQ against RPA-mediated unfolding with increasing number of layers or decreasing loop length. The stability demonstrated by different GQ structures varied by at least three orders of magnitude. Those with shorter loops (less than three nucleotides long) or a greater number of layers (more than three layers) maintained a significant folded population even at physiological RPA concentration (≈ 1 μM), raising the possibility of physiological viability of such GQ structures. Finally, we measured the transition time between the start and end of the RPA-mediated GQ unfolding process to be 0.35 $\pm$ 0.10 s for all GQ constructs we studied, despite significant differences in their steady-state stabilities. We propose a two-step RPA-mediated GQ unfolding mechanism that is consistent with our observations.

INTRODUCTION

Nucleic acid sequences rich in guanosine are capable of adopting four-stranded noncanonical structures called G-quadruplex (GQ) (1–4). GQ structures consist of an arrangement in which each guanosine occupies a corner of a G-tetrad and the G-tetrad layers stack together to form the GQ structure. GQ structures are stabilized by several mechanisms, including Hoogsteen hydrogen bonding between the four guanosines of each G-tetrad, stacking of the G-tetrad layers, shielding of repulsions between guanosines (due to negative charges on the O6 molecules) by monovalent cations that intercalate in or between the tetrad layers, and hydration (5–8). Genome-wide computational analysis has identified several hundred thousand potential GQ-forming sites (PQS) in the human genome, and in vitro assays have demonstrated GQ formation by these PQSs (9–11). In particular, PQSs are overrepresented in or near the promoter regions of numerous genes and at the ends of chromosomes (telomeres) (11–15). The telomeric GQ and capping proteins associated with telomeres are considered to protect the chromosome ends (17), whereas the roles of the nontelomeric GQ sequences have not been well characterized. Over-representation of these nontelomeric sequences in or near promoter sites suggests that they might be involved in transcription-level gene-expression regulation. RNA sequences can also form GQ structures. In particular, it has been demonstrated that RNA GQs located in the 5′-UTR play a role in translational level gene-expression regulation (18–21). The unique structure of GQ has motivated studies in which GQ has been used as a specific drug target (22,23).

Despite the abundance of PQSs in the genome, particularly nontelomeric GQ, and various in vitro demonstrations of GQ formation, it has been more challenging to unambiguously prove the existence and relevance of GQ structures in vivo. Most recently, GQs were visualized in human cells at both telomeric and nontelomeric locations, and their formation was shown to be modulated during the cell cycle (24). In addition, recent genome-wide studies have provided significant evidence on the formation and function of nontelomeric GQ structures in vivo. These studies demonstrate that eliminating certain helicases that are known to have GQ unfolding activity, such as Pif1 or BLM, results in increased DNA breaks in regions of the genome containing PQSs and severe retardation of DNA replication (9,25,26). Hence, in the absence of proteins that unfold them, GQ structures form and act as blocks to the replication system. These results are important indicators of the significance of protein-GQ interactions and their effects on genomic stability. Generic sequence constraints and in vitro thermal stability measurements have traditionally been used to characterize stability of GQ structures. However, within the context of cellular environment, the main factors that destabilize nontelomeric GQ are the proteins interacting with them and competition with the C-rich complementary strand. In particular, stability against GQ-unfolding proteins does not necessarily correlate with thermal stability (27). Thermal stability measurements provide a measure of the free energy difference between the folded GQ and the unfolded conformation, which has to be overcome by the protein while it is unfolding the GQ. However, these measurements are not

Submitted October 21, 2012, and accepted for publication April 2, 2013.

*Correspondence: hbalci@kent.edu

Editor: Ashok Deniz.

0006-3495/13/05/2235/11 $2.00

http://dx.doi.org/10.1016/j.bpj.2013.04.004

necessarily sensitive to structural complexities that could be very important in the context of protein-GQ interactions. For example, the footprint of one of the DNA-binding domains of replication protein A (RPA) is about three nucleotides long. Therefore, having a loop size of two or four nucleotides would affect the efficiency of establishing an initial contact in a way that may not be adequately described by the thermal stability of the GQ. After an initial contact is established, the thermal stability of the GQ would then determine the stability of the GQ against protein-mediated unfolding. Therefore, it is essential to measure the stability of GQ structures against destabilizing proteins for better estimates of their physiological viability. The abundance of proteins with GQ unfolding activity presents a considerable challenge for in vivo viability of these structures. With an emerging interest in GQ structures as potential drug targets and their role in gene-expression regulation (22,23), it is essential to probe protein-GQ interactions in systematic studies of GQ constructs with different structural properties.

In this study, we probed protein-GQ interactions by studying the unfolding of systematically varied GQ structures by RPA. The sequences under study were selected with the aim of quantifying the influence of layer and loop structures on the stability of GQ against protein-mediated unfolding. Such variations in the structure are particularly important in the case of nontelomeric GQ, which could be formed by a broad class of different sequences. RPA was selected as the model protein because it is the most abundant single-strand DNA (ssDNA)-binding protein in eukaryotes (~1 μM concentration in vivo (28)) and plays important roles in DNA metabolism, including DNA replication and repair (29–31). In particular, RPA protects the ssDNA created during replication or repair from enzymatic attack and prevents Watson-Crick pairing with the complementary strand before the completion of the process. In addition, RPA is involved in resolving certain secondary DNA structures that are formed during replication or repair, either by directly unfolding them or by initially binding to such structures and recruiting other proteins for unfolding (32,33). Unless resolved, such structures, including GQ, might act as roadblocks against the normal progression of replication or repair. RPA has very high affinity for ssDNA and GQ, with a dissociation constant, k_D, on the order of 1 nM, and its affinity for dsDNA is approximately three orders of magnitude lower (32). Previous bulk studies have shown that RPA can unfold several different telomeric and nontelomeric GQs (32,34,35); however, no study has systematically compared this unfolding activity in various GQ structures. The number of G-tetrad layers, loop and tail length (overhang sequences at the end of GQs) and their sequence, and folding conformation are some of the potentially important parameters that could determine the stability of GQ against protein-mediated unfolding.

In this study, we concentrated on the effects of two variables on the stability of GQs against RPA-mediated unfolding: the number of tetrad layers and the length of the loops. Four different constructs that have the same loop length and sequence (TTT) but different numbers of G-tetrad layers—two, three, four, and five—were used to determine the effect of number of layers on GQ stability against RPA-mediated unfolding. These GQ structures are named L2–L5 and have sequences of the form $TT(G_nT_3)_3G_nTT$, where $2 \leq n \leq 5$. In addition, five GQ constructs with a fixed number of G-tetrad layers (three) but with loop lengths varying between one and five nucleotides (T to TTTTT) were used to determine the effect of loop length on GQ stability. These GQ structures are named O1–O5 and have sequences of the form $TT(G_3T_n)_3G_3TT$ where $1 \leq n \leq 5$. We employed single-molecule Förster resonance energy transfer (smFRET) in these studies (36). Our experiments were performed at room temperature, near physiological pH (pH 7.5) and ionic strength (150 mM KCl and 2 mM $MgCl_2$), whereas RPA was titrated to its physiological concentration. Our results show that GQ steady-state stability against RPA-mediated unfolding systematically increases with increasing number of tetrad layers or decreasing loop length. We also monitored real-time unfolding of GQ by RPA and measured the unfolding time. Interestingly, all GQ constructs, L2–L5 and O1–O5, were unfolded within very similar times by RPA once an initial binding was established. Finally, we propose a two step model for RPA-mediated GQ unfolding that is consistent with these results.

MATERIAL AND METHODS

DNA constructs

The following DNA constructs were purchased from Integrated DNA Technologies (Coralville, IA) and used for FRET studies:

Stem: 5′-biotin-GCCTCGCTGCCGTCGCCA-Cy5-3′
L2: 5′-Cy3-TT**GGTTTGGTTTGGTTTGG**TTTGGCGACGGCAGCGAGGC-3′,
L3: 5′-Cy3-TT**GGGTTTGGGTTTGGGTTTGGG**TTTGGCGACGGCAGCGAGGC-3′,
L4: 5′-Cy3-TT**GGGGTTTGGGGTTTGGGGTTTGGGG**TTTGGCGACGGCAGCGAGGC-3′,
L5: 5′-Cy3-TT**GGGGGTTTGGGGGTTTGGGGGTTTGGGGG**TTTGGCGACGGCAGCGAGGC-3′,
O1: 5′-Cy3-TT**GGGTGGGTGGGTGGG**TTTGGCGACGGCAGCGAGGC-3′,
O2: 5′-Cy3-TT**GGGTTGGGTTGGGTTGGG**TTTGGCGACGGCAGCGAGGC-3′,
O3: 5′-Cy3-TT**GGGTTTGGGTTTGGGTTTGGG**TTTGGCGACGGCAGCGAGGC-3′,
O4: 5′-Cy3-TT**GGGTTTTGGGTTTTGGGTTTTGGG**TTTGGCGACGGCAGCGAGGC-3′,
O5: 5′-Cy3-TT**GGGTTTTTGGGTTTTTGGGTTTTTGGG**TTTGGCGACGGCAGCGAGGC-3′,

The DNA sequences shown in bold letters are the GQ-forming segment. Note that L3 and O3 are the same constructs. The underlined thymines are used as spacers to minimize the interaction between the GQ structures and the nearby fluorophore. The stem strand was annealed to L2–L5 and O1–O5 and a partial duplex DNA with an 18-basepair (bp) duplex stem was formed (see Fig. 2). The two DNA oligonucleotides were mixed at 2.0 μM

concentration in 10 mM Tris buffer, pH 7.5, and 50 mM NaCl. The mixture was heated and maintained at 95°C for 5 min and cooled down gradually by placing a heated block at room temperature until the sample reached room temperature.

smFRET assay

Cleaned optical quartz slides and glass coverslips were used for preparing the imaging chambers. Surfaces were coated with a mixture of polyethylene-glycol and biotinylated polyethylene-glycol (m-PEG-5000 and biotin-PEG-5000, respectively; Laysan Bio, Arab, AL) to prevent nonspecific binding of DNA and RPA to the surface. Biotinylated DNA, at 15 pM concentration, was immobilized onto the biotin-PEG surface via neutravidin at 0.05 mg/ml concentration. For RPA titration studies, the GQ structures were first formed by incubating the surface-immobilized DNA constructs in 150 mM KCl for 15 min. RPA was then added to the chamber in an imaging solution (50 mM Tris, pH 7.5, 0.8 mg/ml glucose, 0.1 mg/ml bovine serum albumin, 140 mM β-mercaptoethanol, 0.1 mg/ml glucose oxidase, 0.02 mg/ml catalase, Trolox at saturating concentration, 2 mM $MgCl_2$, and 150 mM KCl). Images were acquired after 15 min of incubation.

Imaging and data analysis

A prism-type total internal reflection fluorescence microscope built around an Olympus IX-71 microscope was used for these measurements. Movies 1000–2000 frames long were collected using an Andor Ixon electron-multiplying CCD camera (iXon DV 887-BI, Andor Technology, South Windsor, CT). An integration time of 100 ms was used for RPA titration experiments, and an integration time of 35 ms was used for flow experiments. Control flow experiments were also performed at 18-ms acquisition time. FRET time traces for individual molecules were analyzed to generate FRET histograms using a custom analysis program, and Origin Pro 8 was used for statistical analysis and curve fitting of FRET histograms.

Circular dichroism assay

Circular dichroism (CD) measurements were performed at room temperature using a Jasco J-810 spectrophotometer and a cuvette with a 0.1-cm pathlength. The measurements were performed at 150 mM K^+ and 2 mM Mg^{2+}. In addition, control measurements were performed at 150 mM Li^+ and 2 mM Mg^{2+} to demonstrate effective GQ folding in the presence of K^+ but not Li^+. The DNA concentration was kept at 4 μM in all measurements.

RPA preparation

The RPA purification procedure was adapted from previous works (37,38). Briefly, *Escherichia coli* cells were transformed with a p11d-tRPA construct containing the coding sequences of RPA 70, RPA 14, and RPA 32. Upon reaching an OD_{600} of 0.6, protein expression was induced by adding isopropylthio-β-galactoside to a final concentration of 0.4 mM. Cells were lysed by pelleting and sonication. Cellular debris was pelleted by centrifugation at 12,000 rpm for 30 min. The supernatant was loaded onto an Affi-Gel Blue column (Bio-Rad, Hercules, CA). Protein was eluted using 1.5 M NaSCN in Hepes-Inositol buffer, pH 7.8. Eluted fractions containing protein were loaded onto a hydroxyapatite column to further concentrate the protein and eluted with Hepes-Inositol-80 mM phosphate buffer, pH 7.5. RPA purity was assayed by sodium dodecyl sulfate polyacrylamide gel electrophoresis. RPA functionality was confirmed in the context of its role in a DNA checkpoint complex using the assay described by Choi et al. (39).

