Multiple POT1–TPP1 Proteins Coat and Compact Long Telomeric Single-Stranded DNA

Derek J. Taylor  
Case Western Reserve University, derek.taylor@case.edu

Elaine R. Podell  
University of Colorado Boulder

Dylan J. Taatjes  
University of Colorado Boulder

Thomas R. Cech  
University of Colorado Boulder, thomas.cech@colorado.edu

Follow this and additional works at: https://scholar.colorado.edu/chem_facpapers

Part of the Molecular Biology Commons

Recommended Citation  
https://scholar.colorado.edu/chem_facpapers/24

This Article is brought to you for free and open access by Chemistry & Biochemistry at CU Scholar. It has been accepted for inclusion in Chemistry & Biochemistry Faculty Contributions by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.
Multiple POT1–TPP1 Proteins Coat and Compact Long Telomeric Single-Stranded DNA

Derek J. Taylor¹,², Elaine R. Podell²,³, Dylan J. Taatjes² and Thomas R. Cech²,³

¹Department of Pharmacology, Case Western Reserve University, Cleveland, OH 44106, USA
²Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO 80309-0215, USA
³Howard Hughes Medical Institute

Telomeres are nucleoprotein complexes that cap and protect the ends of linear chromosomes. In humans, telomeres end in 50–300 nt of G-rich single-stranded DNA (ssDNA) overhangs. Protection of telomeres 1 (POT1) binds with nanomolar affinity to the ssDNA overhangs and forms a dimer with another telomere-end binding protein called TPP1. Whereas most previous studies utilized telomeric oligonucleotides comprising single POT1–TPP1 binding sites, here, we examined 72- to 144-nt tracts of telomeric DNA containing 6–12 POT1–TPP1 binding sites. Using electrophoretic mobility gel shift assays, size-exclusion chromatography, and electron microscopy, we analyzed telomeric nucleoprotein complexes containing POT1 alone, POT1–TPP1, and a truncated version of POT1 (POT1-N) that maintains its DNA-binding domain. The results revealed that POT1-N and POT1–TPP1 can completely coat long telomeric ssDNA substrates. Furthermore, we show that ssDNA coated with human POT1–TPP1 heterodimers forms compact, potentially ordered structures.

© 2011 Elsevier Ltd. All rights reserved.

Telomeres are nucleoprotein complexes that comprise the ends of eukaryotic chromosomes. Telomeres contribute to genomic stability, in part, by preventing deleterious events such as chromosome end-to-end fusions and degradation. In addition, telomeres are intimately involved in recruiting and regulating enzymatic complexes necessary for telomeric DNA modification, replication, and repair.¹ Both the protective capabilities of telomeres and their regulation of cellular processes including senescence, DNA damage response, subnuclear localization, and gene expression involve specific proteins that bind to the telomeric DNA.²⁻⁵

In humans, telomeric DNA has a repeating, hexameric sequence of TTAGGG that extends for several thousand base pairs.⁶ Telomeric DNA ends in 3′ single-stranded overhangs, which are about 50–300 bases in mammals.⁷⁻⁹ Both the double- and single-stranded regions of the telomeric DNA are bound by various telomeric proteins. Telomere-repeat binding factors (TRFs) 1 and 2 are sequence-specific proteins that bind the double-stranded telomeric DNA.¹⁰ The single-stranded portion of the telomere DNA is recognized and bound with nanomolar affinity by protection of telomeres 1 (POT1), also in a sequence-specific manner.¹¹⁻¹³ Other proteins interact with TRF1, TRF2, and POT1 to form a six-membered core complex called shelterin that contributes to chromosome-end

Abbreviations used: ssDNA, single-stranded DNA; TRF, telomere-repeat binding factor; EMSA, electrophoretic mobility shift assay; EM, electron microscopy; POT1, protection of telomeres 1.
protection and telomere homeostasis. Sub-complexes have been identified that consist of only three to five of the shelterin components. The multimerization state may be important for conducting interactions with enzymes functioning at the telomere. For example, the binding of one shelterin protein (TIN2) has been shown to protect another member of the shelterin complex, TRF1, from selective ubiquitination.

