Published online 12 September 2011 *Nucleic Acids Research, 2012, Vol. 40, No. 1* **235–244**
doi:10.1093/nar/gkr721

DNA-induced dimerization of the single-stranded DNA binding telomeric protein Pot1 from *Schizosaccharomyces pombe*

Jayakrishnan Nandakumar and Thomas R. Cech*

Department of Chemistry and Biochemistry, Howard Hughes Medical Institute, University of Colorado, Boulder, CO 80309-0215, USA

Received August 1, 2011; Revised and Accepted August 22, 2011

ABSTRACT

Eukaryotic chromosome ends are protected from illicit DNA joining by protein–DNA complexes called telomeres. In most studied organisms, telomeric DNA is composed of multiple short G-rich repeats that end in a single-stranded tail that is protected by the protein POT1. Mammalian POT1 binds two telomeric repeats as a monomer in a sequence-specific manner, and discriminates against RNA of telomeric sequence. While addressing the RNA discrimination properties of SpPot1, the POT1 homolog in *Schizosaccharomyces pombe*, we found an unanticipated ssDNA-binding mode in which two SpPot1 molecules bind an oligonucleotide containing two telomeric repeats. DNA binding seems to be achieved via binding of the most N-terminal OB domain of each monomer to each telomeric repeat. The SpPot1 dimer may have evolved to accommodate the heterogeneous spacers that occur between *S. pombe* telomeric repeats, and it also has implications for telomere architecture. We further show that the *S. pombe* telomeric protein Tpz1, like its mammalian homolog TPP1, increases the affinity of Pot1 for telomeric single-stranded DNA and enhances the discrimination of Pot1 against RNA.

INTRODUCTION

Eukaryotic chromosome ends are protected from being detected as DNA breaks by the formation of protein–DNA structures known as telomeres (1). Telomeric DNA is composed of a long array of G-rich repeats (exact repeats of GGTTAG in mammals and sequences related to GGTTAC in *S. pombe*) that originate from the sub-telomeric region of chromosomes as double-stranded (ds) DNA and terminate as a G-rich single-stranded (ss) tail (2).

The telomeric ssDNA, which is roughly 200 nt long in humans (3,4), is bound in a sequence-specific manner by an Oligonucleotide/Oligosaccharide-binding (OB) domain-containing protein called POT1, which is conserved across mammals and present in fission yeast (5–10). The POT1–DNA interaction is strengthened by the telomeric protein TPP1, which is a binding partner of mammalian POT1 (11). Yeast genetics has revealed a functional homolog of TPP1 in *S. pombe* called Tpz1 (12), although it is not known if this protein mimics the POT1–DNA binding activity of mammalian TPP1.

Each molecule of human POT1 (hPOT1) binds two consecutive repeats of ss telomeric DNA using two tandem N-terminal OB domains in an asymmetric fashion (7). The first OB domain of hPOT1 interacts with the last four nucleotides (TTAG) of the first telomeric repeat and with the first two nucleotides (GG) of the second repeat, whereas the second OB domain binds to the remaining four nucleotides (TTAG) of the second repeat. Additionally, the self-recognition patterns of DNA in each telomeric repeat and the nature of the protein–DNA interactions between each OB domain and its DNA target in the crystal structure are distinct, adding to the asymmetry in binding. *S. pombe* Pot1 (SpPot1) contains a N-terminal OB domain (OB1) that optimally binds a single GGTTAC telomeric repeat with sub-micromolar affinity and exhibits cooperative multiple binding to longer DNA (6,13). Superposition of the most N-terminal OB domains in the hPOT1-**TTA**GGGT TAG (7) and SpPot1 OB1-GG**TTA**C (6) crystal structures shows precise overlap of the **TTA** trinucleotide in the two structures, suggesting that the most N-terminal OB domains of both the POT1s are structurally and functionally conserved.

