Single Molecule Studies of Physiologically Relevant Telomeric Tails Reveal POT1 Mechanism for Promoting G-quadruplex Unfolding*[S]

Hong Wang[‡§1,2], Gerald J. Nora[¶‖1], Harshad Ghodke[§‖], and Patricia L. Opresko[¶‖3]

From the ‡Department of Pharmacology and Chemical Biology, University of Pittsburgh School of Medicine and §University of Pittsburgh Cancer Institute, Hillman Cancer Center, Pittsburgh, Pennsylvania 15213, the ¶Department of Environmental and Occupational Health, University of Pittsburgh Graduate School of Public Health, Pittsburgh, Pennsylvania 15219, and the ‖Molecular Biophysics and Structural Biology Graduate Program, University of Pittsburgh and Carnegie Mellon University, Pittsburgh, Pennsylvania 15260

Human telomeres are composed of duplex TTAGGG repeats and a 3′ single-stranded DNA tail. The telomeric DNA is protected and regulated by the shelterin proteins, including the protection of telomeres 1 (POT1) protein that binds telomeric single-stranded DNA. The single-stranded tail can fold into G-quadruplex (G4) DNA. Both POT1 and G4 DNA play important roles in regulating telomere length homeostasis. To date, most studies have focused on individual quadruplexes formed by four TTAGGG repeats. Telomeric tails in human cells have on average six times as many repeats, and no structural studies have examined POT1 binding in competition with G4 DNA folding. Using single molecule atomic force microscopy imaging, we observed that the majority of the telomeric tails of 16 repeats formed two quadruplexes even though four were possible. The result that physiological telomeric tails rarely form the maximum potential number of G4 units provides a structural basis for the coexistence of G4 and POT1 on the same DNA molecule, which is observed directly in the captured atomic force microscopy images. We further observed that POT1 is significantly more effective in disrupting quadruplex DNA on long telomeric tails than an antisense oligonucleotide, indicating a novel POT1 activity beyond simply preventing quadruplex folding.

Cells with linear chromosomes must solve the following two problems: the progressive lagging strand shortening with each cycle of DNA replication and the need to protect the ends of linear chromosomes from unwanted DNA damage responses (1). As a solution to both these problems, telomeres stand at the junction between aging, genomic stability, and cancer.

Telomeres are composed of the "shelterin complex" of proteins and TTAGGG repeats of duplex DNA along with an ssDNA overhang or "tail" of 50–500 nucleotides (1). The ssDNA tail can fold into G-quadruplex DNA (G4 DNA),[4] which consists of three tetrads of four guanines that form Hoogsteen base pairs with each other (Fig. 1A). These tetrads are in a square planar conformation and are stacked atop one another with the TTA sequences forming linker loops (2, 3). The formation of G4 DNA has been shown to inhibit the telomere-lengthening enzyme complex telomerase *in vitro* (4), although a recent *in vivo* study of *Saccharomyces cerevisiae* telomerase found that G4 DNA can promote the activity of yeast telomerase (5).

Protection of telomeres 1 (POT1) is part of the shelterin protein complex and binds to single-stranded telomeric TTAGGG repeats (6, 7). POT1 protects mammalian chromosome ends from the ataxia telangiectasia mutated and Rad3-related (ATR)-dependent DNA damage response, inhibits 5′ end resection at telomere termini, and regulates telomerase-mediated telomere extension (8). Although POT1 was shown to trap an oligonucleotide with four telomere repeats in an unfolded state to prevent G4 formation (4), the biological significance of this result is unclear. First, POT1 could not bind the short four telomere repeat substrate when the oligonucleotide was pre-folded into G4 DNA (4), and second, the telomeric tail has upwards of 30 tandem repeats in human cells (1). Thus, these studies imply that POT1 cannot actively load on telomeric tails *in vivo* unless the G4 structures are melted by a helicase, yet POT1 cellular function is not reported to depend on G4 unwinders and helicases. On the contrary, we reported that POT1 pre-loading on telomeric DNA regulates the unwinding activity of WRN helicase (9–12). At the late G_2 phase of the cell cycle, POT1 levels at the telomeres decrease, and the telomeres are temporarily unprotected and recognized as DNA damage before POT1 relocalizes to the telomeres (13). Because the unprotected tail can spontaneously fold into G4 DNA and block POT1 binding, the mechanism of POT1 reloading on the exposed telomeric tail is unknown.

Studying POT1 loading on physiological telomeric tails is complicated by a lack of information on G4 DNA formation

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Table S1, Figs. S1–S7, and additional references.

[1] Both authors contributed equally to this work.
[2] To whom correspondence may be addressed: 5117 Centre Ave., Research Pavilion, Ste. 2.6, Pittsburgh, PA 15213. Tel.: 412-623-7803; Fax: 412-623-7761; E-mail: wangh4@upmc.edu.
[3] To whom correspondence may be addressed: Bridgeside Point, 100 Technology Dr., Ste. 326, Pittsburgh, PA 15219. Tel.: 412-624-8285; Fax: 412-624-9361; E-mail: plo4@pitt.edu.

[4] The abbreviations used are: G4 DNA, G-quadruplex DNA; AFM, atomic force microscopy; ssDNA, single-stranded DNA; oligo, oligonucleotide.
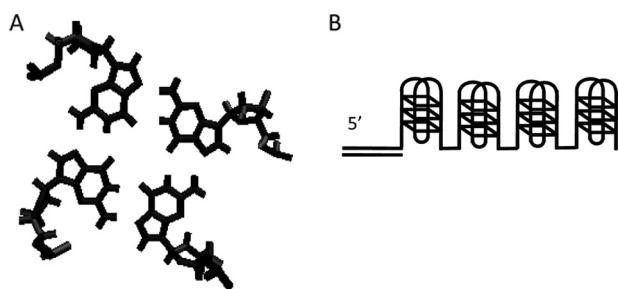
FIGURE 1. **Base pairing in G-quadruplex DNA and the beads-on-a-string model.** *A*, planar tetrad of guanines bound by Hoogsteen base pairing. Image was created on Visual Molecular Dynamics (Urbana, IL) using the Protein Bank 2JPZ structure (20). *B*, schematic illustration of the beads-on-a-string model (18, 19). In this model, long single-stranded telomeric DNA form a beads-on-a-string G4 assembly in which individual quadruplexes are connected by an ssDNA linker.

and distribution on long ssDNA strands. X-ray crystallographic and NMR studies of G4 DNA have focused on individual quadruplexes formed from four TTAGGG repeats (3, 14–18). Possible heterogeneity of the long telomeric ssDNA substrates makes them unamenable to conventional crystallographic and NMR studies (19). Furthermore, bulk biochemical assays, such as native gel electrophoresis, circular dichroism, and UV melting analysis, can only provide a mean value. Results from thermal melting assays support the hypothesis that longer telomeric ssDNAs form a beads-on-a-string G4 assembly in which individual quadruplexes are separated from each other by a TTA linker (Fig. 1*B*) (19), although some data and extrapolations from an NMR structure of individual G4 support a "stacked" arrangement of quadruplexes (20, 21). The discrepancies between these studies underscore the need to examine the formation of G4 structures on physiologically relevant telomeric tails.

Atomic force microscopy (AFM) offers a powerful single molecule approach that allows one to examine distinct nucleic acid structures (single-, double-, and triple-stranded) and their distribution within a heterogeneous population (22, 23). Previous AFM studies established the visualization of human telomeric single G4 DNA units by AFM (24). However, the quantitative distributions of various quadruplex numbers and arrangement ensembles of individual molecules within a potentially heterogeneous population of long single-stranded telomeric molecules have not been addressed. Even more importantly, POT1 coats the 3′ ssDNA tail of the telomere (6, 25, 26). However, the potential modulation of G4 folding by POT1 on physiologically relevant telomeric tails has not been investigated, and whether G4 DNA and POT1 can coexist on a telomeric tail is unknown. AFM has been used extensively to study protein-DNA interactions (27, 28), validating its application for the visualization of telomeric tail structures in the presence and absence of POT protein at the single molecule level.