It is proposed that telomere length in the yeast *Saccharomyces cerevisiae* is regulated by the number of Rap1p proteins coating the double-stranded, telomeric DNA. In such a protein-counting mechanism, a greater number of Rap1p molecules stabilize the formation of a compact structure that prevents access of telomerase, the ribonucleoprotein enzyme that synthesizes telomeric DNA. As telomere length increases, more Rap1p molecules coat the telomeric DNA, and telomerase activity is inhibited. A similar counting mechanism has been proposed in mammals, where the number of TRF1 and TRF2 proteins regulates the length of the double-stranded region of telomeres. Multiple splicing variants of human POT1 have been identified in vivo. The primary products of the splicing variants are the full-length protein and a.

---

**Fig. 1.** EMSAs depicting the ability of POT1-N to coat long tracts of telomeric DNA. ssDNA oligos were synthesized by Integrated DNA Technologies. Oligos ranged from 24 nt to 48 nt and contain four to eight telomere hexameric repeats. Every two hexameric repeats (5'-GGTTAG GGTTAG-3') is suitable for binding individual POT1 molecules. DNA was 5'-radiolabeled, and gel shifts were performed essentially as previously described. For each reaction, a 10-μL solution was prepared containing 400 nM ssDNA spiked with approximately 4% 5'-32P-labeled DNA. POT1-N was added to each reaction at molar ratios indicated on top of the gel. The molar ratio indicated is calculated for the ratio of protein added per dodecameric repeat (5'-GGTTAG GGTTAG-3') in the oligonucleotide in each lane. Reactions were conducted in buffer containing 25 mM Tris (pH 8.0), 150 mM NaCl, 10 mM DTT, and 5% glycerol. After 30 min on ice, reactions were analyzed by gel electrophoresis on 6% acrylamide gels in 0.5× Tris–borate–ethylenediaminetetraacetic acid buffer. Gels were dried and imaged on a Typhoon phosphorimager (GE Life Sciences).
truncated isoform, representing only the N-terminal half of the protein. Both isoforms bind single-stranded, telomeric DNA. The crystal structure of the truncated version of POT1, referred to here as POT1-N, has been solved at atomic resolution in complex with a telomeric decanucleotide. This structure reveals the basis for the sequence specificity of POT1 for telomeric DNA. The C-terminal portion of full-length POT1 is important for coordinating protein–protein interactions with TPP1 and other members of the shelterin complex. Structural information regarding TPP1 is currently limited to the oligosaccharide–oligonucleotide binding fold domain (residues 90–250 of 544) of the protein. While TPP1 by itself does not interact with ssDNA, the POT1–TPP1 heterodimer exhibits an approximate 10-fold increase in affinity compared to POT1 alone. Our studies use a truncated version of TPP1 (89–334) that maintains both its POT1- and telomerase-interacting domains and therefore retains its protective and recruitment properties, respectively. We refer to the truncated TPP1 simply as TPP1 throughout this study.

The minimal DNA bound with high affinity by a single human POT1 protein in vitro is 9 or 10 nt, with the sequence 5′-TTAGGGTTAG-3′. POT1 binds short oligonucleotides (<2 telomere hexameric repeats, 12 nt) with higher affinity when the telomeric sequence resides at the 3′ terminus of the oligonucleotide substrate. The inclusion of TPP1 increases the affinity for the dodecamer telomere sequence, even in the presence of a 3′ non-telomeric extension up to 4 nt. Whether multiple human POT1 binding events are cooperative and what TPP1 contributes to POT1 binding to long strands of telomeric DNA have not been reported.

Previous studies of POT1 binding focused on DNA substrates containing a small number of telomeric repeats, but within human cells, the single-stranded telomeric DNA is 50–300 nt. We therefore characterized telomere protein assembly on long ssDNAs containing multiple POT1 binding sites. Electrophoretic mobility shift assay (EMSA) data indicated that these longer oligonucleotides could be successively bound by POT1-N until the binding sites were saturated (Fig. 1). The EMSSAs suggest that, at suitable ratios of protein–DNA binding sites, POT1-N molecules bind consecutively to every two hexameric DNA repeats. Full-length