Qualitative gel-shift assays with full-length SpPot1 showed that the stability of SpPot1–DNA complexes increased significantly as additional nucleotides were appended to the 3′-end of GGTTAC, suggesting the

*To whom correspondence should be addressed. Tel: +1 303 492 8606; Fax: +1 303 492 6194; Email: thomas.cech@colorado.edu

presence of a second DNA binding module on SpPot1 that bound the newly added DNA sequence albeit with relaxed sequence specificity (14). Indeed, a biochemical construct of SpPot1, SpPot1-1–389, which encompasses OB1 and a contiguous putative second OB domain (OB2), showed a large increase in affinity for GGTTACGGTTAC compared to the affinity of OB1 for GGTTAC (10). Hence, it appeared the SpPot1-DNA complex, like the hPOT1–DNA complex, employs two OB domains to engage two ss telomeric repeats. Additionally, OB2 in isolation shows a modest affinity (~400 nM) and sequence specificity for the GGTTACGGT oligonucleotide (15).

TERRA, non-coding RNA containing multiple G-rich telomeric repeats transcribed from chromosome ends, is found in mammals, budding yeast and fission yeast, and is implicated in telomerase regulation and chromatin remodeling (16–18). How does POT1 avoid TERRA and bind faithfully to telomeric DNA? A recent structural and biochemical study on hPOT1 and mouse POT1A with ribo-substituted GGTTAGGGTTAG oligonucleotides revealed that the thymidine at position 4 defines RNA discrimination by mammalian POT1s and that RNA discrimination is greatly enhanced by binding of POT1 by TPP1 (19). Previous binding experiments using the OB1 domain of SpPot1 and singly ribo-substituted GGTTAC oligonucleotides showed that dT3/rU3 and dT4/rU4 substitutions resulted in large reductions in binding, highlighting the importance of thymidines in RNA discrimination by OB1 (6). However, the determinants of RNA discrimination have not been re-evaluated in light of the expanded GGTTACGGTTAC binding site and in the context of full-length SpPot1.

In this study, we initially addressed how SpPot1 achieves RNA discrimination. To our surprise, the RNA discriminators, dT3 and dT4, are symmetrically distributed on both repeats (GG**TT**ACGG**TT**AC) such that dT to rU substitution in neither repeat alone causes a reduction in binding, whereas simultaneous dT to rU substitutions in both repeats lead to reduced affinity. The discrimination pattern is a duplication of the pattern observed for OB1 with GG**TT**AC, leading us to speculate that two SpPot1 molecules (using two OB1 domains) associate with GGTTACGGTTAC. Using native gel-shift, filter-binding assays, gel-filtration experiments with SpPot1 and SpPot1-1–389, and OB1 tandem-fusion proteins, we show that two SpPot1 molecules bind GGTTACGGTTAC simultaneously. Finally, we show that Tpz1, like its mammalian homolog TPP1, enhances Pot1's ability to bind telomeric ssDNA and discriminate against RNA of telomeric sequence.

MATERIALS AND METHODS

Oligonucleotides

Oligonucleotides were purchased from Integrated DNA Technologies. For gel-shift and filter-binding experiments, the oligonucleotides were 5′ ^{32}P-labeled using T4 polynucleotide kinase (New England Biolabs) and [γ-^{32}]ATP, and purified using Sephadex G-25 spin columns (Roche). For gel-filtration experiments, oligonucleotides were purchased from Integrated DNA Technologies after synthesis on a 1 micromole scale and HPLC purification.

Cloning, protein expression, and purification

The SpPot1 gene and Tpz1-1–234 gene fragment were PCR-amplified from cDNA and cloned into a pET–His–Smt3 expression vector (20). The OB1 gene construct, which codes for amino acids 1–185 of SpPot1, was PCR-amplified from the pET-His-Smt3-SpPot1 plasmid and cloned into the pET-His-Smt3 expression vector. OB1–OB1 and OB1–OB1–OB1 are tandem fusions of two and three OB1 gene-constructs, respectively, constructed by sub-cloning PCR-amplified OB1 fragments (containing 5′ and 3′ HindIII restriction sites) into the HindIII site at the 3′- end of the OB1 ORF in the pET–His–Smt3–OB1 plasmid. All proteins were overexpressed in Bl21 (DE3) cells. His–Smt3–fusions were purified from the soluble cell lysates using Ni-affinity chromatography and the His–Smt3 tag was removed with the Smt3-specific protease, Ulp1 (20). Untagged proteins were further purified by size-exclusion chromatography (Superdex200, GE or Superdex75, GE). Full-length SpPot1 was further purified using anion exchange (MonoQ, GE) chromatography. The fractions containing pure proteins were flash frozen in liquid nitrogen and stored at −80°C.