RESULTS AND DISCUSSION

Circular dichroism measurements

We performed CD measurements to confirm GQ formation by L2–L5 and O1–O5. Fig. 1 shows the results of these measurements, which were performed under physiologically germane ionic-strength and pH conditions (150 mM K^+, 2 mM Mg^{2+}, pH 7.5). A peak at 260 nm and a trough at 240 nm in the ellipticity measured by CD is consistent with parallel GQ conformation, whereas a peak at 290 nm and a trough at 260 nm is consistent with antiparallel GQ conformation (40). On the other hand, an ellipticity that has a peak around 290 nm and a shoulder around 260 nm has been interpreted as the hybrid conformation or a mixture of GQ molecules possessing parallel or antiparallel conformations (41,42). Given these interpretations, L2, O4, and O5 have antiparallel GQ conformation, whereas O1 and O2 have parallel conformation. On the other hand L3–L5 data are consistent with the hybrid conformation (a mixture of parallel and antiparallel conformations). The data on

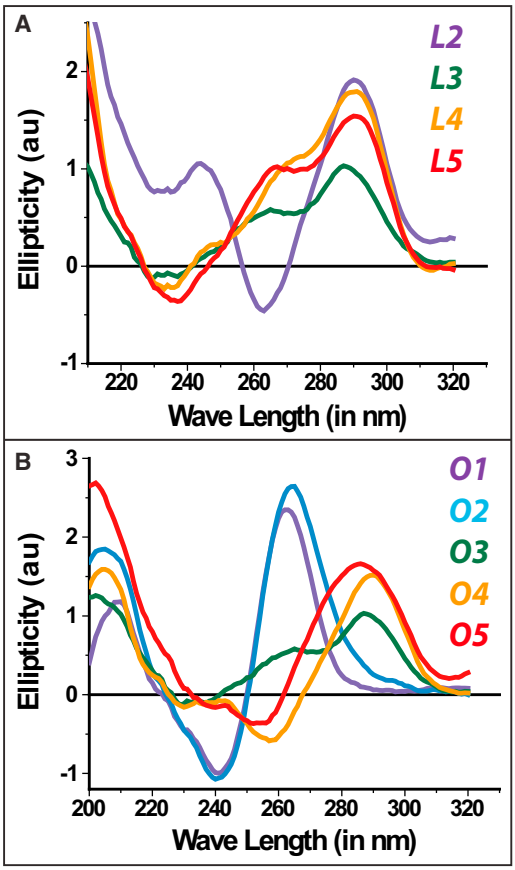

FIGURE 1 CD measurements on all the GQ constructs studied at 150 mM K^+ and 2 mM Mg^{2+}. (*A*) L2–L5 constructs. (*B*) O1–O5 constructs. Data on O1 and O2 are consistent with parallel GQs, whereas that on L2, O4, and O5 are consistent with antiparallel GQ conformations. Data on L3–L5 have signatures of both parallel and antiparallel GQ (the hybrid conformation).

O1–O5 show an interesting trend in which the short loop constructs (O1 and O2) are consistent with the parallel GQ conformation, whereas the longer loop constructs (O4 and O5) are consistent with the antiparallel conformation. The O3 construct, with intermediate loop length, is the intermediate state between these two conformations and has an ellipticity consistent with that of the hybrid conformation. Control CD measurements in the absence of salt or in the presence of 150 mM Li^+ were performed for all constructs (see Fig. S1 in the Supporting Material). These measurements show no signature of GQ formation for some constructs or a significantly weaker signal for others, compared to incubation at 150 mM K^+. Therefore, these results are consistent with K^+ being a more efficient stabilizer of the GQ structure than Li^+. The CD data are merely used here to confirm GQ formation. Studying all the folding conformations independently is beyond the resolution of the methods used in this work, both CD and smFRET. For example, NMR measurements have shown as many as five different folding conformations for a construct similar to L3 (43). In particular, smFRET is sensitive to distance between donor and acceptor fluorophores that are placed at the ends of the GQ-forming sequence so as to cause minimal disturbance to the structure. However, different folding conformations typically result from different arrangements of the loops without a significant change in the end-to-end distance of GQ. All the results presented for a given GQ construct should be considered as the average of all possible conformations of that GQ construct. For demonstration purposes, we illustrate one of the possible folding conformations for all the GQ constructs (Fig. S2, *B* and *C*).

Steady-sate smFRET measurements of RPA-mediated GQ unfolding

Prism-type total internal reflection fluorescence (TIRF) microscopy was used to perform smFRET measurements, as schematically shown in Fig. S2 *A*. The DNA constructs are in the form of a partial duplex with a double-stranded stem and a single-stranded overhang that folds into a GQ structure. The donor fluorophore (Cy3) is placed at the 5′ end of the single-stranded extension and the acceptor fluorophore (Cy5) is placed at the 3′ end of the double-stranded stem. When structured as GQ, the construct brings the donor-acceptor pair into close proximity, resulting in high FRET efficiency. On the other hand, unfolding of the GQ by RPA increases the distance between the donor-acceptor pair, resulting in low FRET efficiency. This low FRET efficiency peak is clearly distinguishable from the donor-only peak that is due to leakage through the dichroic (see Fig. S3 for an example). The duplex stem sequence is used to reduce the interaction of the GQ-forming sequence with the surface. As affinity of RPA to dsDNA is about three orders of magnitude smaller than its affinity to ssDNA or GQ structures, the interaction between RPA and the duplex stem is negligible (32,44,45). A control measurement was performed to demonstrate that RPA does not bind or modify this 18-bp duplex DNA (Fig. S4). A similar single-molecule assay has been used to study GQ formation of human telomeric repeats (27,46,47).

Proper folding of the oligonucleotides into G-quadruplex structure was established by monitoring the increase in FRET signal between the donor and acceptor fluorophores with increasing K^+ ion concentration. As the GQ structure is stabilized with increasing K^+ concentration, the donor and acceptor molecules move closer to each other, resulting in higher FRET. With a gradual increase in K^+ concentration, various secondary structures form, which are manifested as different FRET peaks. All of these peaks eventually converge to a specific high-FRET state as K^+ concentration is increased to 150 mM. We established formation of a stable structure by monitoring this progressive folding for all the constructs we studied. As shown in the example in Fig. S5, various structures are stabilized at different ionic strengths before the GQ becomes the dominant stable structure at 150 mM K^+.

In the steady-state smFRET experiments, the constructs were incubated at 150 mM K^+ and 2 mM Mg^{2+} for 15 min to attain a stable GQ conformation, except in the case of the O5 construct, which was incubated for 1 h. After proper folding of the GQ, different concentrations of RPA were introduced to the sample chamber while maintaining the same ionic strength and pH. Data were collected after 15 min of RPA incubation to attain a steady-state unfolding of GQ and binding of RPA to the unfolded DNA. Incubation times of RPA with DNA were varied between 5 and 30 min to determine the time necessary for attaining the steady state. Incubation times beyond 15 min did not make a difference in the fraction of GQ unfolded and bound by RPA.

Fig. 2 summarizes RPA-mediated GQ unfolding for L2–L5 constructs at different RPA concentrations. The high-FRET peaks ($E_{FRET} \approx 0.7$–0.8) in these data (Fig. 2 *A*) represent folded GQ structures, whereas the low-FRET peaks ($E_{FRET} \approx 0.1$–0.2) represent the RPA-bound unfolded conformations. The unfolded DNA that is bound by RPA demonstrates a significantly lower FRET efficiency compared to the unfolded DNA that is not bound by RPA ($E_{FRET} \approx 0.3$–0.6 depending on the length of the DNA construct) (48). The two states are distinguishable, as shown in the example in Fig. S6. Therefore, the low FRET peak we observe in Fig. 2 *A* represents the RPA-bound unfolded DNA. Larger fractions of GQ structures were unfolded and bound by RPA as the RPA concentration was increased, and eventually a saturating fraction was reached at a certain RPA concentration. We observed large variations among different GQ constructs in terms of their stability against RPA-mediated unfolding. As shown in Fig. 2 *A*, all of the L2 molecules were unfolded by <20 nM RPA, whereas a certain fraction of molecules in L3–L5 remain stably folded even at 1 μM RPA concentration. To quantify the unfolding

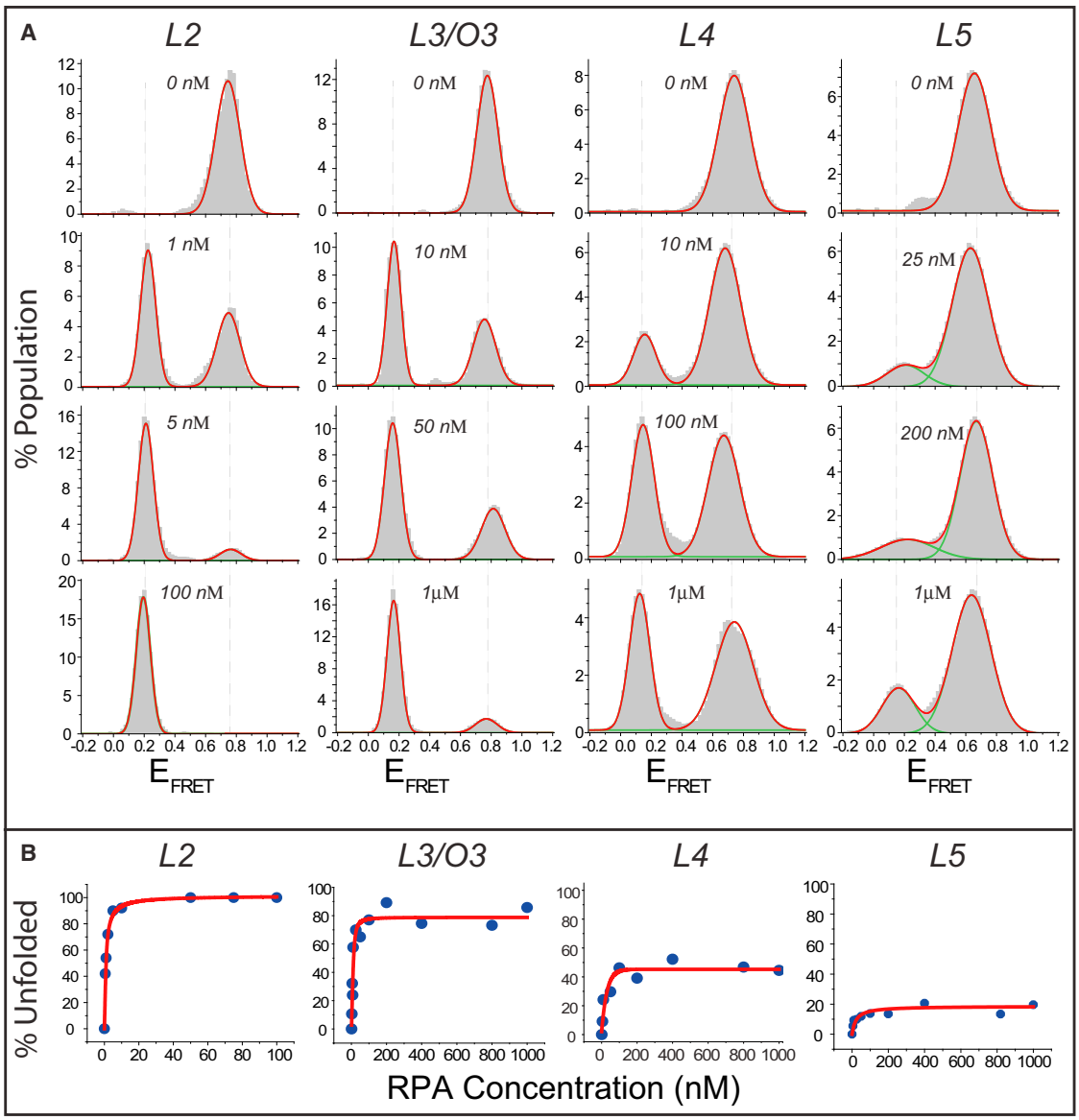

FIGURE 2 (*A*) smFRET data on unfolding of L2–L5 GQ constructs by varying concentrations of RPA. The concentration quoted on each histogram represents the concentration of RPA used for that measurement. High-FRET peaks represent the folded structure and low-FRET peaks the RPA-bound unfolded DNA. Each histogram represents RPA-mediated GQ unfolding at a particular RPA concentration. (*B*) Percentage of GQ molecules unfolded and bound by RPA as a function of RPA concentration for L2–L5. The red line shows a Langmuir binding isotherm fit to the data. The stability of GQ structures systematically increases as the number of G-tetrad layers is increased. L3–L5 molecules maintain a certain fraction of folded GQ molecules at RPA concentrations as high as 1 μM, whereas all L2 molecules are unfolded by 10 nM of RPA.

phenomena, we plotted the percentage of GQ molecules unfolded and bound by RPA as a function of RPA concentration (Fig. 2 *B*). These curves were analyzed using a Langmuir binding isotherm of the form $y = \alpha[\text{RPA}]/([\text{RPA}] + K_{eq})$, where y describes the percentage of GQs unfolded and bound by RPA, [RPA] is the RPA concentration, K_{eq} is the equilibrium constant, and α represents the percentage of unfolded and RPA-bound DNA at saturating RPA concentration. The α parameter accounts for the incomplete unfolding of L3, L4, and L5 at saturating RPA concentration. A summary of the fitting parameters for different constructs is given in Table 1. The steady-state stability of the GQ structures systematically increases as the number of G-tetrad layers is increased. In addition, the percentage of GQ molecules unfolded and bound by RPA at saturation (α parameter) systematically decreases as the number of G-tetrad layers is increased (see Fig. 4 *A*). The slope of the linear fit (see Fig. 4 *A*) suggests that each additional G-tetrad layer decreases the α parameter by 27%.