First, to visualize the formation of G4 DNA on realistic telomeric tails, we performed single molecule AFM imaging of defined DNA substrates with a duplex stem followed by single-stranded TTAGGG repeats (4, 8, or 16) and conducted detailed quantitative analysis of the length and height of the G4 structures. At physiological salt concentrations, the majority of $(TTAGGG)_{16}$ molecules form only two G4 structures, instead

of the maximum of four, so that not all the POT1-binding sites are occluded. Consistent with this, the AFM images revealed that POT1 coexists with G4 DNA on some 3′ tails. We report that POT1 addition shifts the population distribution toward telomeric molecules that have fewer G4 units or that are completely unfolded. Importantly, POT1 was significantly more effective in disrupting G4 DNA on $(TTAGGG)_{16}$ molecules than an antisense oligonucleotide, indicating an activity beyond simply preventing G4 folding as proposed previously (4). Our data are consistent with a model in which POT1 acts as a "steric driver" on long telomeric ssDNA to promote unfolding of neighboring G4 structures.

MATERIALS AND METHODS

DNA Substrates—All oligonucleotides were purchased from Integrated DNA Technologies and were purified using PAGE by the manufacturer. The sequences of the oligos are listed in supplemental Table S1. DNA substrates that contain a 5′ duplex region and a 3′ ssDNA tail were formed by incubating equal molar amounts of oligonucleotides in 1× phosphate buffer (10 mM potassium phosphate and 150 mM KCl) or 1× POT1 buffer (40 mM Hepes, pH 7.5, and 50 mM KCl) at 85 °C for 5 min, followed by slow cooling to room temperature. Linear dsDNA substrate, PCR517, used as an internal size standard was made by PCR amplification of nucleotides 1374–1890 on pUC18 plasmid and purification using Illustra GFX™ PCR DNA and a gel band purification kit (GE Healthcare).

Protein Purification—Recombinant GST-tagged and untagged POT1 proteins were purified using a baculovirus/insect cell expression system and an AKTA Explorer FPLC (GE Healthcare) as described previously (10). Protein concentrations were determined using Coomassie staining along with a standard of known concentration. Proteins used in this study are more than 90% pure based on SDS-PAGE and Coomassie staining (supplemental Fig. S2*F*).

AFM Sample Preparation and Imaging—All DNA substrates and POT1 protein were diluted in 1× POT1 buffer containing additional 10 mM MgCl₂ for AFM imaging. All buffers were heated at 65 °C for 15–30 min to dissolve small salt particles that may have accumulated during storage. Samples of DNA with and without POT were prepared using the same buffer. POT1 and DNA were incubated at 37 °C for 10 min before deposition onto mica. The G-wire solution was prepared by incubating a 270 μM solution of $G_4T_2G_4$ monomer in 100 mM potassium phosphate buffer, pH 7, at 90 °C for 10 min and slow cooling to room temperature, followed by incubation at 4 °C for 12 h. For experiments using the antisense oligo, C-oligo (supplemental Table S1) was incubated with Tel16 DNA (prepared by annealing Tel16 top and bottom oligos, supplemental Table S1) at 37 °C for 10 min. All samples for AFM imaging were prepared by depositing samples onto a freshly cleaved mica (SPI Supply, West Chester, PA), followed by washing with Milli-Q water and drying under a stream of nitrogen gas. All images were collected using a MultiMode V microscope (Veeco Instruments, Plainview, NY) using E scanners in tapping mode. Pointprobe® plus noncontact/tapping mode silicon probes (PPP-NCL, Agilent) with spring constants of ~50 newtons/m and resonance frequencies of ~190 kHz were used. Images

were captured at a scan size of $1 \times 1\ \mu\text{m}$, a scan rate of 2–3 Hz, a target amplitude of 0.30 to 0.35 V, and a resolution of 512×512 pixels.

Combinatoric Model for G4 Formation—Statistical analyses of G4 formation on Tel8, Tel13, Tel14, Tel15, and Tel16 were calculated by treating them as a sequence of 8 and 13–16 lattices, respectively. It was assumed that G4 structures can form by four consecutive TTAGGG repeats and that individual G4s can fold randomly along the entire length of the lattice. The number of possible arrangements of the *h* items (G4s and unstructured repeats) can be described as shown in Equations 1 and 2,

$$^{h}C_{i} = \frac{h!}{i!(h-i)!} \qquad \text{(Eq. 1)}$$

where *i* is the number of G quadruplexes. For example, for Tel8,

$$^{5}C_{1} = \frac{5!}{1!(5-1)!} = 5 \qquad \text{(Eq. 2)}$$

there are five ways to arrange a single quadruplex, and only one way to arrange two quadruplexes.

Statistical Analysis of AFM Images—The length measurement was done using the Nanoscope7.30 software; unless stated otherwise, structures over 1 nm were noted as G4 DNA on Tel4, Tel8, and Tel16. G4 length was measured along the longest axis at the cutoff height. On Tel16, ~92% of the G4 structures form straight lines, whereas 8% of the molecules display a curvature of less than 30°. For the latter molecules, two intersecting lines were drawn following the center line of the contours. Consequently, the alignment of multimers of G4 on Tel16 does not significantly affect the measurement of G4 length. Two discernable G4 peaks on Tel16 were defined as the presence of two local maxima over 1 nm with a trough in-between that was at least 0.2 nm lower than the shorter peak. When using PCR517 fragments as internal standards for the height and length measurements, at least 20 measurements were done of peak height or full-width at half-maximum height on 517PCR. The adjusted peak height or G4 length was calculated as $F = D \times R$, where *F* is the adjusted value for height or G4 length; *D* is the value from direct measurement, and *R* is the ratio of the mean value measured from multiple depositions of PCR517 alone ($n = 20$) using different imaging probes to the mean value of the PCR517 internal standards ($n = 20$). The mean values of height and full-width at half-maximum height for PCR517 are 0.44 and 10 nm, respectively. For AFM volume analysis, the dimensions of proteins were measured using Image SXM software (28–30). The AFM volume of a particle was calculated as $V = S \times (H - B)$, where *V* is the AFM volume; *S* is the area generated at the base of a protein using "density slice" function of the SXM software; *H* is the average height, and *B* is the background height. Two-tailed Student's *t* test was conducted for statistical analysis of the height measurement.

RESULTS

Physiological Telomere Tails Rarely Form the Maximum Number of Quadruplexes—Prior to studying POT1 modulation of G4 DNA on physiological telomere tails, we set out to eluci-

date G4 DNA structures on these molecules in the absence of POT1. Previous AFM studies of G4 DNA used either short telomeric sequence (four repeats), 3′ tails of unknown lengths, or did not provide quantitative or distribution analysis of the images (31–33). Consequently, detailed information regarding the distribution and types of conformations of physiological telomeric tails was lacking. We designed a series of defined DNA substrates that have a 34-bp duplex stem at the 5′ end followed by a 3′ ssDNA overhang of 4, 8, or 16 TTAGGG repeats (Tel4, Tel8, and Tel16, respectively, supplemental Table S1). Tel4, Tel8, and Tel16 can potentially form a maximum of 1, 2, and 4 G4 units, respectively. We reasoned that comparison of G4 structures formed on these substrates as visualized through AFM imaging would provide quantitative information regarding the number of G4 units present on each molecule. AFM field view image and surface plots of Tel4, Tel8, and Tel16 show that all three telomeric substrates formed structures with heights between 1 and 2 nm (Figs. 2, *A–C*, and 3 and supplemental Fig. S1), which were not observed in images of duplex DNA or an ssDNA substrate that lacks G4-forming sequences (supplemental Fig. S2*A* and Fig. 2*E*, respectively). The heights of the peaks observed for the Tel4, Tel8, and Tel16 substrates are consistent with the height measurements from previous AFM studies of single G4 units (31). Evaluation of the AFM height at different target amplitudes indicated that within the range of target amplitudes used in this study (0.30 to 0.35 *V*), the height variation in our AFM images is ~15% of the total height (supplemental Fig. S3*A*). Because the height difference between G4 (1.32 ± 0.22 nm) and duplex DNA (0.44 ± 0.11 nm) exceeds the possible variation in height measurement, we used 1 nm as the height cutoff to measure the length of DNA with G4 character (Fig. 3). A previous AFM study reported a very similar average and standard deviation of G4 peak height on nontelomeric G4-forming sequences (1.30 ± 0.07) (34).

The number of G4 units formed on Tel4, Tel8, and Tel16 molecules was delineated by comparing lengths of G4 regions. To standardize the length measurement, we measured the full-width at half-maximum height of the PCR fragments (517 bp) deposited along with the telomeric DNA substrates (supplemental Fig. S2). The standardized G4 lengths of Tel4, Tel8, and Tel16 (see under "Materials and Methods") are shown in Fig. 2*D*, and yielded similar patterns as the nonstandardized lengths (supplemental Fig. S1*D*). The mean standardized lengths of G4 DNA at 1-nm height of Tel4 and Tel8 are 10 nm. The mean length of DNA with G4 character on Tel16 (20 nm) is only about twice that of Tel4, even though Tel16 could theoretically form a maximum of four quadruplexes as compared with Tel4 which can only form one G4. Further analysis of G4 DNA at higher salt and DNA concentrations and incubation times of up to 2 days did not yield an increase in G4 DNA formation, as judged by the AFM G4 DNA length and volume of Tel16 (data not shown). Together, our data indicate that the majority of molecules with 8 or 16 telomeric repeats only fold into one and two G4 units, respectively, which is 50% of the expected number.