![Fig. 2. EMSAs of POT1-N (a), POT1 (b), and POT1–TPP1 (c and d) binding to a 72-mer telomeric DNA substrate. Gel shifts were performed essentially as described in Fig. 1. Briefly, 200 nM DNA solution (containing ~4% 5′-[32P]-labeled DNA) was incubated with increasing concentrations of protein for 15 min on ice in 10 μL buffer containing 25 mM Hepes (pH 8.0), 150 mM NaCl, 5 mM DTT, and 5% glycerol. The protein concentration in each reaction ranged from 0 to 6 μM in 1.2-μM increments. POT1 and POT1-N were expressed as glutathione S-transferase fusion proteins in baculovirus-infected insect cells as previously described. TPP1(89–334) was expressed with an N-terminal 6×His tag. POT1–TPP1 heterodimers were purified after co-expression in insect cells using both affinity tags and size-exclusion chromatography. After incubation, the protein–DNA complexes were subject to electrophoresis in a non-denaturing polyacrylamide gel and dried. In (a) to (d), line markers on the left side of gel indicate one, two, three, four, five, or six proteins bound to the DNA; arrow indicates position of wells. (e) The relative abundance of each complex was quantified using a Typhoon Imager (GE Life Sciences) and the ImageQuant TL software (GE Life Sciences). For quantification, the number of counts representing ssDNA forming the specific complex with six proteins was divided by the total number of counts in each lane (including the aggregates) and plotted as a percentage.](image-url)
POT1 demonstrated similar binding properties on long tracts of ssDNA (Fig. 2). The observation of sequential loading of POT1 to give intermediate complexes also indicates little, if any, cooperativity in multiple binding events.

Additional gel shift experiments were used to monitor the binding of POT1-N (Fig. 2a) and full-length POT1 (Fig. 2b) to a 72-nt telomeric DNA substrate. When a molar ratio of 0.5 POT1-N molecules per POT1 binding site on the ssDNA was used (Fig. 2a, lane 2), a clear ladder was seen, indicating one to six POT1-N molecules loaded onto the ssDNA substrate. At increasing concentrations, POT1-N largely saturated the 72-nt substrate with the maximum six POT1-N proteins per ssDNA molecule. Full-length POT1 exhibits binding affinities for a 10- to 12-nt ssDNA substrate with a single POT1 binding site that are similar to that of POT1-N. At molar ratios of POT1 to DNA binding site up to 2:1, we observed efficient formation of a specific complex corresponding to six POT1 proteins per 72-mer. However, at higher protein concentrations, the POT1-ssDNA complexes began to aggregate and were no longer able to enter the non-denaturing gels (Fig. 2b, lanes 6–11).

We reasoned that TPP1, which binds to the C-terminal domain of POT1, might facilitate proper folding and solubility of telomeric DNA–protein complexes. Indeed, the presence of TPP1 (Fig. 2c) prevented aggregation of the complex so that it now entered the non-denaturing gel, even at 5× molar excess of the POT1–TPP1 heterodimer over POT1 binding sites on the ssDNA substrate. This phenomenon was also observed when TPP1 was added to the reaction mixture after 15 min of POT1–ssDNA incubation (Fig. 2d). At stoichiometric concentrations of POT1–TPP1 to POT1 binding sites on the ssDNA substrate, approximately 85% of the ssDNA molecules were fully saturated with six POT1–TPP1 heterodimers bound (Fig. 2c, lane 3; Fig. 2e). Surprisingly, even at higher molar excess of POT1–TPP1, the ssDNA was never completely saturated by the heterodimer. This observation was in contrast to POT1-N alone, which was able to saturate the DNA substrate entirely (Fig. 2a). We speculate that the discrepancy between POT1-N and POT1–TPP1 saturations may be due to a higher affinity for DNA by the heterodimer, which would limit its ability to rearrange and optimize its interactions with the DNA substrate (see below).