The O1–O5 constructs were analyzed using methods similar to those presented for the L2–L5 constructs, as shown in Fig. 3. All of the O4 and O5 molecules were

TABLE 1 Summary of Langmuir binding isotherm analysis for all GQ constructs studied

Fitting equation: $y = \alpha[RPA]/([RPA] + K_{eq})$		
Construct	α (%)	K_{eq} (nM)
L2	100	0.8 ± 0.1
L3 (O3)	78.1 ± 3.5	11.5 ± 2.2
L4	44.7 ± 2.6	23.1 ± 7.8
L5	18.6 ± 1.0	22.1 ± 7.5
O1	16.0 ± 0.9	81.2 ± 19.7
O2	56.9 ± 2.9	7.1 ± 2.3
O4	100	1.0 ± 0.1
O5	100	0.6 ± 0.1

unfolded by 20 nM RPA, whereas some O1–O3 molecules remained folded even at 1 μM RPA (Fig. 3 *A*). Langmuir binding isotherm analysis was performed on these data and the results (α and K_{eq}) are summarized in Table 1. The steady-state stability of the GQ structures systematically increases as the length of the loops is decreased and the α parameter systematically decreases as the length of the loops is decreased, as shown in Fig. 4 *B*. A linear fit to α parameters of O1–O4 constructs demonstrates a 26% increase with one nucleotide added to each loop (Fig. 4 *B*). O5 data were not included in the fitting, as 100% of O4 molecules were already unfolded and bound by RPA. Interestingly, for the GQ structures we studied, the effect on steady-state stability of reducing the length of each loop by one nucleotide is similar to that of adding a G-tetrad layer (26% vs. 27% change in α). Determining whether this is a coincidence due to the structures studied or a general feature of a broader class of GQ constructs would require systematic studies on a larger number of GQ constructs.

Several control measurements were performed to assure the validity of the analysis, and conclusions presented in Figs. 2 and 3. In one of these measurements, we confirmed that all secondary DNA structures were removed in the RPA-bound unfolded state (see Fig. S7). A GQ-forming DNA (the L4 construct) and a polythymine DNA of very similar length (30 vs. 29 nucleotides long, respectively) were used for this study. The polythymine DNA was selected because it does not form any secondary structure. We demonstrated that the FRET peak representing the unfolded and RPA-bound state of L4 is identical to that of the RPA-bound polythymine DNA, suggesting that all secondary structures have been removed from the L4 construct. In another measurement, of the γ parameter for the Cy3-Cy5 pair before and after introducing RPA, we confirmed that the lower FRET peak observed upon introduction of RPA is not due to interactions of the fluorophores with RPA or the GQ structure (see Fig. S8).

An interesting aspect of the steady-state smFRET measurements is that a certain fraction of L3–L5 and O1–O3 molecules remain stably folded even at the highest RPA concentration studied. For all these constructs, a steady state is reached at ~100 nM RPA, and adding more RPA to the environment does not give rise to any further significant unfolding of GQ. This suggests a dynamic binding/unbinding of RPA, although such was not observed in our 2- to 3-min-long single-molecule time traces. In particular, RPA binding to the unfolded DNA is very stable and essentially irreversible within our 2- to 3-min observation time. It is possible that RPA would dissociate and GQ would refold over longer time periods. To test this idea, movies 10–20 min long with 0.5-s integration time and very low laser power were acquired. Back-and-forth transitions between the folded, unfolded, and RPA-bound states were rarely observed even in these long movies. Examples of these rare transitions and stability of RPA binding are shown in Fig. S9.

A related issue is the stoichiometry of RPA binding. The length of our DNA constructs varies between 19 and 33 nucleotides, all of which lengths support stable binding of a single RPA, which can bind to 8- to 30-nucleotide-long ssDNA (as discussed in more detail below). The only possible exception to this is the L5 construct, which is 33 nucleotides long and therefore might possibly accommodate stable binding of one RPA and binding of one of the DNA-binding domains of a second RPA. In the smFRET histograms in Fig. 2 A, a single peak is observed for the RPA-bound and unfolded state of L5, suggesting that binding of a second RPA either does not take place or is a very rare event.

Dynamics of the RPA-mediated GQ unfolding process

Finally, single-molecule buffer-exchange measurements (called RPA flow) were performed to measure transition time from the folded GQ state to the RPA-bound unfolded state. In RPA flow experiments, the DNA constructs were incubated in a buffer that contains 150 mM K^+ (in the absence of RPA) for 15 min to ensure proper folding of GQ, with the exception of O5, which was incubated for 1 h. A buffer containing 100 nM RPA and 150 mM K^+ was flowed into the chamber by a microfluidic syringe pump while the folded GQ construct was being imaged, enabling us to monitor RPA-mediated GQ unfolding in real time. A representative single-molecule time trace of this unfolding process is shown in Fig. 5 *A*. The initial high-FRET state represents the folded GQ conformation, and the low-FRET state that follows corresponds to the RPA-bound unfolded state. The duration of time between these two states is called the unfolding time for brevity. The low-FRET state matches with the RPA-bound state obtained in the steady-state experiments, and it is different from the donor-only level (the FRET level after acceptor photobleaching takes place (Fig. 5 *A*)).

The unfolding time, shown as Δt in Fig. 5, *A–C*, is measured from the start of FRET decrease. The unfolding is considered to be complete when the FRET signal stabilizes at a low-FRET state representing the RPA-bound DNA structure. As shown in Fig. 5, *A–C*, several different

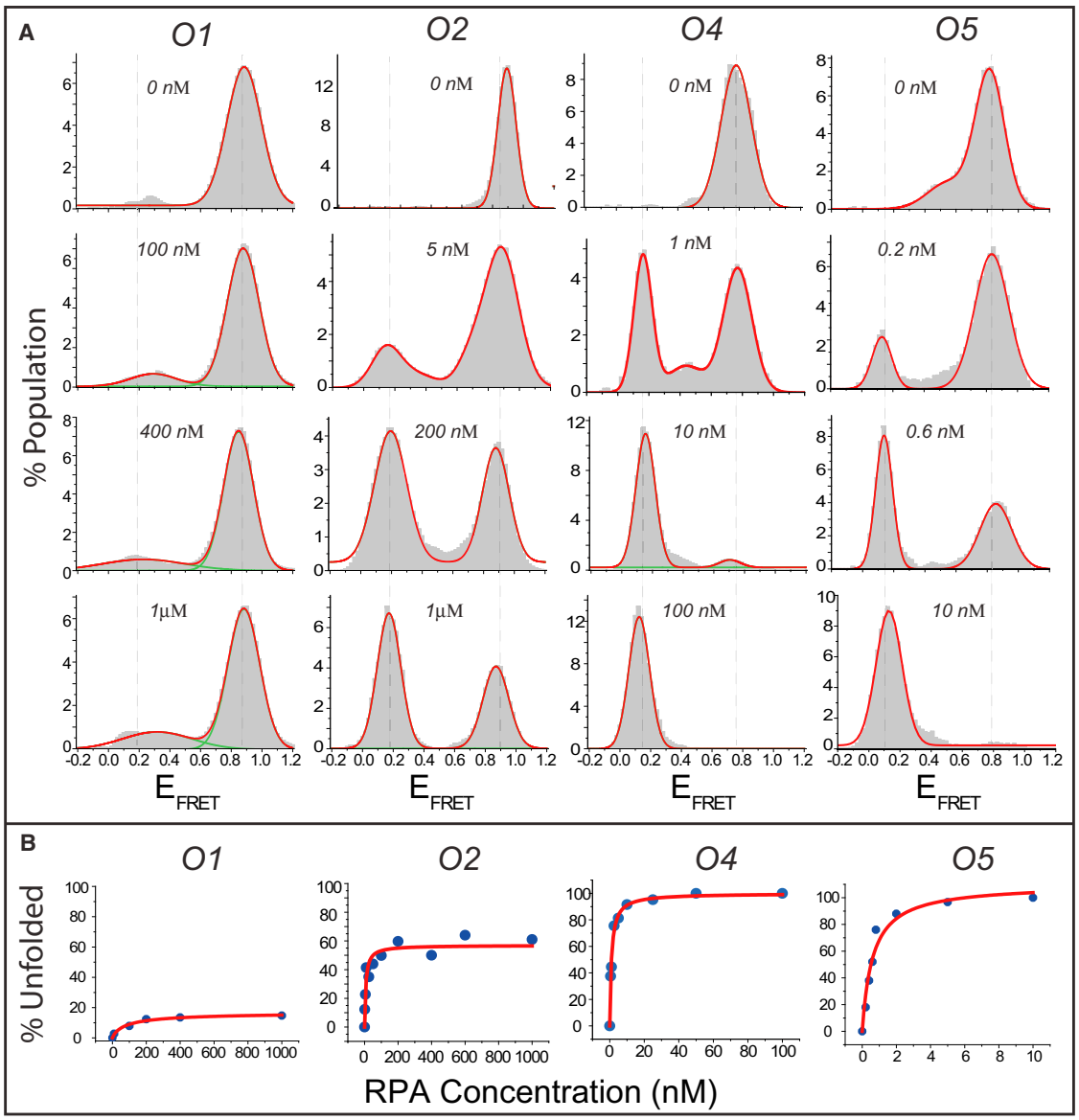

FIGURE 3 (*A*) smFRET data on unfolding of O1–O5 GQ constructs by varying concentrations of RPA. The concentration quoted on each histogram represents the concentration of RPA used for that measurement. (*B*) Percentage of GQ molecules unfolded and bound by RPA as a function of RPA concentration for O1–O5. The red line shows a Langmuir binding isotherm fit to the data. The stability of GQ structures systematically decreases as the loop length is increased. O1–O3 molecules maintain a certain fraction of folded GQ molecules at RPA concentrations as high as 1 μM, whereas all of the O4 and O5 molecules are unfolded by <20 nM of RPA.

types of unfolding patterns are observed in these measurements. In some cases, a clear intermediate FRET state is observed (Fig. 5 *A*, *lower*), whereas in other cases, the FRET signal transitions back and forth between the folded and unfolded states before stabilizing at the unfolded state (Fig. 5 *B*). Finally, in some cases, the FRET signal gradually reduces from the folded state to the unfolded state without an obvious intermediate state, as shown in Fig. 5 *C*. These intermediate states might be giving rise to the FRET populations between the folded and RPA-bound unfolded states in the steady-state histograms in Figs. 2 and 3. Regardless of the details of the unfolding process, the unfolding time is determined as the time between the stable high-FRET state and the stable low-FRET state, as shown in Fig. 5, *A–C*.