To investigate the mechanism underlying the underfolding (*i.e.* formation of less than the maximum number of quadruplexes) for Tel8 and Tel16, we constructed a first-principles
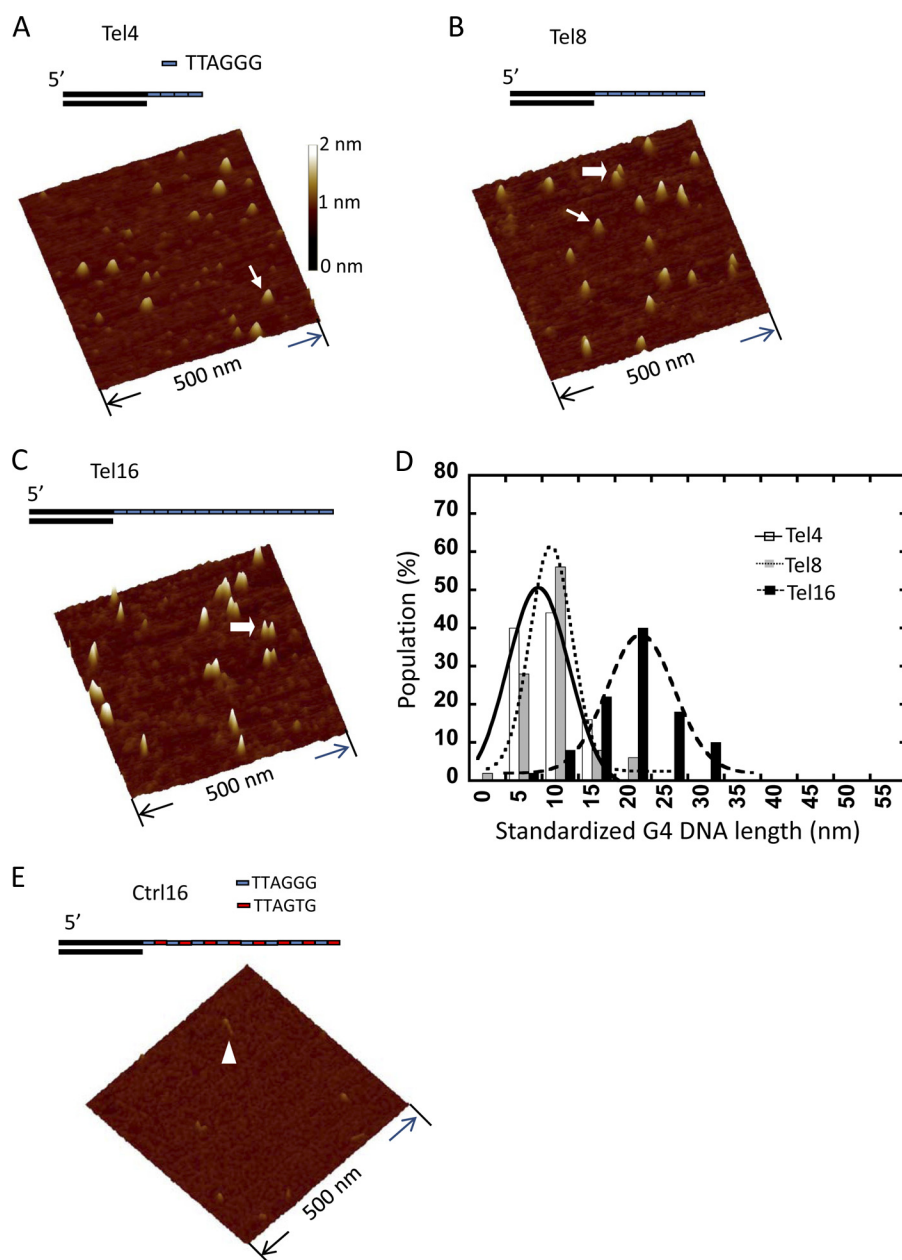
FIGURE 2. **Quantification of the number of G4 structures formed on Tel4, Tel8, and Tel16.** *A–C,* schematic drawings and representative AFM surface plots of Tel4 (*A*), Tel8 (*B*), and Tel16 (*C*) DNA substrates. *Thin arrows* point to single G4 structures, and *wide arrows* denote two distinct G4 structures on individual Tel8 or Tel16 molecules. See supplemental Table S1 for sequences. All DNA substrates were incubated in a buffer containing 150 mM KCl and deposited at 500 nM concentration (see under "Materials and Methods"). Minor particles in *A* are likely contaminants in the Tel4 preparation (*i.e.* acrylamide from the gel purification) rather than unfolded molecules because these images differ from unfolded Ctrl16 structures. *D, histogram* of G4 length (cross-section at 1-nm height) standardized using the mean full-width at half-maximum height of PCR fragments from AFM images of Tel4 (*open bars*, *n* = 50 molecules), Tel8 (*gray bars*, *n* = 50 molecules), and Tel16 (*black bars*, *n* = 50 molecules). The *black lines* represent the Gaussian fit to the data ($R^2 > 0.93$), which are centered at 10 nm (Tel4 and Tel8) and 20 nm (Tel16), respectively. *E,* representative AFM surface plot of Ctrl16 DNA, which contains eight TTAGGGTTAGTG repeats (supplemental Table S1) and does not form G4 structures. The *triangle* points to an individual Ctrl16 molecule. All images are 500 × 500 nm, and the *color bar* corresponds to height from 0 to 2 nm (from dark to bright).

combinatoric model (see "Materials and Methods") considering each telomeric repeat as a lattice point which can either be extended or folded into G4 DNA (Fig. 4*A*). The model shows that the formation of a single G4 in Tel8 is nearly five times more probable than two G4 structures. For Tel16, the most striking insight from the combinatoric model is that formation of four G4 structures on Tel16 is a rare event, which is consistent with our experimental data. In addition, the folding of two quadruplexes was the most probable conformation, but three