It has been reported that the binding constant of POT1 with a 12-nt telomeric substrate increases 9-fold with the addition of TPP1 (\(K_d = 63\) nM for POT1 versus 7 nM for POT1–TPP1). We found a similar situation with the 72-mer telomeric substrate, with the POT1–TPP1 heterodimer binding at lower protein concentrations than full-length POT1 or POT1-N alone (Fig. 2e). When POT1 and TPP1 were added sequentially, with a 15-min delay, more protein was needed to reach maximum binding (Fig. 2e). We attribute this observation to the fact that full-length POT1 appears to aggregate and thus inhibits POT1–ssDNA binding interactions. While this adverse effect is fully reversible upon the addition of TPP1, more TPP1 is required presumably to disrupt aberrant POT1–POT1 aggregations and facilitate proper complex formation. These data support the conclusion that TPP1 assists in proper folding of the C-terminal domain of POT1, thus preventing aggregation and enhancing the ability of POT1 to coat long single-stranded telomeric tracts.
The increased affinity of POT1–TPP1 for DNA may explain why it only saturates about 85% of the DNA molecules, whereas POT1-N is capable of saturating 100% of the POT1 binding sites. POT1-N might be able to dissociate and rebind or may slide along the substrate to insure that consecutive POT1 molecules bind every 12 nt on the DNA. Conversely, the tighter-binding POT1–TPP1 heterodimer may not slide along the DNA and thus would be unable to accommodate additional POT1–TPP1 dimers at increasing stoichiometries. Unexpectedly, a small percentage of 72-mer telomeric DNA never bound more than one or two POT1–TPP1 dimers (see lower bands, Fig. 2c). This observation indicates that as few as one POT1–TPP1 heterodimer is capable of interacting with a long piece of DNA and affects subsequent binding events, possibly by looping the DNA into unique structures that reduce the accessibility of the remaining POT1 binding sites.

We next characterized the POT1–TPP1–ssDNA complexes using size-exclusion chromatography. To assemble nucleoprotein complexes, we incubated a 2-fold molar excess of highly purified, recombinant POT1–TPP1 with ssDNA containing 72, 120, 132, or 144 nt. Each assembled telomeric complex eluted from the gel-filtration column as a sharp peak, indicative of a homogenous population (sample data for the 132-mer shown in Fig. 3a). In contrast, the free ssDNA eluted as a broad peak (Fig. 3b). In the chromatogram of the POT1–TPP1 assembled complexes, a smaller peak representing excess POT1–TPP1 heterodimer follows the peak representing the complex. The OD\textsubscript{260}/OD\textsubscript{280} ratios of the two peaks are consistent with the larger faster-eluting complex containing DNA. The slower-eluting peak represents protein only and therefore possesses a lower OD\textsubscript{260}/OD\textsubscript{280} ratio. Whereas a small percentage of complexes (<15% total) appeared to be less than saturated in the EMSA analysis, there was no indication of smaller complexes in the size-exclusion chromatograms. This is likely due to the relative abundance of any one subcomplex being too low for detection in our experiments.

Multiple complexes were assembled using different lengths of DNA and saturated with either POT1-N or POT1–TPP1 heterodimer. Complexes assembled with 72-nt, 132-nt, and 144-nt telomeric DNA substrates and proposed to contain 6, 11, and 12 telomere proteins, respectively, were analyzed using gel filtration. The peak representing the telomeric complex eluted at different volumes in the chromatogram (Supplemental Fig. 1). The total mass of complexes ranged from \(\sim\)250 kDa for a complex consisting of a 72-mer DNA coated with six POT1-N proteins up to \(\sim\)1.2 MDa for a complex composed of 144-nt telomeric DNA coated with 12 POT1–TPP1 dimers. Calibration of the gel-filtration column using molecular weight standards revealed that the telomeric complexes eluted at volumes corresponding to the predicted molecular masses.

Next we used electron microscopy (EM) to visualize assembled telomeric complexes consisting of 12 POT1–TPP1 heterodimers bound to a 144-mer

**Fig. 4.** Negative-stain EM of telomeric complexes. (a) A typical micrograph of telomeric complexes assembled from 144-nt telomeric DNA saturated with 12 POT1–TPP1 heterodimers. The particles are globular in shape and monodisperse on the grid. The scale bar represents 100 nm. (b) Two-dimensional class averages of the 144-nt POT1–TPP1 telomeric complexes. The purified complexes were brought to a final concentration of approximately 50 \(\mu\)g/mL in 25 mM Hepes, pH 8.0, and 150 mM NaCl. Grids were prepared using the sandwich method and stained in 0.75% uranyl formate as described previously. Micrographs were recorded on a 4K×4K Gatan CCD camera with a Tecnai F20 electron microscope operating at an accelerating voltage of 200 keV and a magnification of 39,000×. Pixel sizes, after binning, were 5.52 Å per pixel on the object scale. All image processing was performed with the SPIDER software. A total of 3664 particles were picked and classified into 57 class averages using reference-free alignment and K-means classification procedures. Class averages are windowed in boxes with 440 Å×440 Å dimensions.
POT1–TPP1 Proteins Coat Telomeric ssDNA