The movies for these measurements were acquired at a rate of 0.035 s/frame. However, the uncertainty in determining the beginning and end of the transition reduces our time resolution to 0.10 s for these measurements. Fig. 5 *D* shows the unfolding-time histograms for the constructs L2–L5, and Fig. 5 *E* shows similar histograms for O1–O5. The unfolding time represents a multistep process including GQ unfolding and RPA binding. The histograms in Fig. 5 *D* were fit by Gaussian curves, and the peak values were as

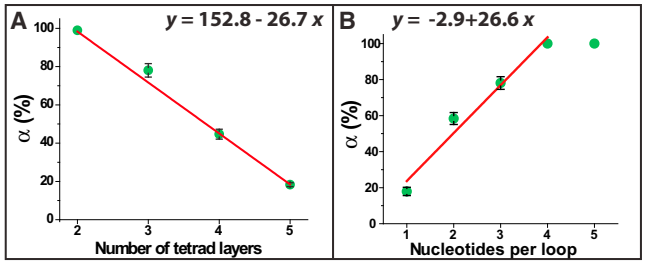

FIGURE 4 (*A*) The α parameter, the percentage of GQ molecules unfolded and bound by RPA at saturating RPA concentration, is plotted for L2–L5, which have two to five G-tetrad layers, respectively. (*B*) The α parameter for O1–O5 constructs, which have two to five nucleotides, respectively, in each loop. The α parameter is obtained from the Langmuir binding isotherm analysis, as shown in Figs. 2 *B* and 3 *B*. The red lines are linear fits to the data, which are consistent with both data sets. The O5 data are not included in the fit, as 100% unfolding is already attained at the O4 construct.

follows: 0.24 ± 0.10 s, 0.27 ± 0.10 s, 0.27 ± 0.10 s, and 0.33 ± 0.10 s for L2–L5, respectively. Gaussian fits to histograms in Fig. 5 *E* resulted in peaks of 0.39 ± 0.10 s, 0.33 ± 0.10 s, 0.27 ± 0.10 s, 0.27 ± 0.10 s, and 0.42 ± 0.10 s for O1–O5, respectively. Given the uncertainty in determining the transition time, all constructs are essentially unfolded by RPA within very similar times of ~0.35 ± 0.10 s. To ascertain whether the observed unfolding time was due to a limitation of the image acquisition time used for these measurements, we used half the CCD screen for image acquisition and reduced the image acquisition time to 0.018 s. The RPA flow experiment for L2 was repeated using this improved time resolution, and an unfolding time of 0.27 ± 0.05 s was found, which is consistent with the measurements at longer image-acquisition time (see Fig. S10 for a histogram and a sample trace). An example of a trace at this higher time resolution is shown in Fig. 5 *C*.

A model of RPA-mediated GQ unfolding

Unfolding of all the GQ constructs within very similar times is an interesting result, especially considering the orders-of-magnitude difference in their steady-state stabilities. It is important to consider the specifics of the RPA structure to better interpret these results. RPA is a heterotrimeric protein with subunits RPA1 (70 kDa), RPA2 (32 kDa), and RPA3 (14 kDa), and it has six ssDNA-binding domains (DBDs). Four of these DBDs (DBD-A, DBD-B, DBD-C, and DBD-D) are active and can bind to 8- to 30-nucleotide-long ssDNAs depending on the number of DBDs involved (49–52). For a recent review of the structure, DNA binding properties, and role in replication of RPA, see Prakash and Borgstahl (53). It is widely accepted that initial binding of RPA to ssDNA is achieved via binding of DBD-A and DBD-B, both in RPA1, and this initial interaction is further stabilized by successive binding of DBD-C and DBD-D. DBD-A and DBD-B each have a footprint of three nucleotides, and these three-nucleotide stretches are separated by a two-nucleotide stretch (52). This binding mode effectively enables RPA to stably bind to ssDNAs as short as eight nucleotides. However, three nucleotides may be enough for achieving an initial binding of DBD-A or DBD-B to ssDNA. In the rest of the discussion, we assume that DBD-A binds first and DBD-B binds second to make the description of our model easier to follow.

Based on the variation of stabilities we observed for different GQ constructs, and taking into account the structure of RPA, we propose a two-step model for RPA-mediated GQ unfolding (Fig. 6). The first step starts with RPA establishing contact with the GQ structure via binding of DBD-A to the available single-stranded regions, e.g., either the loops or the overhangs. The longer the loops or the overhangs, the easier it is to establish this initial contact. After this initial contact is achieved, RPA interacts with the folded GQ structure, which eventually leads to destabilization of the GQ and binding of DBD-B. The thermal stability of the GQ is an important parameter to consider here, as it would be more difficult for RPA to destabilize a GQ with higher thermal stability, e.g., more layers or shorter loops. The other important parameter is the affinity of DBD-B for ssDNA, which determines the level of the interactions between RPA and GQ. The second step of this model starts with destabilization of the GQ and is followed by binding of DBD-C and DBD-D to ssDNA as the GQ unfolds. Destabilization of the GQ is the onset of the FRET change we observe in RPA flow measurements. This model is consistent with the findings of SELEX studies suggesting that binding of either DBD-A or DBD-B results in a weakly bound state that would then enable RPA to invade longer stretches of DNA and destabilize the GQ (35). GQ structures that have four or five nucleotide loops (O4 and O5) are efficiently and completely unfolded by RPA, whereas those that have three or fewer nucleotide loops (O1–O3) are unfolded less efficiently and cannot attain complete unfolding of all molecules. For the cases of 1- to 2-nucleotide-long loops or overhangs, transient melting of the hydrogen bonds between the guanines that form the G-tetrad would be necessary before DBD-A or DBD-B can bind to the GQ. Such a melting event could take place due to thermal fluctuations or possibly due to RPA-GQ interactions. For the case of L2–L5 constructs, which all have three nucleotide loops, the stability of the GQ structure against thermal fluctuations or RPA-induced destabilization would increase with each additional layer. This would result in an increased stability against RPA-mediated GQ unfolding, as demonstrated in Fig. 2.

The observed dynamics of RPA-mediated GQ unfolding and unfolding of all the GQ constructs within very similar times can also be understood in the context of the two-step model proposed in Fig. 6. The unfolding time we measure in these measurements is essentially a measure of the second step of this interaction (Fig. 6, *right*). This step starts with measurable destabilization of GQ followed by binding

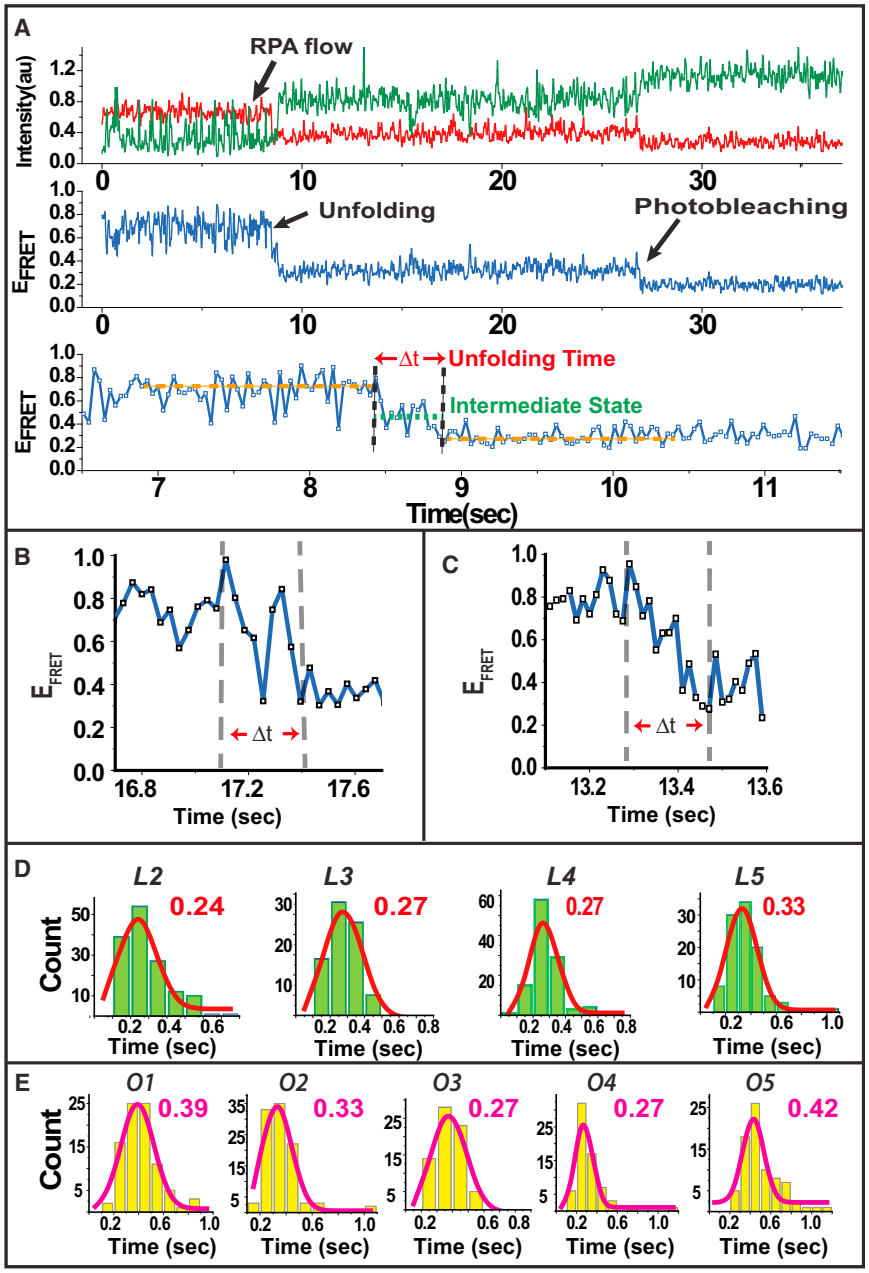

FIGURE 5 (*A*) An RPA flow experiment showing unfolding of a GQ by RPA in real time. As GQ is unfolded by RPA, FRET efficiency transitions from a high level, corresponding to the folded GQ structure, to a low level, corresponding to RPA-bound unfolded DNA. The bottom panel shows the section of the time trajectory that is in the vicinity of RPA binding and unfolding of the GQ. The unfolding time is marked by red arrows. We occasionally observed intermediate states during unfolding (*green dotted line*). (*B*) Example of an unfolding event in which the GQ molecule transitions back and forth between the folded and unfolded states. (*C*) Example of an unfolding event in which the transition from the folded to the unfolded state occurs gradually, without a distinct stable intermediate state. (*D*) Histograms of unfolding times (Δt) for L2–L5. (*E*) Histograms of unfolding times (Δt) for O1–O5. The red curves in *D* and *E* are Gaussian fits to the data. Interestingly all the constructs have similar unfolding times despite the large difference in their steady-state stability.

of DBD-C and DBD-D as more ssDNA becomes available. The first step is not detected, as the GQ is still folded and FRET efficiency does not change significantly. As we do not observe any variation in the duration of the second step of the unfolding mechanism, we conclude that the first step determines the stability of GQ against RPA-mediated

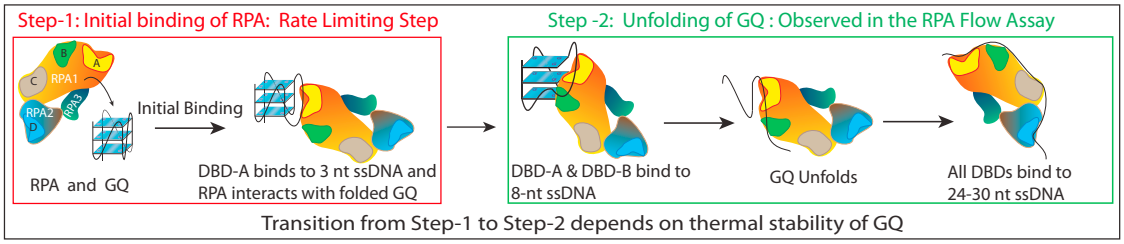

FIGURE 6 Cartoon of a model describing RPA-mediated GQ unfolding as a two-step process.

unfolding. The stability of GQ that is described in the first step of this model depends on both the thermal stability of the GQ and the length of available ssDNA in the loops or overhangs, which enable RPA to position in the vicinity of GQ and interact with it.

CONCLUSION

In conclusion, this study provides a systematic way to characterize the stability of GQs in terms of protein-mediated unfolding. RPA, the most abundant ssDNA-binding protein in eukaryotes, is expected to frequently interact with nontelomeric PQSs. Our data identify the effects of number of tetrad layers and loop lengths on the stability of GQ against RPA-mediated unfolding. Clearly, live-cell environment is significantly more complicated than the environment in in vitro assays, and there are various other factors that are known to destabilize GQ structures, e.g., helicases or competition with the complementary DNA strand. It is also possible that there might be direct or indirect mechanisms or factors that could stabilize the GQ in the context of chromatin. Nevertheless, this study provides what we believe to be a new perspective on the interactions of an abundant and very important protein with systematically varied GQ structures. Our results impose strict constraints on physiologically viable GQ structures. For example, GQs with four- or five-nucleotide-long loops are completely unfolded by RPA, suggesting that GQ structures with such long loops are not viable in a physiological setting. Finally, we propose a two-step model of RPA-GQ interactions that is consistent with our data. The quantitative data we present would be particularly important in providing guidelines for computational work on modeling the interaction of an arbitrary GQ structure with RPA. Extension of these studies to a broader class of GQs with systematically varied structures or with inclusion of multiprotein complexes would have the potential to make better estimates for the physiological viability of GQ structures.

SUPPORTING MATERIAL

Ten figures and supporting methods are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(13)00427-X.

We thank Prof. Aziz Sancar and members of his group for providing us with RPA. M.H.Q. thanks ICAM for funding during his training in the laboratory of Prof. Aziz Sancar. H.B. designed the experiments, S.R., M.H.Q., D.W.M., J.B.B., and U.Ç. performed the experiments and the analysis and edited the manuscript, and H.B. and S.R. wrote the article.