quadruplexes were almost as probable as two (Fig. 4*B*). This did not fit the normalized experimental data in which the lengths of G4 regions on Tel16 were divided by the mean G4 length from the Tel4 data (Fig. 4*B*). Similarly, a previous study suggested an oligonucleotide with 13 telomeric repeats formed only two quadruplexes based on circular dichroism spectra with a G4 ligand (35). To assess whether the combinatoric model was consistent with our data, we calculated the probability distributions for DNA containing 13–15 repeats. Tel13 and Tel14 both exhib-
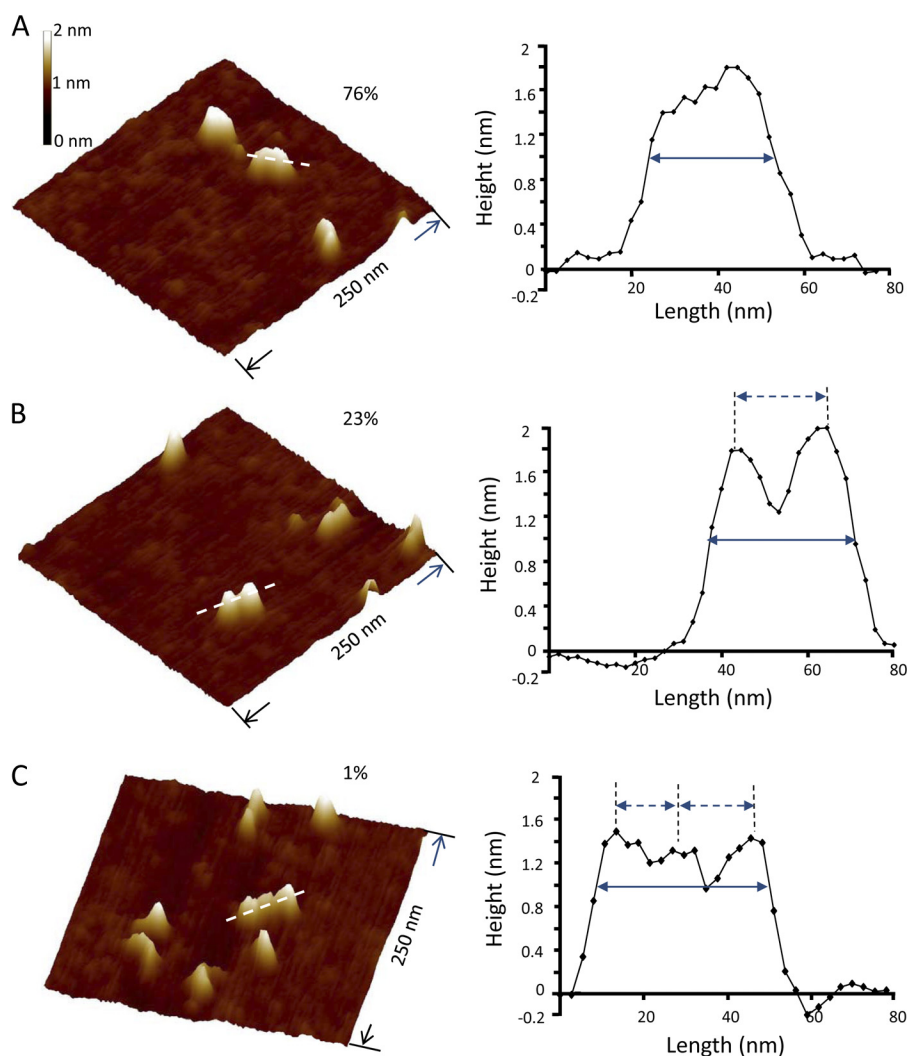
FIGURE 3. **Subpopulations of Tel16 molecules display structures that resemble beads-on-a-string.** A representative AFM image (*left panel*) and sectional analysis (*right panel*) of a Tel16 molecule in which individual G4 structures cannot be resolved (*A*), a molecule which contains two distinct peaks (*B*), or one with three distinct peaks (*C*) are shown. The *white lines* in AFM images indicate the lines drawn for section analysis. The *solid lines* with *arrows* in the section analysis indicate the length of G4 measured at 1-nm height; the *dashed lines* with *arrows* indicate the interpeak distances. The *number* at the *top right corner* of each image indicates the percent of each molecule conformation in the total population of Tel16 molecules. The AFM images are 250 × 250 nm, and the *color bar* corresponds to a height from 0 to 2 nm (from dark to bright).

ited maxima for two G4s, but for Tel15 three G4s was highly probable as well (Fig. 4*B*).

Physiologically Relevant Telomeric Tails Form Structures That Resemble Beads-on-a-String—Different models have been proposed to describe the intra-molecular assembly of multiple G4 units on long telomeric ssDNA (19, 36, 37). In a beads-on-a-string model, two G4 units are connected by one linker without stacking interactions between the units (Fig. 1*B*). In the stacking model, every G4 unit stacks onto adjacent G4, with residues on the TTA loops interacting with each other (19, 21, 38). Among all the Tel16 molecules observed, 23% displayed two distinct peaks in the AFM images (Fig. 3*B*). Although the height difference between the two distinct peaks on individual Tel16 molecules is 0.3 nm, the heights of the lower peaks are still above 1 nm at 1.3 ($\pm$ 0.3) nm. The mean interpeak distance of Tel16 molecules with two distinct peaks is 20 nm, which corresponds to ~7 TTAGGG repeats between the individual quadruplexes (supplemental Fig. S4). In the AFM images of Tel16 molecules, a small population (1%) of molecules exhib-

ited three distinct peaks (Fig. 3*C*). The assembly of multiple defined peaks resembles individual beads-on-a-string. It is worth noting that because of limitations in the AFM resolution, results from AFM imaging could underestimate the number of Tel16 molecules forming the beads-on-a-string structure (see supplemental calculations).

To further differentiate between the beads-on-a-string and the stacking models, we imaged G-wires that are long complexes of highly ordered self-assembly of inter-molecular G4 units (Fig. 5*A*). G-wires are long, uniformly quadruplectic structures with heights greater than 1 nm in AFM images (39). AFM images of G-wires formed by the short oligonucleotides $G_4T_2G_4$ are shown in Fig. 5, *B* and *C*. Because the G-wires involve stacking of the adjacent G4 units, regular well separated peaks were not apparent in the AFM images as expected, even for G-wires that were the same length as Tel16 molecules (Fig. 5, *C* and *D*). In addition, G-wires exhibited a statistically significant ($p < 0.008$) greater average height (1.63 $\pm$ 0.17 nm) compared with the Tel16 structures (1.32 $\pm$ 0.22 nm) (nonstan-
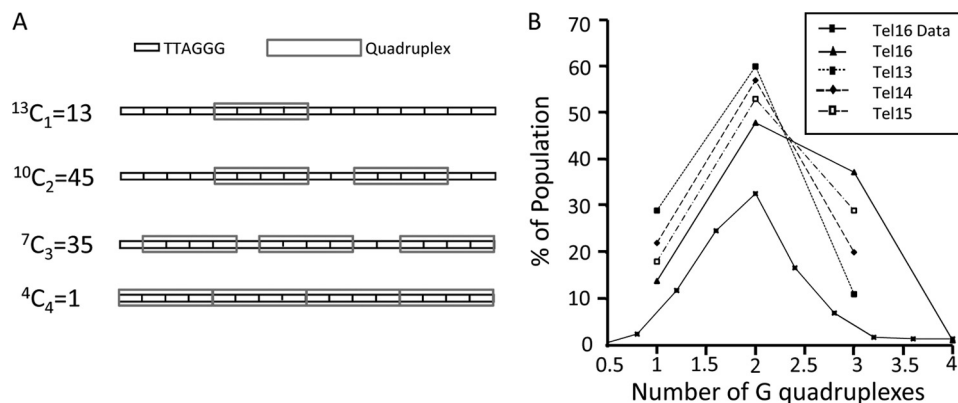
FIGURE 4. **Longer telomeric tails rarely form the maximum potential number of quadruplexes.** *A,* schematic examples and calculations of the number of possible arrangements of G4 DNA on Tel16 (see under "Materials and Methods" for equations). *B,* probability of forming 1–4 G4 structures on substrates with 13–16 (Tel13–16, respectively) based on the combinatoric calculations detailed under "Materials and Methods." The Tel16 data curve is based on the length of G4 regions on Tel16 molecules normalized using the length of single G4 measured from AFM images of Tel4 (supplemental Fig. S1*D*).
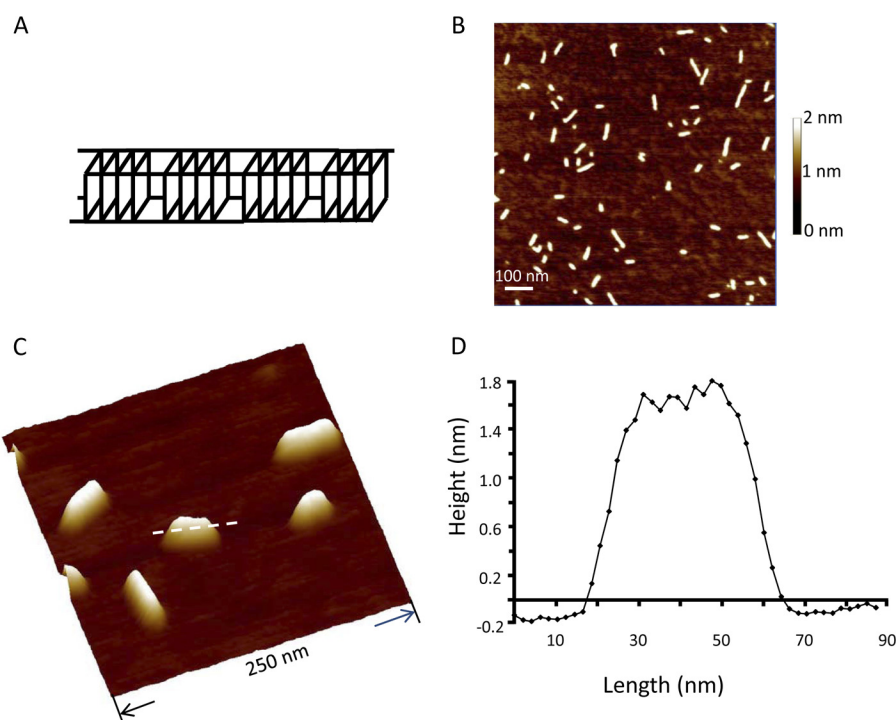

FIGURE 5. **AFM imaging of G-wires reveals a smooth surface without distinct peaks.** *A,* proposed model for G-wire formation. *B,* AFM field view image of G-wires. The image is 1 × 1 μm at 2-nm height scale. *C,* AFM surface plot of G-wires. The image is 250 × 250 nm at 2-nm height scale. The *white line* denotes the line for section analysis. *D,* section analysis of G-wire highlighted in *C.*

dardized). These results suggest that the G-wires appeared to be more rigid possibly because of the direct stacking interactions between adjacent G4 units, which lead to less compression by the mechanical AFM imaging process. The distinctly different structure of the G-wires compared with the Tel16 molecules revealed by AFM imaging suggest that G4 structures on Tel16 molecules are inconsistent with a stacked model of multiple G4 units.

