PNAS COMMENTARY

Dynamic folding and accessibility of telomeric overhang

Ahmet Yildiz[a,b,c,1]

The evolution of linear chromosomes in eukaryotes poses several challenges for the natural ends of chromosomes, such as misrecognition as DNA damage repair (DDR) sites, DNA degradation, and incomplete DNA replication. A solution to these challenges has been the formation of a protective cap structure comprising a tandem array of telomeric repeats and associated proteins. Human telomeres consist of 2 to 20 kb of double-stranded TTAGGG repeats. Telomeres are not blunt-ended but instead terminate with a 100- to 150-base-long 3′ single-stranded TTAGGG overhang. These overhangs serve two key functions: They provide primers for telomere extension by telomerase and recruit sequence-specific telomeric proteins to cap chromosome ends and to ensure genome stability. Deletion of the overhangs leads to chromosome fusions. The mechanism of telomere end protection has been attributed primarily to the sequestration of the 3′ overhang within large duplex loops (t-loops) (1). Another potentially significant factor that could influence the protection of the overhang is that these G-rich repeats spontaneously fold into G-quadruplex (GQ) structures (2). In vitro studies have shown that GQ formation of single-stranded telomeric DNA reduces access of several proteins that target these sites (3, 4), raising the possibility that GQ formation suppresses unwanted DDR or degradation of these overhangs. However, how these long overhangs fold into a complex array of GQ structures and how GQ formation affects the accessibility of single-stranded telomeric sites are not well understood. Most of the previous in vitro work focused on DNA substrates long enough to form a single GQ. However, telomeric overhangs can form an array of more than 10 GQ structures, and these GQs can dynamically fold, unfold, and alter their conformation. In PNAS, Shiekh et al. used the single-molecule FRET-PAINT (Förster resonance energy transfer-point accumulation for imaging in nanoscale topography) method to investigate the accessibility of long, single-stranded telomeric DNA (5).

In GQs, four guanines pair by Hoogsteen hydrogen bonding to form a planar tetrad. GQ folding requires at least four tandem GGGXXX (X refers to any base) repeats and stabilization of the central cavity by monovalent cations. Although the ability of guanine-rich oligonucleotides to fold into tetrameric structures has been known for decades, studies in GQ formation had been mostly limited to in vitro, and their physiological significance remained elusive. This view started to change when GQ formation was shown to down-regulate the transcription of several oncogenes (6). The development of antibodies that target DNA GQ structures also enabled the visualization of GQ structures in the genome of mammalian cells (7). We now know that GQs are enriched at promoters of multiple genes and play a role in the regulation of gene expression (8). GQs are also enriched at telomeres, but their precise role in telomere biology remains to be demonstrated. The intermolecular GQs that form at the double-stranded telomeric DNA may serve as a major obstacle to the replication of telomeric DNA, and resolving these structures requires helicase activity (9). The intramolecular GQs that form at the overhang may reduce the accessibility of single-stranded telomere binding proteins or contribute to the protection of telomere ends against DDR signals.

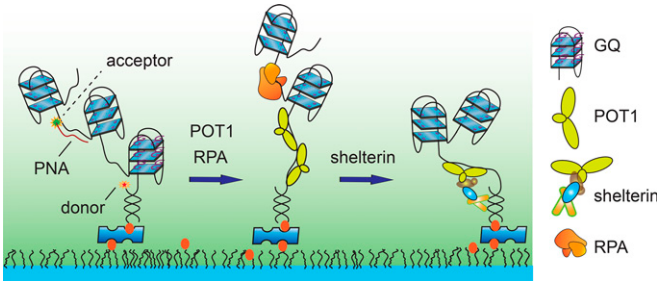

Fig. 1. Single-molecule observation of GQ folding and accessibility of telomeric overhangs. A model telomeric substrate with a long overhang is immobilized to a glass surface. (*Left*) The overhang spontaneously folds into multiple GQs and the available sites between adjacent GQs are targeted with short complementary PNA. The FRET signal between PNA probes and the duplex DNA reveals the spatial distribution of accessible sites. (*Middle*) The accessible sites can be targeted with single-stranded telomeric binding proteins, such as POT1 and RPA. (*Right*) Tethering of POT1 to the duplex DNA via the rest of the shelterin complex may enable POT1 to antagonize the binding of PNA probes or proteins that target the overhang.