This work was supported by the start-up funds of H.B. from Kent State University.

REFERENCES

1. Blackburn, E. H. 1991. Structure and function of telomeres. *Nature.* 350:569–573.
2. Gellert, M., M. N. Lipsett, and D. R. Davies. 1962. Helix formation by guanylic acid. *Proc. Natl. Acad. Sci. USA.* 48:2013–2018.
3. Gilbert, D. E., and J. Feigon. 1999. Multistranded DNA structures. *Curr. Opin. Struct. Biol.* 9:305–314.
4. Williamson, J. R. 1994. G-quartet structures in telomeric DNA. *Annu. Rev. Biophys. Biomol. Struct.* 23:703–730.
5. Chaires, J. B. 2010. Human telomeric G-quadruplex: thermodynamic and kinetic studies of telomeric quadruplex stability. *FEBS J.* 277:1098–1106.
6. Gray, R. D., and J. B. Chaires. 2008. Kinetics and mechanism of K^+- and Na^+-induced folding of models of human telomeric DNA into G-quadruplex structures. *Nucleic Acids Res.* 36:4191–4203.
7. Lane, A. N., J. B. Chaires, …, J. O. Trent. 2008. Stability and kinetics of G-quadruplex structures. *Nucleic Acids Res.* 36:5482–5515.
8. Stegle, O., L. Payet, …, J. H. Leon. 2009. Predicting and understanding the stability of G-quadruplexes. *Bioinformatics.* 25:i374–i382.
9. Lipps, H. J., and D. Rhodes. 2009. G-quadruplex structures: in vivo evidence and function. *Trends Cell Biol.* 19:414–422.
10. Verma, A., V. K. Yadav, …, S. Chowdhury. 2009. Evidence of genome-wide G4 DNA-mediated gene expression in human cancer cells. *Nucleic Acids Res.* 37:4194–4204.
11. Burge, S., G. N. Parkinson, …, S. Neidle. 2006. Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.* 34:5402–5415.
12. Reference deleted in proof.
13. Eddy, J., and N. Maizels. 2006. Gene function correlates with potential for G4 DNA formation in the human genome. *Nucleic Acids Res.* 34:3887–3896.
14. Eddy, J., and N. Maizels. 2008. Conserved elements with potential to form polymorphic G-quadruplex structures in the first intron of human genes. *Nucleic Acids Res.* 36:1321–1333.
15. Huppert, J. L., and S. Balasubramanian. 2005. Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.* 33:2908–2916.
16. Reference deleted in proof.
17. Blackburn, E. H., C. W. Greider, and J. W. Szostak. 2006. Telomeres and telomerase: the path from maize, Tetrahymena and yeast to human cancer and aging. *Nat. Med.* 12:1133–1138.
18. Kumari, S., A. Bugaut, …, S. Balasubramanian. 2007. An RNA G-quadruplex in the 5′ UTR of the NRAS proto-oncogene modulates translation. *Nat. Chem. Biol.* 3:218–221.
19. Morris, M. J., and S. Basu. 2009. An unusually stable G-quadruplex within the 5′-UTR of the MT3 matrix metalloproteinase mRNA represses translation in eukaryotic cells. *Biochemistry.* 48:5313–5319.
20. Morris, M. J., Y. Negishi, …, S. Basu. 2010. An RNA G-quadruplex is essential for cap-independent translation initiation in human VEGF IRES. *J. Am. Chem. Soc.* 132:17831–17839.
21. Huppert, J. L., A. Bugaut, …, S. Balasubramanian. 2008. G-quadruplexes: the beginning and end of UTRs. *Nucleic Acids Res.* 36:6260–6268.
22. Balasubramanian, S., L. H. Hurley, and S. Neidle. 2011. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat. Rev. Drug Discov.* 10:261–275.
23. Balasubramanian, S., and S. Neidle. 2009. G-quadruplex nucleic acids as therapeutic targets. *Curr. Opin. Chem. Biol.* 13:345–353.
24. Biffi, G., D. Tannahill, …, S. Balasubramanian. 2013. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* 5:182–186.
25. Paeschke, K., J. A. Capra, and V. A. Zakian. 2011. DNA replication through G-quadruplex motifs is promoted by the *Saccharomyces cerevisiae* Pif1 DNA helicase. *Cell.* 145:678–691.
26. Paeschke, K., K. R. McDonald, and V. A. Zakian. 2010. Telomeres: structures in need of unwinding. *FEBS Lett.* 584:3760–3772.
27. Qureshi, M. H., S. Ray, …, H. Balci. 2012. Replication protein A unfolds G-quadruplex structures with varying degrees of efficiency. *J. Phys. Chem. B.* 116:5588–5594.

28. Kim, C., and M. S. Wold. 1995. Recombinant human replication protein A binds to polynucleotides with low cooperativity. *Biochemistry.* 34:2058–2064.

29. Wold, M. S. 1997. Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu. Rev. Biochem.* 66:61–92.

30. Iftode, C., Y. Daniely, and J. A. Borowiec. 1999. Replication protein A (RPA): the eukaryotic SSB. *Crit. Rev. Biochem. Mol. Biol.* 34:141–180.

31. Oakley, G. G., and S. M. Patrick. 2010. Replication protein A: directing traffic at the intersection of replication and repair. *Front. Biosci.* 15:883–900.

32. Fan, J. H., E. Bochkareva, …, D. M. Gray. 2009. 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.

33. Wu, Y., N. Rawtani, …, R. M. Brosh, Jr. 2008. Human replication protein A melts a DNA triple helix structure in a potent and specific manner. *Biochemistry.* 47:5068–5077.

34. Salas, T. R., I. Petruseva, …, C. Saintomé. 2006. Human replication protein A unfolds telomeric G-quadruplexes. *Nucleic Acids Res.* 34:4857–4865.

35. Prakash, A., A. Natarajan, …, G. E. Borgstahl. 2011. Identification of the DNA-binding domains of human replication protein A that recognize G-quadruplex DNA. *J. Nucleic Acids.* 2011:896947.

36. Ha, T., T. Enderle, …, S. Weiss. 1996. Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc. Natl. Acad. Sci. USA.* 93:6264–6268.

37. Binz, S. K., A. M. Dickson, …, M. S. Wold. 2006. Functional assays for replication protein A (RPA). *Methods Enzymol.* 409:11–38.

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

39. Choi, J. H., L. A. Lindsey-Boltz, …, A. Sancar. 2010. Reconstitution of RPA-covered single-stranded DNA-activated ATR-Chk1 signaling. *Proc. Natl. Acad. Sci. USA.* 107:13660–13665.

40. Hazel, P., J. Huppert, …, S. Neidle. 2004. Loop-length-dependent folding of G-quadruplexes. *J. Am. Chem. Soc.* 126:16405–16415.

41. Phan, A. T., and D. J. Patel. 2003. Two-repeat human telomeric d(TAGGGTTAGGGT) sequence forms interconverting parallel and antiparallel G-quadruplexes in solution: distinct topologies, thermodynamic properties, and folding/unfolding kinetics. *J. Am. Chem. Soc.* 125:15021–15027.

42. Zhang, N., A. T. Phan, and D. J. Patel. 2005. (3 + 1) Assembly of three human telomeric repeats into an asymmetric dimeric G-quadruplex. *J. Am. Chem. Soc.* 127:17277–17285.

43. Heddi, B., and A. T. Phan. 2011. Structure of human telomeric DNA in crowded solution. *J. Am. Chem. Soc.* 133:9824–9833.

44. Lao, Y., C. G. Lee, and M. S. Wold. 1999. Replication protein A interactions with DNA. 2. Characterization of double-stranded DNA-binding/helix-destabilization activities and the role of the zinc-finger domain in DNA interactions. *Biochemistry.* 38:3974–3984.

45. Walther, A. P., X. V. Gomes, …, M. S. Wold. 1999. Replication protein A interactions with DNA. 1. Functions of the DNA-binding and zinc-finger domains of the 70-kDa subunit. *Biochemistry.* 38:3963–3973.

46. Lee, J. Y., B. Okumus, …, T. Ha. 2005. Extreme conformational diversity in human telomeric DNA. *Proc. Natl. Acad. Sci. USA.* 102:18938–18943.

47. Okumus, B., and T. Ha. 2010. Real-time observation of G-quadruplex dynamics using single-molecule FRET microscopy. *Methods Mol. Biol.* 608:81–96.

48. Murphy, M. C., I. Rasnik, …, T. Ha. 2004. Probing single-stranded DNA conformational flexibility using fluorescence spectroscopy. *Biophys. J.* 86:2530–2537.

49. Fanning, E., V. Klimovich, and A. R. Nager. 2006. A dynamic model for replication protein A (RPA) function in DNA processing pathways. *Nucleic Acids Res.* 34:4126–4137.

50. Bastin-Shanower, S. A., and S. J. Brill. 2001. Functional analysis of the four DNA binding domains of replication protein A. The role of RPA2 in ssDNA binding. *J. Biol. Chem.* 276:36446–36453.

51. Bochkarev, A., and E. Bochkareva. 2004. From RPA to BRCA2: lessons from single-stranded DNA binding by the OB-fold. *Curr. Opin. Struct. Biol.* 14:36–42.

52. Cai, L., M. Roginskaya, …, Y. Zou. 2007. Structural characterization of human RPA sequential binding to single-stranded DNA using ssDNA as a molecular ruler. *Biochemistry.* 46:8226–8233.

53. Prakash, A., and E. O. Borgstahl. 2012. The structure and function of replication protein A in DNA replication. *In* The Eukaryotic Replisome: A Guide to Protein Structure and Function. S. MacNeill, editor. Springer, New York. 171–196.

Supporting Material

RPA Mediated Unfolding for Systematically Varying G-quadruplex Structures

Sujay Ray[1], Mohammad H. Qureshi[2], Dominic Malcolm[1], Jagat B. Budhathoki[1], Uğur Çelik[3], Hamza Balci[1,*]

[1] Physics Department, Kent State University, Kent, OH 44242 USA
[2] Biological Sciences Department, Kent State University, Kent, OH 44242 USA
[3] Department of Genetics and Bioengineering, Fatih University, Istanbul, Turkey

* To whom correspondence should be addressed. Tel: +1-330-672-2577; Fax: +1-330-672-2959; Email: hbalci@kent.edu

The authors declare no competing financial interests.

Circular Dichroism

Samples were heated at 95°C for 10 minutes and then were allowed to gradually cool down to room temperature over 2-3 hours by placing the heating block at ambient temperature. CD samples contained 4 µM DNA and either no salt (Fig. S1-A and Fig. S1-B), 150 mM LiCl (Fig. S1-C and Fig. S1-D), or 150 mM KCl (Fig.1). 10 mM Tris at pH 7.5 was used as buffer in all measurements, and 2 mM $MgCl_2$ was added to the Li^+ and K^+ measurements to simulate physiological conditions. All measurements were performed at room temperature on a Jasco 810 Spectrapolarimeter (Easton, MD), using a quartz cuvette with a 1mm optical path length. The spectra obtained are averages of three scans taken over a range of 200nm to 320nm excitation wavelength with a measurement taken every 0.5 nm at a scanning rate of 50nm/minute. Representative spectra have been baseline subtracted from a buffer-only sample, and smoothed using a Savitzky-Golay function. It is clear from the spectra shown in Fig. S1-A and Fig. S1-B that most constructs do not demonstrate a signature of GQ formation in the absence of salt. Exception to this are constructs such as L5 and O1. These constructs show a weak signature of GQ formation suggesting that they form GQ even in the absence of salts. Incubating the constructs in 150 mM Li^+ stabilizes the GQ structure for some constructs, such as L4 and L5, but most constructs do not show a signature of GQ (Fig. S1-C and Fig. S1-D).