Oligomeric State of POT1—A key issue in understanding the mechanism of action by POT1 is its oligomeric state. Despite evidence showing a monomeric state for the N-terminal domain of human POT1 (7), information on the oligomeric state of full-length human POT1 proteins was lacking. To evaluate the oligomeric state of full-length POT1, we measured the volume of POT1 in AFM images compared with other known

proteins of various sizes. AFM-derived volumes of proteins can be correlated to their molecular masses, permitting determination of oligomeric states (see under "Materials and Methods") and protein-protein interactions (28, 30). Purified POT1 protein after removal of the GST tag appeared as monodispersed particles in the AFM images (Fig. 6*A*). At three different concentrations (20, 200, and 1000 nM), the distribution of the calculated AFM-derived volumes of POT1 is Gaussian and centered at ~22 nm^3 (for 200 nM POT1, see Fig. 6*B*, other data not shown), which is consistent with the expected value for a POT1 monomer based on the calibration curve for globular proteins (supplemental Fig. S5). These results demonstrate that POT1 exists as a monomer in solution under the AFM imaging conditions tested. In contrast, AFM images of GST tagged POT1 protein (GST-POT1) revealed particles consistent with GST-
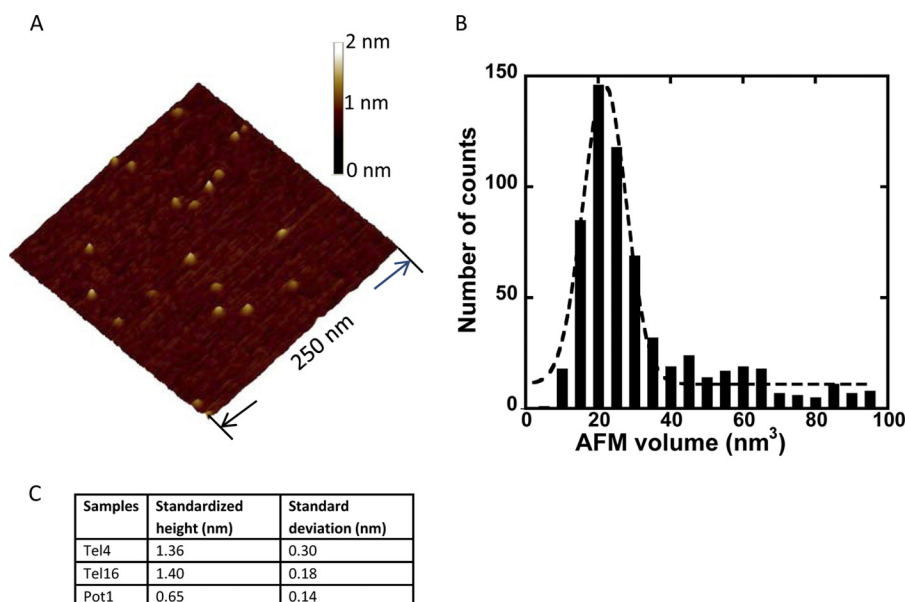
FIGURE 6. **Full-length POT1 is a monomer in solution and exhibits a height in AFM images distinct from G4 DNA.** *A,* representative AFM image of untagged POT1 protein at 200 nM concentration. The image is 250 × 250 nm at 2-nm height scale. *B,* AFM volume distribution of POT1 from images of POT1 at 200 nM concentration. The *dashed line* represents the Gaussian fit to the data ($n = 664$ molecules, $R^2 = 0.96$), which is centered at 22 nm³ and corresponds to POT1 monomer based on the standard calibration curve (supplemental Fig. S5). *C,* comparison of the standardized peak heights of Tel4, Tel16, and POT1 molecules ($n = 50$ each) in AFM images. The peak height was standardized using PCR517 DNA fragments as internal standards (see under "Materials and Methods").

POT1 dimers and tetramers (data not shown). Therefore, only untagged POT1 was used in all the imaging experiments with the DNA substrates. Importantly, the standardized height of POT1 (0.65 ± 0.14 nm) is significantly different from the standardized height for G4 DNA on Tel4 (1.36 ± 0.30 nm) and Tel16 (1.40 ± 0.18 nm) (Fig. 6C). The nonstandardized heights showed the same result (supplemental Fig. S3B). Thus, height measurement provides a robust criterion to differentiate between POT1 and G4 structure when POT1 and Tel16 are mixed together.

POT1 Binding Competes with G4 Formation on Physiologically Relevant Telomeric Tails—To study the binding of POT1 to physiological telomeric tails using AFM, we utilized two DNA substrates, Tel16 and Ctrl16 (supplemental Table S1). Ctrl16 is the same length as the Tel16 DNA substrate, but every other TTAGGG sequence in Ctrl16 is changed to TTAGTG, which eliminates G4 folding (Fig. 2E). The minimum DNA sequence that is required for high affinity binding of human POT1 *in vitro* is TTAGGGTTAG (7). Accordingly, both Tel16 and Ctrl16 substrates have a maximum of eight POT1 DNA binding sites. Electrophoresis mobility shift assays (EMSA) showed that under the same conditions POT1 binds Tel16 and Ctrl16 substrates to a similar extent (supplemental Fig. S6B). The appearance of more than one shifted band suggests that multiple POT1 molecules can bind to the Tel16 or Ctrl16 substrates.

In the AFM images of Ctrl16 with POT1, arrays of tandem POT1 proteins were observed (*thin arrow*, Fig. 7A), which were not present in the POT1-alone images (Fig. 6A). The mean height of these POT1 arrays is statistically similar to the POT1 height in the protein-alone images (supplemental Fig. S3B). We used the statistically significant height difference between POT1 and G4 DNA to differentiate between POT1 and G4 structures (Fig. 6C for standardized and S3B for nonstandar-

dized heights). When POT1 (200 nM) was incubated with a 5-fold molar excess of Tel16 (1 μM), the percent of molecules that exhibited G4 DNA structures (peak heights >1 nm) was greatly reduced from 100% of the Tel16-alone molecules, to 24% (98:405) of the molecules visualized after coincubating Tel16 with POT1 (Fig. 7C). The majority of molecules (76%, 307:405) showed only structures that were characteristic of POT1. Importantly, of the G4 DNA structures observed (98:405), 23 molecules displayed multiple peaks with differing heights that were consistent with G4 DNA and bound POT1 on the same molecule (compare Fig. 7D for POT1 + Tel16 and Fig. 7B for POT1 + Ctrl16). The height of the lower peaks is 0.7 (± 0.1) nm ($n = 23$ complexes), which is statistically different from the lower peaks on Tel16 molecules displaying two or more peaks in the absence of POT1 (1.3 ± 0.3 nm) and very closely matches the standardized peak for POT1 alone (Fig. 6C). These images indicate that G4 DNA and POT1 can coexist on the same molecule. The length distributions of POT1-bound regions for Ctrl16 and Tel16 (supplemental Fig. S6C) both exhibited a long right-sided "tail" representing similar numbers of POT1 proteins bound to Tel16 and Ctrl16 molecules. The length of longer POT1 arrays (45–60 nm) is consistent with the length of ssDNA (48 nm, assuming ssDNA as 0.5 nm/base) on fully extended Tel16 molecules. Together, these data indicate that POT1 binding can successfully compete with G4 DNA folding on telomeric ssDNA.