Electrophoretic mobility shift assays

Binding mixtures (10 µl) containing 50 mM Tris–HCl (pH 8.5), 25 mM NaCl, 5 mM DTT, 6% glycerol and specified concentrations of 5′ ^{32}P-labeled oligonucleotides and SpPot1/Tpz1 proteins were incubated for 30 min at 4°C. The mixtures were then analyzed by electrophoresis for 1.5 h at 4°C at 200 V through a non-denaturing 4–20% gradient polyacrylamide gel (Invitrogen). The gel was dried and visualized using a phosphorimager (Typhoon Trio, GE).

Filter-binding assays for K_D determination

Filter-binding experiments were performed in a 96-well dot blot apparatus. A total of 10 pM labeled oligonucleotides were incubated with SpPot1 (active protein concentration of 0, 0.001, 0.003, 0.01, 0.03, 0.1, 0.3, 1, 3, 10, 30, 100 nM) in the presence of 0.1 µg/ml bovine serum albumin (BSA; New England Biolabs), 0.1 µg/ml yeast tRNA (Sigma) and 5 mM NaCl in a 100 µl total volume of binding buffer (50 mM Tris–HCl at pH 8.0, 5 mM DTT) for 30 min on ice. The mixtures (90 µl) were then filtered through a pre-cooled (at 4°C) membrane sandwich containing a nitrocellulose membrane (BA85, Whatman), a positively charged nylon membrane (Hybond N$^+$, GE), and a filter paper (Whatman). The filters were pre-washed with 90 µl of binding buffer before sample application and were washed with 90 µl of binding buffer after sample application. The membranes were air-dried and quantified using a Phosphorimager. The data were analyzed with MS Excel (Microsoft Office) and Kaleidograph (Synergy software). For assays involving SpPot1 + Tpz1-1–234, the binding mixtures minus Pot1 were incubated with 200 nM Tpz1-1–234 for 10 min prior to addition of SpPot1. All other assay conditions and procedures were identical to

that described for SpPot1 alone. We note here that the absolute value of the K_D of SpPot1 with ssDNA seems dependent on the conditions of the assay, insofar as gel-shift experiments with the same SpPot1 construct in the absence of tRNA, and presence of 1 mg/ml BSA and 50 mM NaCl, yield dissociation constants that are in the low picomolar range (Prof. Deborah S. Wuttke, personal communication). All of our conclusions are therefore based on relative binding affinities measured under our conditions (containing 0.1 μg/ml yeast tRNA, 0.1 μg/ml BSA and 5 mM NaCl).

Gel-filtration experiments

The size-exclusion chromatography experiments discussed in Figure 3C were performed using a Superdex 200 column [10/300 (GE); 23.5 ml bed volume] equilibrated with buffer containing 20 mM Tris–HCl (pH 8.0), 250 mM NaCl and 1 mM DTT. SpPot1 (50 μM) was incubated alone or with ssDNA (100 μM) for 15 min on ice before being applied to the column. The elution was analyzed by monitoring UV absorbance at 280 nm and 260 nm. The column was calibrated using gel-filtration protein standards (Bio-rad). Log_{10}(Molecular weight) was plotted against Ve/Vo (Ve is the elution volume at maximum A_{280} absorbance for a given sample and Vo is the void volume of the column determined to be 9.58 ml based on the elution of Dextran blue) for all the protein standards and the straight line: Log_{10}(Molecular weight) = $-1.759(Ve/Vo) + 4.7532$ was used to fit the data (Supplementary Figure S1). The elution volumes (Ve) for SpPot1 and SpPot1–DNA complexes were substituted in the above equation to determine the molecular weights of the complexes.

RESULTS

Disruption of SpPot1-ssDNA binding by ribo-substitution requires dT to rU substitutions in both telomeric repeats of a dodecameric oligonucleotide

The OB1 domain of SpPot1 defines DNA (versus RNA) specificity at positions T3 and T4 of GGTTAC (6). Given that the optimal binding substrate of SpPot1 is not a hexamer, but a dodecamer comprising two GGTTAC repeats, we wished to determine the RNA discrimination criterion for two-repeat DNA and the full-length protein. Full-length SpPot1 was expressed in *E. coli* as a Smt3-fusion and purified to homogeneity (Figure 1A; 'Materials and Methods' section). Recombinant SpPot1 purified in this manner displayed robust DNA binding activity with oligonucleotide T_{18}GGTTACGGTTAC in a gel-shift assay (Figure 1B). The 18 T nucleotides preceding the telomeric repeats provide excess negative charge allowing the DNA–protein complex to enter the gel (19).