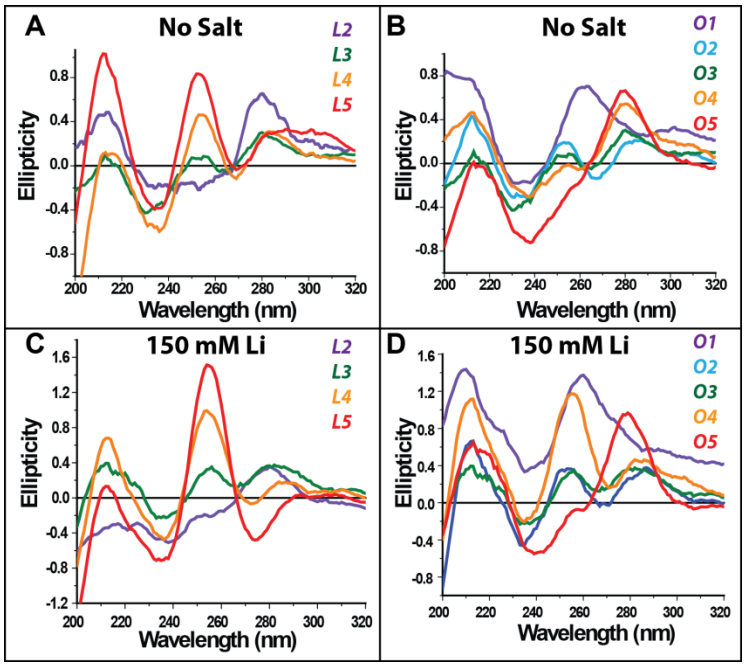

Fig. S1: CD spectra for (A) L2-L5 constructs (B) O1-O5 constructs in absence of salt. CD spectra in 150 mM Li^+ for (C) L2-L5 constructs (D) O1-O5 constructs.

smFRET Assay and Possible Folding conformations

Prism-type total internal reflection microscopy was used to perform the single molecule measurements. A schematic of the assay is shown in Fig. S2-A. The folded, unfolded, and RPA-bound unfolded states of the ssDNA tail have different FRET efficiencies, and can be distinguished from each other. The sequence of the ssDNA tail is varied for different GQ structures. The duplex stem and the donor-acceptor fluorophores (Cy3-Cy5) are kept the same for all studied constructs. Fig. S2-B shows one of the possible folding conformations for each GQ constructs based on the circular dichroism data presented in Fig. 1.

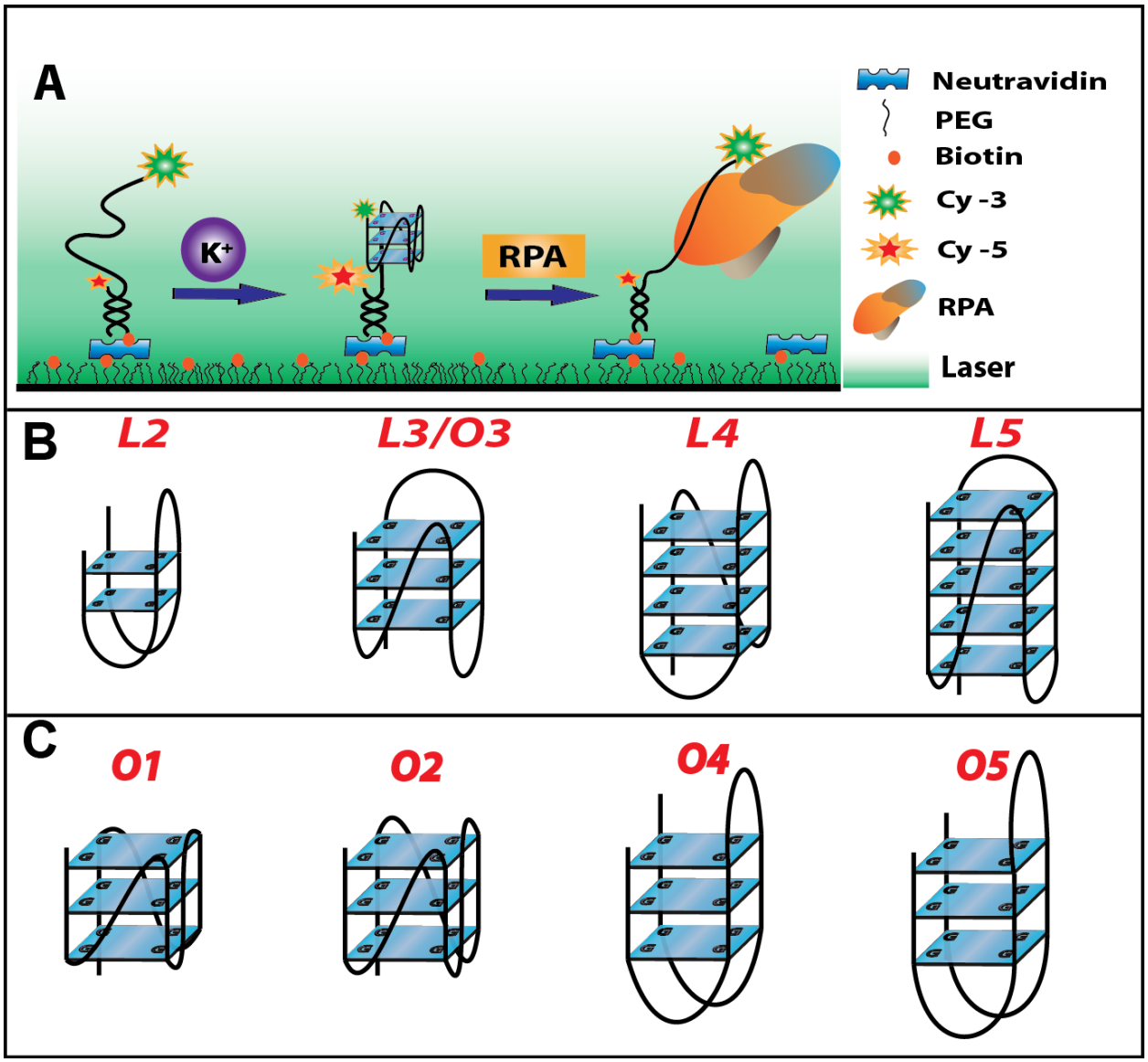

Fig. S2. (A) A schematic of the smFRET assay. (B) One of the possible folding conformations for each GQ constructs L2-L5. **(C)** One of the possible folding conformations for each GQ constructs O1-O5.

smFRET Data Analysis

Background subtraction was done manually for each molecule, by subtracting both donor and acceptor intensities after donor photobleaching. Overlap between donor and acceptor emission spectra gives rise to leakage of the donor signal into the acceptor channel, resulting in a non-zero FRET value for molecules that had a donor but not an acceptor fluorophore (donor only (DO) molecules). The dichroic we use to split the emissions of Cy3-Cy5 pair results in a donor only peak at 0.12 before manual subtraction of background, which reduces to 0.06 after manual subtraction. For all the data presented in Fig.2 and Fig. 3 background has been manually subtracted by inspecting each time trace. This DO peak is then subtracted from the FRET histograms and histograms rescaled to compensate for the leakage into acceptor channel. Fig. S3 is a representative histogram, showing the position of the DO peak, RPA bound unfolded peak and folded GQ peak before correction for the leakage factor. It also shows that the DO peak is distinctly different from the RPA bound unfolded peak, allowing us to subtract the DO peak without affecting any other population.

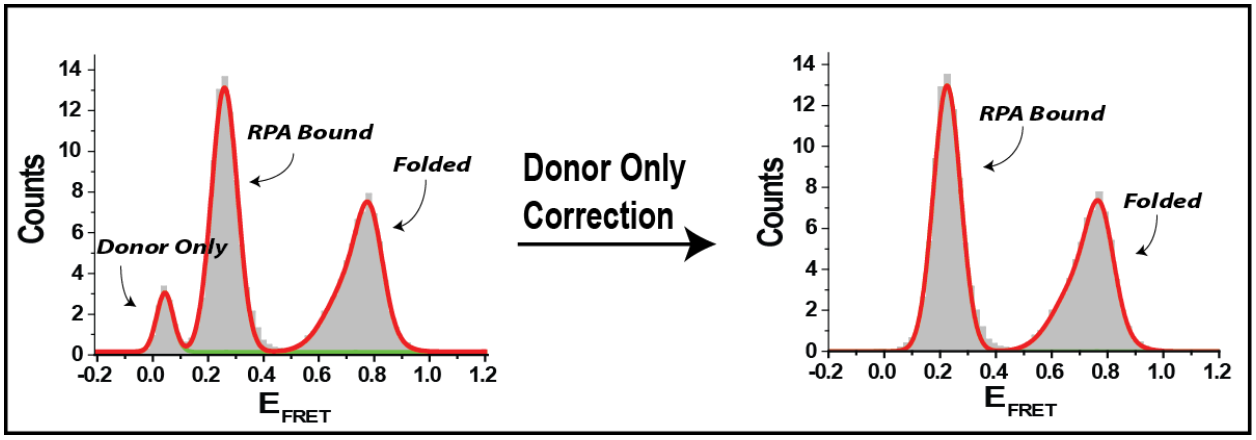

Fig. S3: Representative FRET histogram, showing the position of different peaks.

RPA does not bind to the double stranded DNA Stem

RPA has at least three orders of magnitude lower affinity for dsDNA compared to ssDNA. This is important as we have an 18 bp duplex stem in all our DNA constructs, and modification of this stem by RPA would cause complications in our analysis. In order to experimentally verify that RPA does not interact with dsDNA under our experimental conditions, we prepared a double stranded DNA construct that is internally labeled with a Cy3 fluorophore in one strand and a Cy5 fluorophore at the 3' end of the other strand. The sequences of the oligos we used for these measurements are:

5'-TGG CGA CG/Cy3/G CAG CGA GGC

5'-biotin-GCC TCG CTG CCG TCG CCA-Cy5

These oligos form an identical duplex DNA as that used for all our other measurements. The data shown below in Fig. S4 demonstrates that RPA does not bind to this duplex stem even at the highest protein concentration used in our studies (1 μM RPA).

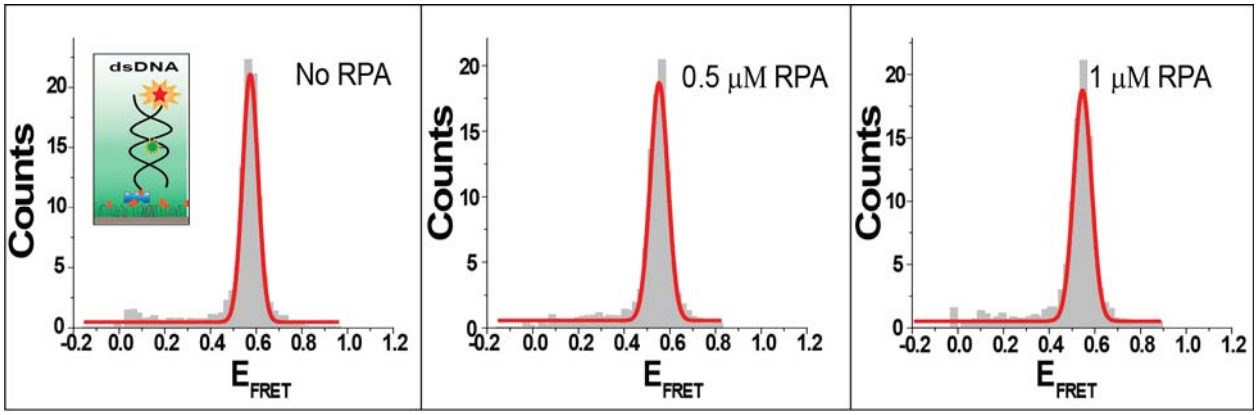

Fig. S4. RPA does not bind to the 18 bp long dsDNA stem . smFRET measurements on a 18 bp duplex internally labeled with Cy3 and end labeled with Cy5. (Left)smFRET data on duplex DNA before RPA is introduced into the chamber. (Middle) 0.5 μM RPA is introduced to the chamber but no change is observed in the FRET efficiency. (Right) 1 μM RPA is introduced to the chamber but no change is observed in the FRET efficiency. This is the highest RPA concentration we used in our measurements.

Folding of GQ as a function of increasing K^+ ion:

Fig. S5 shows progressive folding of L3 construct with the increase of K^+ concentration. smFRET is capable of detecting various folding conformations adopted by DNA at different K^+ concentrations. However, for the current study we are only concerned with the conformation attained by the GQ structures at 150 mM K^+ concentration.

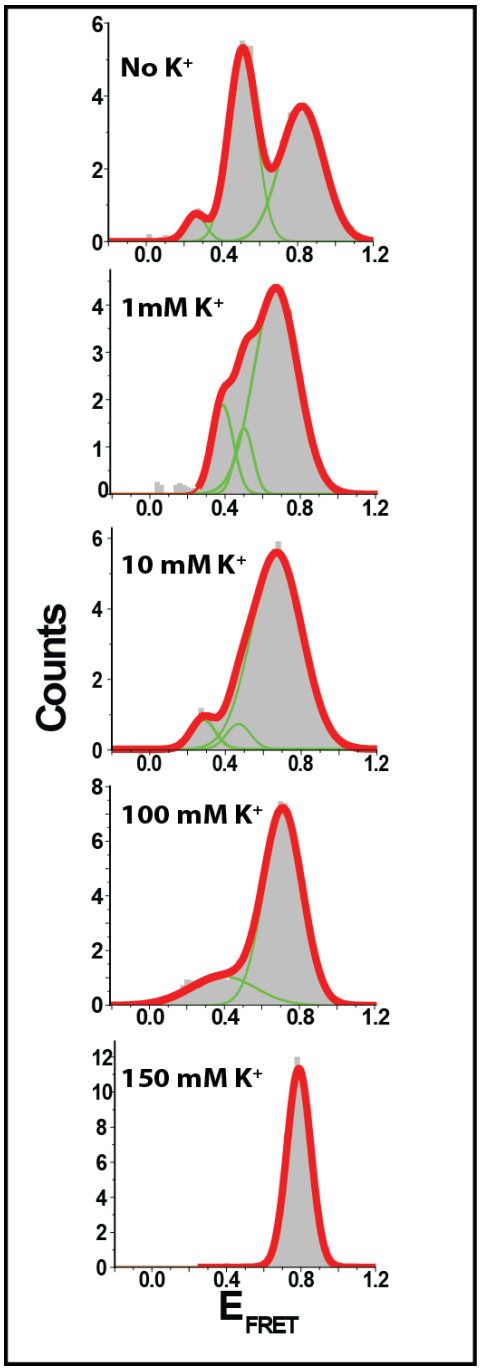

Fig. S5. Different folding conformation adopted by L3 construct as K^+ concentration is titrated to 150 mM.