Previous work suggested that POT1 and an antisense 13-mer oligonucleotide, which base pairs with telomeric ssDNA, share the same mechanism of trapping a short oligonucleotide GGG(TTAGGG)₃ in an unfolded state to prevent G4 formation (4). To further investigate the mechanism of G4 disruption on long telomeric ssDNA, we quantified the G4 structures on the Tel16 substrate after incubation with the antisense oligonucleotide (C-oligo, supplemental Table S1) for comparison with
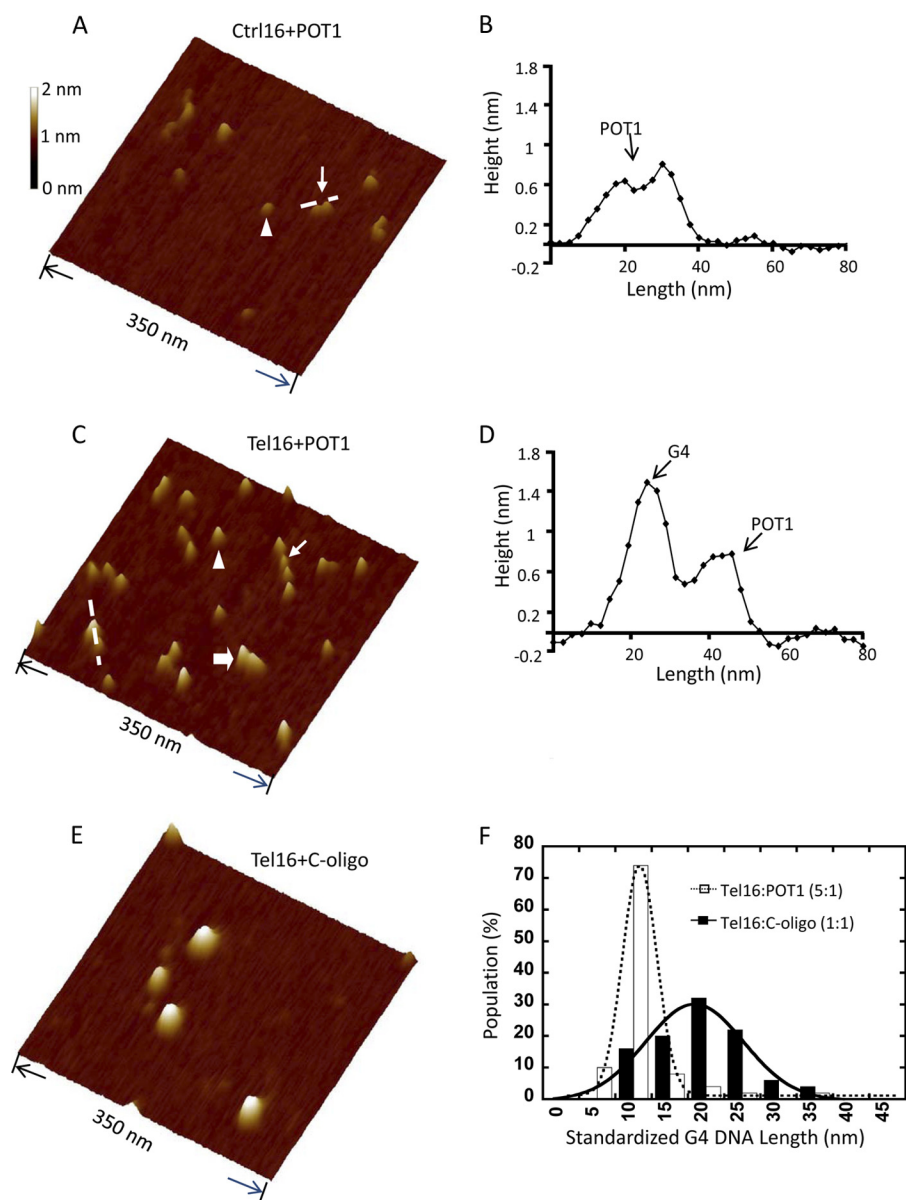
FIGURE 7. **POT1 is more effective at disrupting G4 DNA on Tel16 molecules than an antisense oligonucleotide.** *A,* representative AFM surface plot of the non-G4 forming Ctrl16 (1 μM) substrate in the presence of POT1 (200 nM). The *triangle* points to individual POT1 molecules. The *thin arrow* points to a structure with multiple POT1 proteins. *B,* cross-section of the molecule highlighted in *A* by the *dotted line* showing two POT1 molecules on the same telomeric tail. *C,* representative AFM surface plot of Tel16 (1 μM) in the presence of POT1 (200 nM). The *thick arrow* points to a structure with folded G4. The *triangle* points to an individual POT1 molecule. The *thin arrow* points to a structure with multiple POT1 proteins. *D,* cross-section of the molecule highlighted in *C* by the *dotted line* demonstrating that G4 (*left peak*) and POT1 (*right peak*) coexist on the same molecule. *E,* representative AFM surface plot of Tel16 (1 μM) in the presence of C-oligo (1 μM). *F,* histogram of the length of G4 DNA regions (stretch of DNA with peaks >1 nm) measured from AFM images of Tel16 (1 μM) in the presence of POT1 (200 nM, *open bar, n* = 50 molecules) and C-oligo (1 μM, *black bar, n* = 50 molecules). The *black lines* represent the Gaussian fit to the data (R^2 >0.96), which are centered at 10 (POT1) and 20 nm (C-oligo), respectively. The G4 length values were standardized using PCR517 DNA fragments as internal standards (see "Materials and Methods"). The AFM images are 350 $\times$ 350 nm at 2-nm height scale.

the images of POT1 added to Tel16. When Tel16 and C-oligo were incubated at a 1:1 molar ratio, most (92%) of the molecules displayed peaks at a height consistent with G4 structures (>1 nm). Thus, POT1 was more effective in decreasing the population of molecules with G4 character (24%), even though POT1 was present at lower stoichiometric amounts (5-fold less) compared with the C-oligo. An excess of C-oligo over Tel16 (5:1) is required to fully trap the G4 structures in an unfolded state (supplemental Fig. S7), which indicates that C-oligo can bind the Tel16 ssDNA. However, at this ratio the disruption of G4 structure by C-oligo is through elimination of consecutive sin-

gle-stranded TTAGGG repeats that can form G4. At a 5-fold molar excess, if the oligo is evenly distributed, the distance between individual C-oligos is ~5 nucleotides.

One caveat of our experiment is that a fraction of the Tel16 molecules that lack G4 character (76%) upon POT1 addition may represent POT1 unbound to DNA. This is unlikely because Tel16 is present at a 5-fold excess over POT1, which represents a 40-fold excess of POT1-binding sites. However, for a more rigorous analysis, we measured the length of the G4 regions on the Tel16 molecules that showed G4 peaks in the presence of POT1 (24%) or C-oligo (92%). For the C-oligo, the majority of

the G4 structure lengths were consistent with the existence of two G4 units similar to Tel16 alone (15–20 nm, standardized lengths, Fig. 7F and Fig. 2D). It is worth noting that the peaks for two G4 units on Tel16 with the C-oligo are less well defined compared with Tel16 alone, perhaps because of the oligo annealing to the region (~7 repeats or ~40 nucleotides, supplemental Fig. S4) between the G4 units. In stark contrast, the lengths of the G4 regions remaining on Tel16 after POT1 addition were about half as long as G4 regions on Tel16 with or without C-oligo (Fig. 7F and supplemental Fig. S6D). This is consistent with POT1 inducing a shift from two to one G4 unit on those Tel16 molecules that retain G4 folds. In summary, our data indicate that contrary to results with short telomeric tails (4), POT1 is much more effective at disrupting G4 DNA on long telomeric tails, compared with an antisense oligonucleotide.

DISCUSSION

POT1 binding to $(TTAGGG)_4$ substrates prevents G4 DNA folding (4, 40). However, the arrangement of G4 DNA and the competition with POT1 binding on long, physiologically realistic telomeric tails were unknown. In this study we used single molecule imaging to examine the assembly of G4 units on DNA substrates containing 4 (Tel4), 8 (Tel8), and 16 (Tel16) TTAGGG repeats, with the latter representing the mid range of the telomeric overhang length in human cells (1). Telomeric DNA with well defined lengths allowed us to study the length-dependent formation of G4 structures at the single molecule level. We demonstrated that G4 DNA assemblies on physiologically relevant telomeric tails rarely form the maximum potential number of G4 units. We observed via AFM imaging that full-length POT1 is monomeric and stabilizes the ssDNA, driving the $(TTAGGG)_{16}$ structural equilibrium toward an extended protein-bound state. This study is the first to report that bound POT1 can coexist with G4 DNA on the same Tel16 molecule. Compared with an antisense oligo that statically binds the telomeric ssDNA, POT1 is much more effective in disrupting G4 structures on long telomeric tails. Our results are consistent with a novel and more dynamic mechanism of POT1 G4 disruption, in contrast to a simple static trapping of unfolded DNA.