Recent in vitro studies started to shed light on how telomeric overhangs fold into GQ structures. A long telomeric overhang can fold into multiple heterogenous GQ conformations flanked with unfolded regions (10). Due to the scarcity of direct approaches, it remains unclear whether neighboring GQs interact with each other and whether the sequences between folded GQs become accessible to telomere-binding proteins. Shiekh et al. (5) studied the accessibility of telomeric overhangs by using FRET-PAINT, which combines single-molecule FRET with DNA-PAINT. They immobilized a model telomeric DNA that terminates with 4 to 28 repeats of a TTAGGG overhang (Fig. 1, *Left*). The DNA is labeled with a FRET donor at the junction between single-stranded and double-stranded (ss/ds) telomeric DNA. DNA-PAINT uses an acceptor-labeled peptide nucleic acid (PNA) probe complementary to the G-tracts at the overhang. These PNA probes are seven nucleotides in

Author affiliations: [a]Biophysics Graduate Group, University of California, Berkeley, CA 94720; [b]Physics Department, University of California, Berkeley, CA 94720; and [c]Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720

Author contributions: A.Y. wrote the paper.

The author declares no competing interest.

See companion article, "Emerging accessibility patterns in long telomeric overhangs," 10.1073/pnas.2202317119.

[1]Email: yildiz@berkeley.edu.

Published September 7, 2022.

length, such that they only transiently hybridize to unfolded telomeric tracts. PNA binding produces a FRET signal correlated with the distance between the ss/ds junction and the unfolded overhang site. At sufficiently low PNA concentrations, the frequency, residence time, and FRET signal of individual binding and dissociation events can be detected in real time. The FRET signal does not directly report the spatial distribution of unfolded G-tracts in tandem telomeric sequences because GQ folding patterns and potential stacking interactions between the neighboring GQs affect the FRET signal. Therefore, the method indirectly reports the distribution of PNA probes that bind near the ss/ds junction or elsewhere on the overhang.

Shiekh et al. (5) found that the distributions of FRET values varied considerably depending on the number of telomeric repeats present in the overhang. Broader FRET distribution was observed when the overhang length is a multitude of four repeats (4n, n = 1, 2, 3…), suggesting that the PNA probes anneal to unfolded telomeric repeats throughout the sequence. In comparison, uniformly high FRET values were observed for 4n+1, 4n+2, and 4n+3 long overhangs, which have at least one to three unfolded repeats at a time. In these substrates, PNA prefers to bind to the first couple of G-tracts near the ss/ds junction, indicating that single-stranded TTAGGG repeats next to the ss/ds junction are the most accessible sites for telomere-targeting proteins.

The FRET-PAINT approach developed in this study can be used to investigate how diverse and dynamic folding of telomeric GQs is modulated by proteins that target the overhang. In cells, telomeres are safeguarded by a six-protein complex, named shelterin. Shelterin contains both single-stranded (POT1) and double-stranded (TRF1 and TRF2) telomeric DNA-binding proteins that are connected by TPP1 and TIN2 (11). Previous single-molecule FRET studies showed that POT1 binds and unfolds completely folded telomeric GQs (4). The POT1–TPP1 complex was also shown to slide along the telomeric DNA and induce dynamic folding and unfolding of GQ structures (12). However, these studies were performed using short DNA substrates to keep the end-to-end distance of DNA within the sensitivity range (2 to 8 nm) of the FRET signal. FRET-PAINT would extend these studies to physiologically relevant overhang lengths.

One potential future direction is to investigate whether POT1 increases the overall accessibility of single-stranded telomere binding proteins by unfolding the GQ folds or POT1 binding to these sites protects them against ssDNA binding proteins, as shown for Replication Protein A (RPA) (4). RPA is 1,000 times more abundant in cells than POT1 and has a similar affinity for the telomeric overhang (13–15). This leads to the question of how cells efficiently protect the overhang against RPA binding (Fig. 1, *Middle*). Previous in vitro studies showed that POT1/TPP1 alone is not sufficient to protect telomeres against excess RPA in solution (4). Protection of the telomeric overhang may require tethering of POT1/TPP1 to the rest of shelterin via TIN2 (Fig. 1, *Right*). Consistent with this view, TIN2 knockdown leads to the accumulation of RPA at telomeres in cells (13). Alternatively, TIN2 and TPP1 may increase the telomere binding affinity of POT1 (14), effectively reducing RPA's ability to displace it from telomeric DNA. In addition, TRF2 prefers to bind to the ss/ds junction of telomeric substrates (16), which can alter the folding kinetics of G-tracts adjacent to the junction. The FRET-PAINT approach can be used to monitor the competition between shelterin and RPA to bind telomeric overhangsand determine how this competition is affected by the dynamics nd heterogeneity of GQ folding landscapes.