Comparison of FRET Peaks for Different States

There are several different states involving RPA and DNA that need to be resolved in our measurements. Folded GQ, unfolded DNA, and RPA bound unfolded DNA are some of these states. We performed control measurements to ensure that these different states can be distinguished based on the FRET peaks that represent them.

The first of these measurements aimed distinguishing between the unfolded state and RPA bound unfolded state. Clearly identifying the unfolded state (coiled ssDNA) at 150 mM K^+ using the GQ forming constructs we used for these studies is not possible as they fold into GQ structure at such ionic strength. Even under 150 mM Li^+, which is a weaker stabilizer of GQ structure compared to K^+, a significant proportion of folding occurs. Therefore, we used a partial duplex DNA that has an overhang of 21 thymines (called pdT21) which do not form any secondary structure. This DNA has a length similar to the O2 construct which has an overhang of 22 nucleotides. We compared the FRET peak of pdT21 observed at 150 mM K^+ concentration before and after adding 0.1 nM RPA (Fig. S6A-B). For reference we also show the folded peak and the RPA bound unfolded peak of the O2 construct (Fig. S6C-D). As expected, the RPA bound unfolded peak of O2 and that of pdT21 match and they are both close 0.20 FRET efficiency. The unfolded DNA peak of pdT21 is significantly different from this peak and occurs at 0.55 FRET efficiency. The folded GQ peak is significantly different from both of these peaks and occurs at 0.85 FRET efficiency. Therefore, all three states can be distinguished from each other based on the corresponding FRET peak that represents them. In general, the RPA bound unfolded peak for the eight GQ forming constructs we studied are in the 0.10-0.20 FRET efficiency range. On the other hand, the unfolded DNA peaks for these constructs are in the 0.35-0.55 FRET efficiency range. Finally, the folded GQ states have FRET efficiencies are in the 0.70-0.85 range.

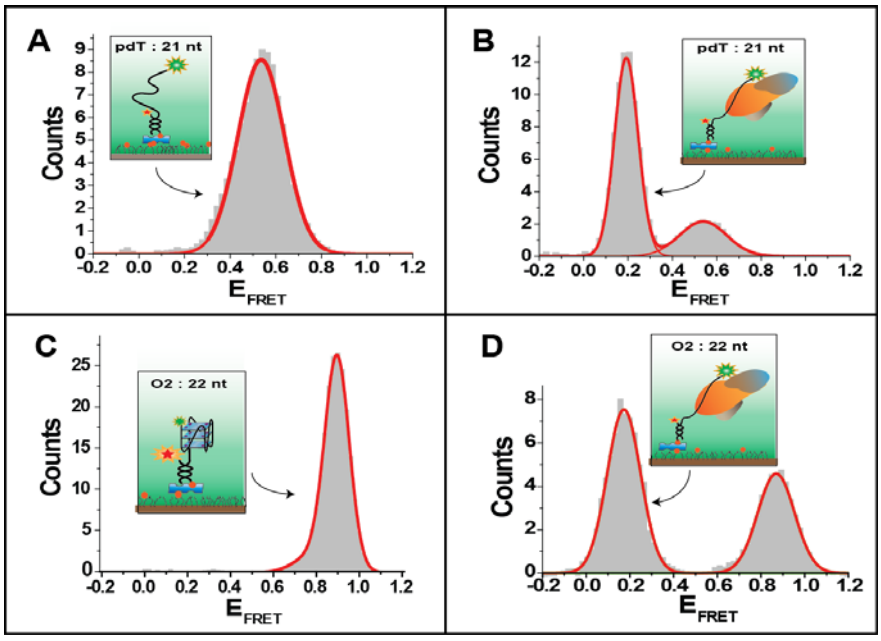

Fig. S6. (A) smFRET data on pdT21 construct in the absence of RPA. A single peak at 0.55 FRET efficiency is observed. (B) smFRET data on pdT21 after introducing RPA to the chamber. A second peak at 0.2 FRET efficiency is observed, which represents the RPA bound peak. (C) smFRET data on O2 in the absence of RPA. A single peak representing the folded GQ is observer. (D) smFRET data on O2 after introducing RPA. A second peak representing the unfolded and RPA bound state emerges at 0.20 FRET efficiency.

Another related question, is whether RPA binding removes all the secondary structure of the folded conformation. This might particularly be an issue with longer constructs such as L4 or L5. In order to check if the RPA bound peak still has some secondary structure associated with it, we performed control measurements on partial duplex DNA with 30 nucleotide long poly-thymine tail (will be called pdT30). The duplex stem and the position of the dyes in pdT30 are identical to those in the GQ constructs and the only difference is in the single stranded tail which is composed of 30 thymines rather than a GQ forming sequence. pdT30 is not expected to have any secondary structure and therefore the FRET value for the RPA bound state of this construct can be taken as a reference. The histogram for this construct is shown in Fig. S7-A which has an RPA bound FRET peak at 0.11, consistent with our GQ constructs that are around 30 nucleotides long. For comparison we also show the data on the GQ construct L4 which has 29 nucleotides at the single stranded tail (Fig. S7-B). The RPA bound peak for L4 is at 0.13 FRET, similar to that for pdT30. Hence, we do not think any secondary structure remains folded once RPA binds to the unfolded GQ constructs. As RPA binds to about 30 nucleotides of ssDNA, it might be expected that the FRET efficiency of the RPA bound state would be lower than 0.13 FRET if the DNA was stretched in a linear fashion. However, the geometry of the complex RPA forms with ssDNA is not linear but more like an L-shape. Fig. S7-C demonstrates this geometry (adapted from Figure 5 of Ref. (1)).

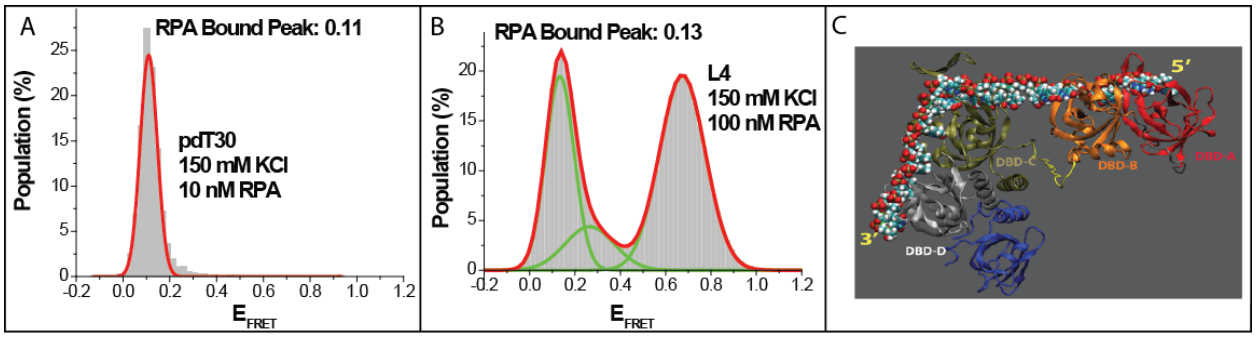

Fig. S7. (A) RPA bound peak for a pdDNA with 30 nucleotide long polythymine ssDNA tail. (B) RPA bound peak for GQ forming construct L4 which has an ssDNA tail of 29 nucleodies long. (C) Structure of RPA-bound DNA. DNA is not linearly stretched by RPA but forms an L-shape. Fig. S7-C is adapted from Figure 5 of Ref. (1).

Influence of RPA binding and GQ Folding on the Fluorescence Properties of Fluorophores

To quantify the influence of RPA binding and GQ folding on the fluorescence properties of the fluorophores we measured the gamma parameter of L3 (O3) molecules at FRET levels corresponding to folded GQ and RPA-bound unfolded GQ. Gamma parameter is defined as: $\gamma = \frac{\phi_A \eta_A}{\phi_D \eta_D}$, where ϕ_A and ϕ_D are the donor and acceptor intensities, respectively, and η_A and η_D are the sensitivities of sensor to the donor and acceptor emissions, respectively (2). Fig. S8-A shows an example trace demonstrating how the gamma parameter is determined experimentally. It is essentially the ratio of the change in the donor intensity to the change in the donor intensity upon photobleaching of the acceptor:

$$\gamma = \frac{I_A^{pre-photobleach} - I_A^{post-photobleach}}{I_D^{post-photobleach} - I_D^{pre-photobleach}}.$$

In order to determine the gamma parameter for the Cy3-Cy5 pair when they are attached to a partial duplex DNA similar to that we used in our experiments, we performed measurements on a partial duplex DNA with the same duplex stem as that used in all our GQ constructs and a polythymine tail of 21 nucleotides (pdT21). The inset in the top panel of Fig. S8B shows a schematic of the construct. Similar to the GQ constructs we studied, the Cy5 was placed at the ssDNA/dsDNA junction and Cy3 was placed at the end of the ssDNA tail. Gamma measurements for this construct were performed under 150 mM KCl and pH 7.4, as in our measurements for the GQ constructs. The FRET measurements for this constructs showed a peak at 0.54 FRET before performing any gamma correction, in agreement with reported measurements in literature (3). Fig. S8-B shows a 2D histogram of gamma measurements on 164 single molecule FRET traces. The x-axis of the histogram represents the FRET efficiency, with the gamma correction taken into account [$E_F=I_A/(I_A+\gamma I_D)$], and the y-axis represents the gamma parameter. These measurements yield $\gamma \approx 1.6 \pm 0.5$. Performing this gamma correction shifts the FRET peak to significantly lower level (peak at 0.30 FRET) as shown in Fig. S8-B. The shift in the FRET peaks are much smaller at lower or higher FRET levels, as we have in our case, due to the nature of the correction. To our knowledge, the gamma parameter for Cy3-Cy5 pair for a partial duplex DNA construct has not been reported before. The gamma parameter for the Cy3-Cy5 pair attached to double stranded DNA has been measured to be in the 0.6-1.3 range for the same EMCCD camera used in this study depending on the details of the fluorophore positions on DNA and the type of dichroics and filters used in the detection path (4). Gamma is typically assumed to be 1.0 in the literature for the Cy3-Cy5 pair (5).

We next measured the gamma parameter of Cy3-Cy5 for folded GQ and RPA-bound DNA cases. We obtained a gamma value of $\gamma \approx 0.6 \pm 0.2$ for the folded GQ state and $\gamma \approx 1.2 \pm 0.3$ for the RPA-bound unfolded state. These gamma value for the RPA-bound DNA case is very similar to the pdT21 case, which suggests that RPA does not have strong interactions with the fluorophores. The gamma value for the folded GQ case is significantly different from the pdDNA case, which suggests that the folded GQ interacts with the fluorophores despite the two thymine spacer placed between the GQ and Cy3 fluorophore. However, these interactions are not significantly different from those observed when these fluorophores are attached at the ends of duplex DNA, in which gamma in the 0.6-1.3 range have been observed.

The spread of the gamma values in the RPA case is significantly larger than that of the folded GQ case, but not different from the pdT21 case. This variation is partially due to the difficulty in determining the gamma at low FRET values when the acceptor signal is weak. The gamma values demonstrated in Fig. S8

were obtained from the photobleaching steps of acceptor fluorophores. Up on acceptor photobleaching, the acceptor intensity reduces to the donor only level (leakage from the donor emission into the acceptor channel) and the donor intensity increases correspondingly (see Fig. S8-A). The ratio of the decrease in acceptor intensity to the increase in donor intensity determines the gamma parameter. At high FRET states where the acceptor signal is strong, the gamma parameter is accurately measured as the acceptor photobleaching gives rise to large intensity changes in both donor and acceptor signals. On the other hand photobleaching of the acceptor gives rise to only small decrease in the acceptor signal and a small increase in the donor signal at low FRET states where the acceptor signal is naturally weak. As the RPA-bound unfolded DNA state has an uncorrected FRET efficiency of $\approx$0.1, the changes in acceptor and donor intensities are small and the measured gamma parameters have a larger errors associated with them. Correcting for the different gamma parameters shifts the folded GQ FRET level to a slightly higher level (0.78 to 0.86 assuming γ=0.6) and the FRET level of the RPA bound unfolded state to a slightly lower level (0.10 to 0.085 assuming γ=1.2). Due to the large difference between these two FRET states, these corrections would not affect any of the equilibrium constants or the unfolding times reported in the manuscript.