We applied a first principles combinatoric approach to understand the mechanism underlying the underfolding, and we found that the model prediction for ssDNA with 13 repeats (Fig. 4) is consistent with a bulk circular dichroism study that suggested oligonucleotides with 13 telomeric repeats formed on average only two quadruplexes (35). However, the normalized G4 distributions of Tel16 images demonstrated a sharp peak at two quadruplexes, whereas the probabilistic model based on the first-principles combinatoric approach predicted a nearly equal quantity of molecules with three quadruplexes as well (Fig. 4). The discrepancy between our experimental observations and the probabilistic models may be explained by differences in the probability of forming G4 at different positions along the length of Tel16 and that the model does not take into account free energy of folding. A previous study using dimethyl sulfate footprinting and exonuclease hydrolysis with $T_{24}(TTAGGG)_7$ DNA substrates revealed that the probability of forming G4 rapidly decreases toward the 5′-flanking sequence (41), from 55.8% at the 3′ end (0 position)

to 21.8, 14.5, and 7.9% at the first, second, and third positions (next to 5′-flanking sequence), respectively. Our model (Fig. 4) presumes that probabilities of forming G4 along the 3′ G-rich tail of Tel16 are the same. The dramatic decrease in the probability of forming G4 units when the repeat positions are close to the 5′-flanking region effectively shortens the number of available repeats for G4 folding on Tel16. This explains the close agreement of the normalized G4 distributions from the experimental data with the theoretical G4 distributions of two shorter substrates with 13 and 14 repeats (Fig. 4B). A previous report indicated that $GGG(TTAGGG)_3$ forms the most stable G4, and as repeat number increases ($n = 7–16$), the quadruplex molecules become less thermostable (42). The presence of loops with various lengths on the tetraplex sides can potentially lead to irregularities in G4 structure and consequently cause structure destabilization. Current literature suggests that loop length and composition strongly influence the quadruplex stability, and quadruplexes formed by $(TTAGGG)_5$ with a 9-nucleotide loop were less stable than quadruplexes formed from four consecutive repeats (41).

The arrangement of G4 DNA on longer physiological telomeric tails has been controversial. One thermal melting study supported a beads-on-a-string conformation whereby long telomeric substrates fold into the maximum number of quadruplexes that do not directly interact with each other (19). Another study found support for a stacked model whereby individual quadruplexes fold in a way that their loop reactions interact, and a more rigid superstructure is formed (21, 38). Direct visualization of individual molecules in our study revealed that 23% and 1% of the measured Tel16 molecules had two and three discernable peaks, respectively. These results support a beads-on-a-string model whereby the quadruplexes form as individual G4 units separated by stretches of ssDNA, creating a more flexible structure with discernable peaks (Figs. 2 and 3, for interpeak distance distribution see supplemental Fig. S4). Although not all the molecules displayed distinct peaks, this was likely due to the resolution limits of the AFM under the current imaging conditions. If two quadruplexes are linked by a TTA linker, the AFM cannot resolve two individual peaks; roughly 1.5 telomeric repeats are required to resolve two peaks (for the calculation of AFM resolution see supplemental material). Also, although the average nonstandardized height of the Tel16 molecules was 1.32 (± 0.22) nm, the average height of the G-wires was 1.63 (± 0.17) nm, suggesting that Tel16 G4 DNA is more flexible, corroborating a beads-on-a-string arrangement.

Previous studies indicated that POT1 binding to substrates with four repeats trapped the molecules in an extended state, shifting the equilibrium from a folded G4 unit to an extended conformation (4, 40). However, POT1 binding to physiologically relevant telomeric tails had not been examined. Our finding that the majority of Tel16 molecules only form two G4 structures has important implications for POT1 loading on realistic telomeric tails. POT1 cannot bind the short $GGG(TTAGGG)_3$ substrates until the equilibrium shifts from G4 structure to an extended state (4). In contrast, on the physiologically relevant Tel16 substrates, an underfolded Tel16 molecule constantly has multiple ssDNA sites available for POT1
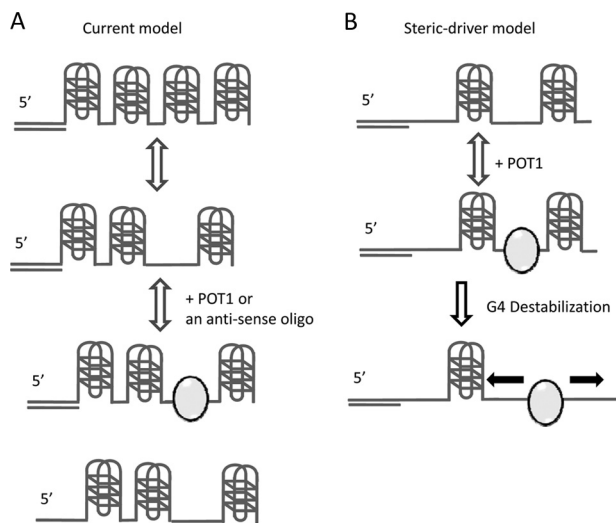
FIGURE 8. **Static passive and dynamic steric driver models of POT1 modulation of G4 DNA at telomeric tails.** A 3′ telomeric tail with 16 TTAGGG repeats is shown as an example. POT1 is shown as *gray ovals*. *A,* passive model whereby G4 DNA is arranged as a beads-on-a-string, maximally folded telomeric tail. POT1 or an antisense oligo could not bind the telomeric sequence until the G4 thermally melted to an extended state. Then POT1 or the antisense oligo traps the molecule in the unfolded state without altering adjacent G4 folds. *B,* in a steric driver model, at equilibrium the telomeric tails rarely form the maximum number of G4 units. POT1 is able to bind unfolded telomeric repeats and destabilize existing G4 DNA on the same molecule possibly through two nonmutually exclusive mechanisms (represented by the *black arrows*) as follows: dynamic movements on DNA including one-dimensional sliding, hopping, and jumping and/or its ability to destabilize adjacent G4 structures.

binding (Fig. 3), and thus, POT1 loading does not require thermal melting of existing G4 DNA.

We propose that POT1 promotion of G4 disruption on long telomeric DNA is not simply by trapping thermally melted G4 structures, as described for short substrates (4). This is because POT1 is more effective in disrupting G4 DNA than a 13-mer antisense oligonucleotide on long telomere tails (Fig. 7*F*) but not on short tails (4). At equal concentrations of antisense oligo and Tel16, the length of the majority of the G4 structures is consistent with two G4 units (Fig. 7*F*). This suggests that similar to the proposed passive model (4), the 13-mer antisense oligo can bind to the unfolded ssDNA on Tel16, but it cannot significantly influence the adjacent remaining G4 folds. On the contrary, for POT1 at a much lower protein to Tel16 ratio (1:5), the majority of molecules were unfolded, and the distribution of G4 length was shifted to one G4 unit. Our results clearly demonstrate that POT1 can disrupt G4 structures more efficiently than the antisense oligo (Fig. 7*F*).

We propose that POT1 binds to the unfolded ssDNA regions and sterically impairs adjacent telomeric repeats from folding into G4 DNA, thereby promoting unfolding into extended ssDNA (Fig. 8*B*). This is in contrast to the previous passive model based on experiments using short oligos, in which POT1 and the antisense oligo share the same ability to trap the short telomeric DNA in an unfolded form (Fig. 8*A*). We propose a steric driver model for the mechanism of G4 disruption by POT1 at 3′ telomeric tails based on the following two nonmutually exclusive mechanisms. First, POT1 binding can destabilize adjacent G4 structures. Recently, it was demonstrated using an isothermal differential hybridization method that binding of

a 46-kDa antidigoxin antibody fragment adjacent to a G4 fold dramatically destabilized the G4 structure (43). Another possible mechanism of G4 disruption by POT1 is through dynamic one-dimensional sliding and/or microscopic dissociation and re-association to adjacent sequences. Precedent for one-dimensional diffusion of single-stranded DNA-binding proteins has been described for *Escherichia coli* single-stranded DNA-binding protein based on the single molecule studies (44). The steric driver model is consistent with results from AFM imaging of Ctrl16 and Tel16 with POT1 (Fig. 7 and supplemental Fig. S6). Specifically, upon addition of POT1, the equilibrium shifts from a majority of Tel16 molecules forming two quadruplexes to one quadruplex and/or multiple POT1 monomers bound (Fig. 7*F* and supplemental Fig. S6). Importantly, multiple POT1 molecules bind Tel16 and the non-G4-forming Ctrl16 substrate to similar extents, leading to protein arrays of roughly equal length distributions (supplemental Fig. S6*C*). If POT1 can only capture the ssDNA when the G4 DNA thermally melts, then we would expect a greater number of molecules with long POT1-bound arrays for Ctrl16 relative to Tel16, because POT1 does not need to compete with G4 folding to bind Ctrl16.