ssDNA substrate. Complexes were assembled, purified using gel filtration, and then stained in uranyl formate. The predominant species observed in the electron micrographs were globular monodisperse particles measuring approximately 230 Å in diameter (Fig. 4a). To improve the signal-to-noise ratio and, thus, the interpretability of the imaged complexes, we classified the particles into two-dimensional averages (Fig. 4b). All class averages presented a similar shape consisting of compact globular structures with cavities or pits throughout the structures. In each of the two-dimensional class averages, the pits are arranged in a criss-crossing pattern. The pattern visualized within the class averages is consistent with the string of telomeric DNA, coated by POT1, which likely reflect limited compositional heterogeneity (as seen with the 72-mer in Fig. 2), and other differences that could be due to orientational or conformational heterogeneity. It is also possible that the telomeric complex projections have slightly different structures due to distortions from adsorption onto the EM grid or alterations from the negative stain. For example, some of the particles may be irregularly compressed due to staining in a sandwich of two carbon films, which has been noted previously with this technique. Nonetheless, the images and two-dimensional class averages obtained by EM indicate that the telomeric complexes form compact globular structures. The direct visualization, combined with characterization of the complexes using gel filtration and gel shifts, indicates that POT1–TPP1 heterodimers coat single-stranded, telomeric DNA to form complexes that appear to have specific, ordered structures. We note that POT1-N saturating a 132-nt telomeric oligonucleotide also revealed a compact structure by EM (data not shown).

Implications for telomere maintenance

The primary role of telomeres is to cap and protect the ends of chromosomes from degradation and repair pathways. If left exposed, the telomeric ssDNA overhang at the extreme 3′ termini of chromosomes would be susceptible to DNA repair or end-to-end fusion events. POT1 shields this ssDNA from deleterious events by binding it in a deep groove within two N-terminal oligosaccharide-oligonucleotide binding folds, and TPP1 increases the affinity of this interaction. Our data demonstrate that multiple POT1–TPP1 molecules can coat a long ssDNA substrate and form compact structures. These data suggest that the individual POT1–TPP1 heterodimers may bend and direct the telomeric DNA in a way that compresses the ssDNA overhang into a compact and ordered nucleoprotein complex.

There are several prior examples of telomeric proteins influencing the overall structure of telomeric DNA. TRF2 has been implicated in the formation of large lariat loops called T-loops that are thought to contribute to protecting telomeric DNA. Later experiments revealed that TRF2 may facilitate formation of T-loops by condensing and supercoiling telomeric DNA. Human TRF1 homodimers have been shown to bind long tracts of double-stranded telomeric DNA, which may also facilitate formation of T-loops.

Aside from the primary role in chromosome protection and stability, another role of the telomere-end binding proteins is to regulate telomerase. This regulation is often inhibitory, as the deletion or truncation of many telomere-end binding proteins leads to longer telomeres. On the other hand, telomere proteins have been shown to stimulate telomerase activity by increasing enzyme processivity and by recruiting the holoenzyme to the 3′ end of the existing telomere. Genetic mutations have since localized G100 as a surface amino acid of telomerase reverse transcriptase (TERT) that is necessary for TPP1 interaction. Forming a compact structure at the 3′ ssDNA overhang may offer the advantage that multiple consecutively bound TPP1 molecules could more efficiently recruit telomerase to the 3′ DNA substrate. Alternatively, the compact structure may represent a state that shields telomere DNA by forming a protective “shell” around the ssDNA.

Conclusion

Our biochemical and structural data reveal that a long single-stranded, telomeric DNA substrate can be coated by multiple POT1–TPP1 heterodimers and can adopt a compact, ordered structure. These compact structures may be beneficial for protecting the ssDNA, for regulating enzyme recruitment, and/or for promoting telomerase processivity.

Supplementary materials related to this article can be found online at doi:10.1016/j.jmb.2011.04.049

Acknowledgements

We thank Dr. Ming Lei (present address, University of Michigan) for his contributions to the early stages of this project. We also thank Dr. Andy Hoenger and the Boulder Lab for 3D Electron Microscopy for use of the facilities. This work was supported by National Institutes of Health Grant GM28039 to T.R.C., National Institutes of Health.
References