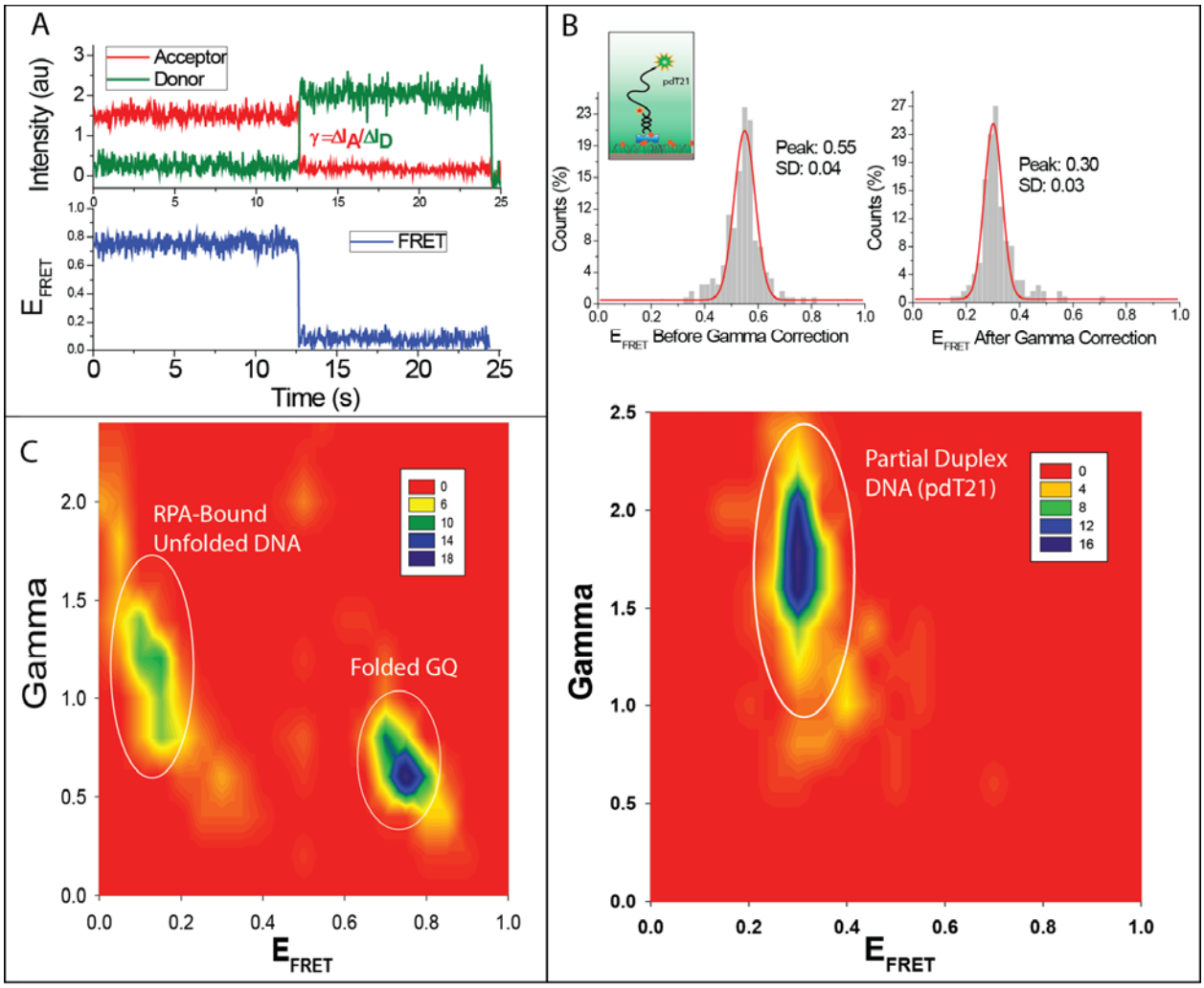

Fig. S8. Gamma parameter for RPA-bound unfolded DNA and folded GQ structures. (A) A single molecule FRET trace demonstrating how gamma is determined. (B) (Top Panel) FRET Histograms for a partial duplex DNA with a tail of 21 thymines (pdT21) before and after the gamma correction. The DNA construct is shown as an inset. (Bottom Panel) Histogram of gamma values for pdT21. (C)Histogram of gamma values for RPA bound unfolded DNA and folded GQ molecules. The numbers in the scale bar represent the number of molecules.

Monitoring RPA Mediated GQ Unfolding for Extended Time Periods.

For all steady state RPA mediated GQ unfolding measurements, we observed that RPA binds to the unfolded DNA very stably. Once GQ is unfolded and RPA binds to the unfolded ssDNA, as evidenced by a transition to the low FRET state, the GQ does not refold and transition to the high FRET state in the time frame of our measurements (120 sec). In order to study the conformational changes over longer times, we took longer movies (1200 sec) with an integration time of 500ms. These data were taken with L4 construct after 15 minutes of incubation with 50 nM RPA. We reduced the laser intensity to increase the fluorophore lifetime, and increased the integration time in order to collect enough photons per frame. Majority of the traces we collected showed either high (folded GQ-top panel of Fig. S9-A) or low FRET state (unfolded and RPA bound GQ-bottom panel of Fig. S9-A), with no transitions between the two states. Fig. S9-B shows an rare example of a molecule that transitions back and forth between the folded and RPA bound unfolded state.

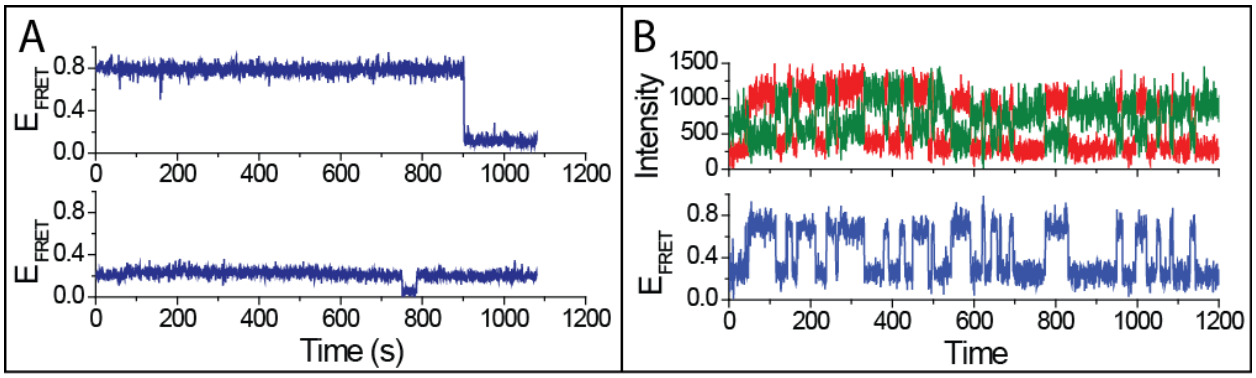

Fig. S9: Steady state measurements that were taken over longer times than those presented in the manuscript. (A)Example FRET traces showing stable folded or RPA bound states with no dynamics between the two states. Top panel shows a stably folded GQ that is not unfolded by RPA for about 900 seconds . Bottom panel shows an unfolded DNA that is bound by RPA for about 1100 seconds. The dip in the FRET level around 780 seconds is due to donor blinking. (B) An rare example trace of a molecule showing back and forth transitions between the folded and RPA bound states.

Unfolding time of GQ : 18 ms acquisition time:

The data presented in Fig. 5, on the time it takes for RPA to unfold GQ structures, were acquired at 35 millisecond image acquisition time. This limits the accuracy of the method to ±0.10 sec. In order to confirm our results with a higher time resolution, we used half of the CCD sensor on the camera for image acquisition and were able to reduce the image acquisition time to 18 ms. This method improved the time resolution by a factor of two, but the smaller area of imaging made it twice as difficult to collect statistically significant number of molecules. We performed this control measurement on the L2 construct and obtained the data presented in Fig. S10. Fig. S10-A shows a sample trace that demonstrates the clear transition from the folded to unfolded state despite the lower photon count (y-axis of top panel is a measure of the photon count). Fig. S10-B shows a histogram of unfolding times. The average unfolding times were consistent for the 18 ms and 35 ms image acquisition times within the uncertainties of the methods: 0.24±0.10 sec for 35 msec image acquisition and 0.27±0.05 sec for 18 msec image acquisition.

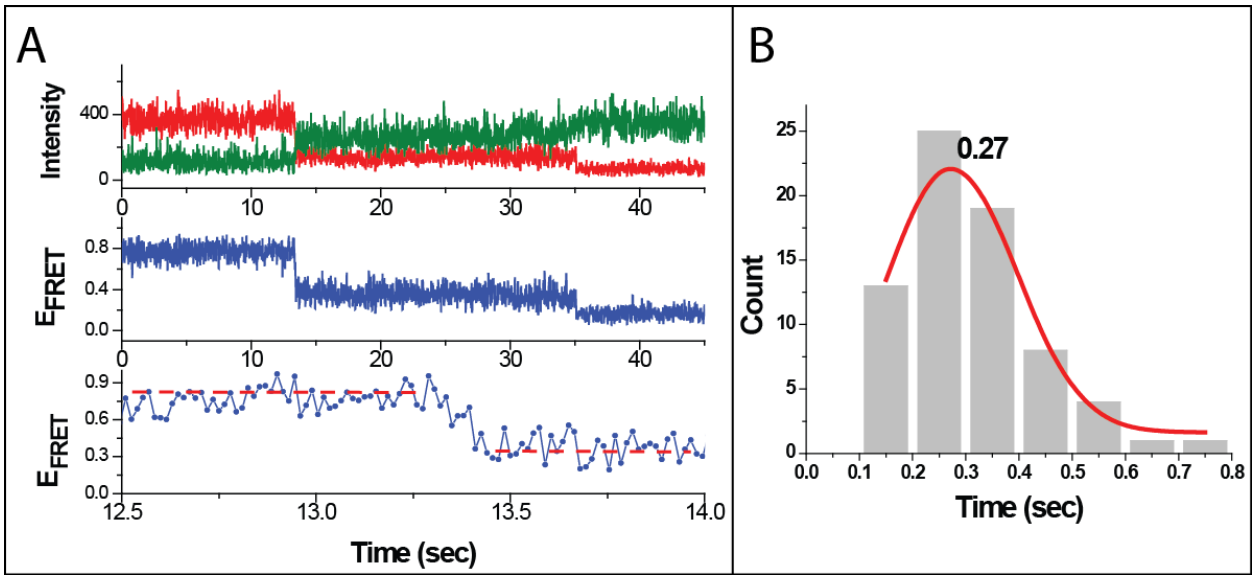

Fig. S10: (A) 18 ms Time resolution stopped-flow type experiment, showing unfolding of a GQ by RPA in real time. (B) Histogram of unfolding time for L3/O3 construct (18 ms data).

References

1. Cai, L., M. Roginskaya, Y. Qu, Z. Yang, Y. Xu, and Y. Zou. 2007. Structural characterization of human RPA sequential binding to single-stranded DNA using ssDNA as a molecular ruler. Biochemistry 46:8226-8233.
2. Dahan, M., Deniz, A., Ha, T., Grunwel, J., Schultz, P.G., Chemla, D.D., and Weiss S. 1999. Ratiometric identification and separation of single molecules diffusing in solution. Chem. Phys. 247:85-106.
3. Murphy, M. C., I. Rasnik, W. Cheng, T. M. Lohman, and T. Ha. 2004. Probing single-stranded DNA conformational flexibility using fluorescence spectroscopy. Biophysical journal 86:2530-2537.
4. McCann, J. J., U. B. Choi, L. Zheng, K. Weninger, and M. E. Bowen. 2010. Optimizing methods to recover absolute FRET efficiency from immobilized single molecules. Biophysical journal 99:961-970.
5. Iqbal, A., S. Arslan, B. Okumus, T. J. Wilson, G. Giraud, D. G. Norman, T. Ha, and D. M. Lilley. 2008. Orientation dependence in fluorescent energy transfer between Cy3 and Cy5 terminally attached to double-stranded nucleic acids. Proc Natl Acad Sci U S A 105:11176-11181.