In summary, we propose a model whereby POT1 acts not as an active DNA unwinder but rather as a steric driver by binding to underfolded telomeric tails and thereby destabilizing the adjacent remaining G4 structures on the molecule (Fig. 8*B*), as evidenced by the reduction of G4 DNA structures upon POT1 addition (Fig. 7*F* and supplemental Fig. S6*C*). Our results demonstrate that on a long telomeric substrate, the mechanism of action by POT1 is different from the simple static trapping mechanism utilized by an antisense oligo. POT1 binding competition with G4 DNA folding on physiologically relevant 3′ telomeric tails suggests an important mechanism for preserving telomere stability. Because a telomeric tail that is exposed during replication of the telomere can spontaneously fold into G4 DNA, this raises the issue of how POT1 reloads on the telomeric tail to promote telomerase activity or telomere remodeling into a capped structure (13). Another study demonstrated that a G4-stabilizing agent induced an ATR-dependent DNA damage response but that POT1 levels at the telomere ends remained unchanged (45), implying that G4 DNA and POT1 may coexist at telomere ends. The AFM images in this study show that the underfolding (*i.e.* less than the maximum number of G4 units) of long telomeric ssDNA provides a route for POT1 binding and a mechanism for POT1 and G4 DNA coexistence on the same molecule. The direct visualization of single molecules that resemble physiologically relevant telomeric tails provide a mechanistic basis for understanding the modulation of telomere structure and function by POT1 and G4 DNA.

REFERENCES

1. Palm, W., and de Lange, T. (2008) *Annu. Rev. Genet.* **42,** 301–334
2. Li, J., Correia, J. J., Wang, L., Trent, J. O., and Chaires, J. B. (2005) *Nucleic Acids Res.* **33,** 4649–4659
3. Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K., and Neidle, S. (2006) *Nucleic Acids Res.* **34,** 5402–5415
4. Zaug, A. J., Podell, E. R., and Cech, T. R. (2005) *Proc. Natl. Acad. Sci. U.S.A.* **102,** 10864–10869
5. Zhang, M. L., Tong, X. J., Fu, X. H., Zhou, B. O., Wang, J., Liao, X. H., Li, Q. J., Shen, N., Ding, J., and Zhou, J. Q. (2010) *Nat. Struct. Mol. Biol.* **17,** 202–209
6. Baumann, P., and Cech, T. R. (2001) *Science* **292,** 1171–1175
7. Lei, M., Podell, E. R., and Cech, T. R. (2004) *Nat. Struct. Mol. Biol.* **11,** 1223–1229
8. Palm, W., Hockemeyer, D., Kibe, T., and de Lange, T. (2009) *Mol. Cell. Biol.* **29,** 471–482
9. Opresko, P. L., Mason, P. A., Podell, E. R., Lei, M., Hickson, I. D., Cech, T. R., and Bohr, V. A. (2005) *J. Biol. Chem.* **280,** 32069–32080
10. Sowd, G., Lei, M., and Opresko, P. L. (2008) *Nucleic Acids Res.* **36,** 4242–4256
11. Sowd, G., Wang, H., Pretto, D., Chazin, W. J., and Opresko, P. L. (2009) *J. Biol. Chem.* **284,** 34682–34691
12. Opresko, P. L., Sowd, G., and Wang, H. (2009) *PLoS One* **4,** e4825
13. Verdun, R. E., Crabbe, L., Haggblom, C., and Karlseder, J. (2005) *Mol. Cell* **20,** 551–561
14. Smith, F. W., and Feigon, J. (1992) *Nature* **356,** 164–168
15. Wang, Y., and Patel, D. J. (1993) *Structure* **1,** 263–282
16. Balagurumoorthy, P., and Brahmachari, S. K. (1994) *J. Biol. Chem.* **269,** 21858–21869
17. Parkinson, G. N., Lee, M. P., and Neidle, S. (2002) *Nature* **417,** 876–880
18. Maizels, N. (2006) *Nat. Struct. Mol. Biol.* **13,** 1055–1059
19. Yu, H. Q., Miyoshi, D., and Sugimoto, N. (2006) *J. Am. Chem. Soc.* **128,** 15461–15468
20. Dai, J., Carver, M., Punchihewa, C., Jones, R. A., and Yang, D. (2007) *Nucleic Acids Res.* **35,** 4927–4940
21. Petraccone, L., Trent, J. O., and Chaires, J. B. (2008) *J. Am. Chem. Soc.* **130,** 16530–16532
22. Hansma, H. G., Revenko, I., Kim, K., and Laney, D. E. (1996) *Nucleic Acids Res.* **24,** 713–720
23. Lyubchenko, Y. L., and Shlyakhtenko, L. S. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94,** 496–501
24. Kan, Z. Y., Lin, Y., Wang, F., Zhuang, X. Y., Zhao, Y., Pang, D. W., Hao, Y. H., and Tan, Z. (2007) *Nucleic Acids Res.* **35,** 3646–3653
25. Hockemeyer, D., Sfeir, A. J., Shay, J. W., Wright, W. E., and de Lange, T. (2005) *EMBO J.* **24,** 2667–2678
26. Price, C. M. (2006) *Nat. Struct. Mol. Biol.* **13,** 673–674
27. Janiæijeviæ, A., Ristic, D., and Wyman, C. (2003) *J. Microsc.* **212,** 264–272
28. Wang, H., Yang, Y., and Erie, D. A. (2007) in *Protein Interactions, Biophysical Approaches for the Study of Complex Reversible Systems* (Schuck, P., ed) pp. 39–78, Springer Science+Business Media, LLC, Boston
29. Ratcliff, G. C., and Erie, D. A. (2001) *J. Am. Chem. Soc.* **123,** 5632–5635
30. Yang, Y., Wang, H., and Erie, D. A. (2003) *Methods* **29,** 175–187
31. Zhang, X. Y., Cao, E. H., Zhang, Y., Chou, C., and Bai, C. (2003) *J. Biomol. Struct. Dyn.* **20,** 693–702
32. Yoshimura, S. H., Maruyama, H., Ishikawa, F., Ohki, R., and Takeyasu, K. (2004) *Genes Cells* **9,** 205–218
33. Xu, Y., Ishizuka, T., Kurabayashi, K., and Komiyama, M. (2009) *Angew. Chem. Int. Ed. Engl.* **48,** 7833–7836
34. Neaves, K. J., Huppert, J. L., Henderson, R. M., and Edwardson, J. M. (2009) *Nucleic Acids Res.* **37,** 6269–6275
35. Chang, C. C., Chien, C. W., Lin, Y. H., Kang, C. C., and Chang, T. C. (2007) *Nucleic Acids Res.* **35,** 2846–2860
36. Ambrus, A., Chen, D., Dai, J., Bialis, T., Jones, R. A., and Yang, D. (2006) *Nucleic Acids Res.* **34,** 2723–2735
37. Renciuk, D., Kejnovská, I., Skoláková, P., Bednárová, K., Motlová, J., and Vorlícková, M. (2009) *Nucleic Acids Res.* **37,** 6625–6634
38. Petraccone, L., Garbett, N. C., Chaires, J. B., and Trent, J. O. (2010) *Biopolymers* **93,** 533–548
39. Marsh, T. C., Vesenka, J., and Henderson, E. (1995) *Nucleic Acids Res.* **23,** 696–700
40. Torigoe, H., and Furukawa, A. (2007) *J. Biochem.* **141,** 57–68
41. Tang, J., Kan, Z. Y., Yao, Y., Wang, Q., Hao, Y. H., and Tan, Z. (2008) *Nucleic Acids Res.* **36,** 1200–1208
42. Vorlícková, M., Chládková, J., Kejnovská, I., Fialová, M., and Kypr, J. (2005) *Nucleic Acids Res.* **33,** 5851–5860
43. Wang, Q., Ma, L., Hao, Y. H., and Tan, Z. (2010) *Anal. Chem.* **82,** 29469–29475
44. Roy, R., Kozlov, A. G., Lohman, T. M., and Ha, T. (2009) *Nature* **461,** 1092–1097
45. Rizzo, A., Salvati, E., Porru, M., D'Angelo, C., Stevens, M. F., D'Incalci, M., Leonetti, C., Gilson, E., Zupi, G., and Biroccio, A. (2009) *Nucleic Acids Res.* **37,** 5353–5364