The 3′-end of the telomeric overhang is the primary site for telomere elongation by both telomerase and alternative lengthening of telomere (ALT) mechanisms. In vitro studies have shown that GQ formation reduces the de novo synthesis of new telomeric tracts by telomerase (17). POT1/TPP1 has been proposed to act as both positive (14, 17) and negative (18) regulators of telomerase, and their role in telomere length regulation appears to be complex. The TPP1 OB-fold was shown to be the telomerase-binding component of shelterin (19), highlighting this domain as a key component that facilitates telomerase recruitment to telomere ends, which are otherwise tightly protected by the shelterin complex (20). A physical TPP1–telomerase interaction on the primer that is being extended has been proposed to statically tether telomerase and prevent its premature release from its substrate (20). Alternatively, POT1–TPP1 may stimulate telomerase recruitment to the 3′-end by sliding along the overhang (12). These models are not mutually exclusive but are both based on indirect evidence. New experimental approaches, such as the one developed by Shiekh et al. (5), would be required to directly visualize how GQ folding/unfolding of the overhang together with binding and assembly of shelterin components regulate the accessibility and enzymatic activity of telomerase and other proteins that target the 3′-end of telomeres.

ACKNOWLEDGMENTS. A.Y.'s research is supported by the NIH (GM136414) and the NSF (MCB-1055017, MCB-1617028).

1. J. D. Griffith *et al.*, Mammalian telomeres end in a large duplex loop. *Cell* **97**, 503–514 (1999).
2. E. Henderson, C. C. Hardin, S. K. Walk, I. Tinoco Jr., E. H. Blackburn, Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine-guanine base pairs. *Cell* **51**, 899–908 (1987).
3. A. J. Zaug, E. R. Podell, T. R. Cech, Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 10864–10869 (2005).
4. S. Ray, J. N. Bandaria, M. H. Qureshi, A. Yildiz, H. Balci, G-quadruplex formation in telomeres enhances POT1/TPP1 protection against RPA binding. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2990–2995 (2014).
5. S. Shiekh *et al.*, Emerging accessibility patterns in long telomeric overhangs. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2202317119 (2022).
6. A. Siddiqui-Jain, C. L. Grand, D. J. Bearss, L. H. Hurley, Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11593–11598 (2002).
7. C. Schaffitzel *et al.*, In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with *Stylonychia lemnae* macronuclei. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 8572–8577 (2001).
8. J. Robinson, F. Raguseo, S. P. Nuccio, D. Liano, M. Di Antonio, DNA G-quadruplex structures: More than simple roadblocks to transcription? *Nucleic Acids Res.* **49**, 8419–8431 (2021).
9. K. Paeschke *et al.*, Pif1 family helicases suppress genome instability at G-quadruplex motifs. *Nature* **497**, 458–462 (2013).
10. H. Wang, G. J. Nora, H. Ghodke, P. L. Opresko, Single molecule studies of physiologically relevant telomeric tails reveal POT1 mechanism for promoting G-quadruplex unfolding. *J. Biol. Chem.* **286**, 7479–7489 (2011).

11. W. Palm, T. de Lange, How shelterin protects mammalian telomeres. *Annu. Rev. Genet.* **42**, 301–334 (2008).
12. H. Hwang, P. Opresko, S. Myong, Single-molecule real-time detection of telomerase extension activity. *Sci. Rep.* **4**, 6391 (2014).
13. K. K. Takai, T. Kibe, J. R. Donigian, D. Frescas, T. de Lange, Telomere protection by TPP1/POT1 requires tethering to TIN2. *Mol. Cell* **44**, 647–659 (2011).
14. F. Wang *et al.*, The POT1-TPP1 telomere complex is a telomerase processivity factor. *Nature* **445**, 506–510 (2007).
15. Q. Yang, Y.-L. Zheng, C. C. Harris, POT1 and TRF2 cooperate to maintain telomeric integrity. *Mol. Cell. Biol.* **25**, 1070–1080 (2005).
16. R. M. Stansel, T. de Lange, J. D. Griffith, T-loop assembly in vitro involves binding of TRF2 near the 3′ telomeric overhang. *EMBO J.* **20**, 5532–5540 (2001).
17. A. J. Zaug, E. R. Podell, T. R. Cech, Mutation in TERT separates processivity from anchor-site function. *Nat. Struct. Mol. Biol.* **15**, 870–872 (2008).
18. D. Loayza, T. De Lange, POT1 as a terminal transducer of TRF1 telomere length control. *Nature* **423**, 1013–1018 (2003).
19. H. Xin *et al.*, TPP1 is a homologue of ciliate TEBP-beta and interacts with POT1 to recruit telomerase. *Nature* **445**, 559–562 (2007).
20. J. Nandakumar, T. R. Cech, Finding the end: Recruitment of telomerase to telomeres. *Nat. Rev. Mol. Cell Biol.* **14**, 69–82 (2013).