Using a filter-binding assay, we evaluated the dissociation constants (K_D) of SpPot1 with substrates containing all deoxyribonucleotides, all ribonucleotides, tri-ribo-substitutions, or hexa-ribosubstitutions in the context of the GGTTACGGTTAC sequence (Figure 2A and B). The all-RNA oligonucleotide showed a severe binding defect

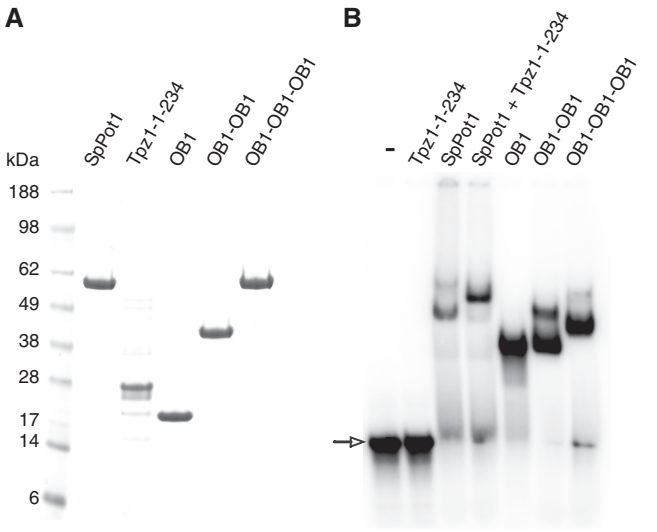


Figure 1. Purity and telomeric ssDNA-binding activity of SpPot1 and Tpz1 constructs. (**A**) A 5 μl of protein standard (SeeBlue Plus2, Invitrogen) and 4 μg each of indicated proteins were analyzed by 4–20% acrylamide SDS–PAGE and visualized using coomassie blue staining. (**B**) EMSA of the proteins (50 nM) shown in (A) with a radiolabeled T_{18}GGTTACGGTTAC oligonucleotide (25 nM). The thymidines were added to the 5′- end of the telomeric sequence to facilitate migration of the protein–DNA complexes into the gel. The arrow indicates the position of the free DNA on the gel and '–' indicates that no protein was added.

(K_D ~324 nM; Table 1) compared to the all-DNA oligonucleotide (5.0 ± 1.3 nM; Table 1). Surprisingly, all the oligonucleotides where ribo-substitutions were limited to one of the two hexameric repeats, showed binding affinities for SpPot1 comparable to that of the all-DNA substrate (Figure 2A and Table 1).

To ask whether ribo-substitution of both repeats is necessary to observe a binding defect, we tested SpPot1 binding to oligonucleotide substrates that had tri-ribonucleotide substitutions in both repeats (Figure 2B). Interestingly, all four tested ribo-substituted substrates showed a significant SpPot1 binding defect (12- to 30-fold; K_D's are displayed in Table 1). All the hexaribonucleotide-containing substrates shown in Figure 2B have in common one dT to rU substitution in each telomeric repeat. We wondered whether the RNA discrimination profile of full-length protein on a dodecameric repeats is equivalent to a duplication of the RNA discrimination profile of OB1 on GGTTAC. Fully consistent with this idea, the dT3/dT9 to rU3/rU9 or the dT4/dT10 to rU4/rU10 substitutions elicited 7- and 20-fold reductions in binding to SpPot1, whereas all the other tested di-ribo-substituted oligonucleotides showed DNA-like SpPot1-binding (Figure 2C and Table 1).

Schizosaccharomyces pombe telomeres are different from mammalian telomeres, in that they often contain a 1–3 nt spacer separating consecutive GGTTAC repeats (14). However, the dissociation constants with SpPot1 of the oligonucleotides containing deoxyribonucleotide spacers found commonly in *S. pombe* telomeres (A and AGG) were unaffected by ribo-substitutions in the

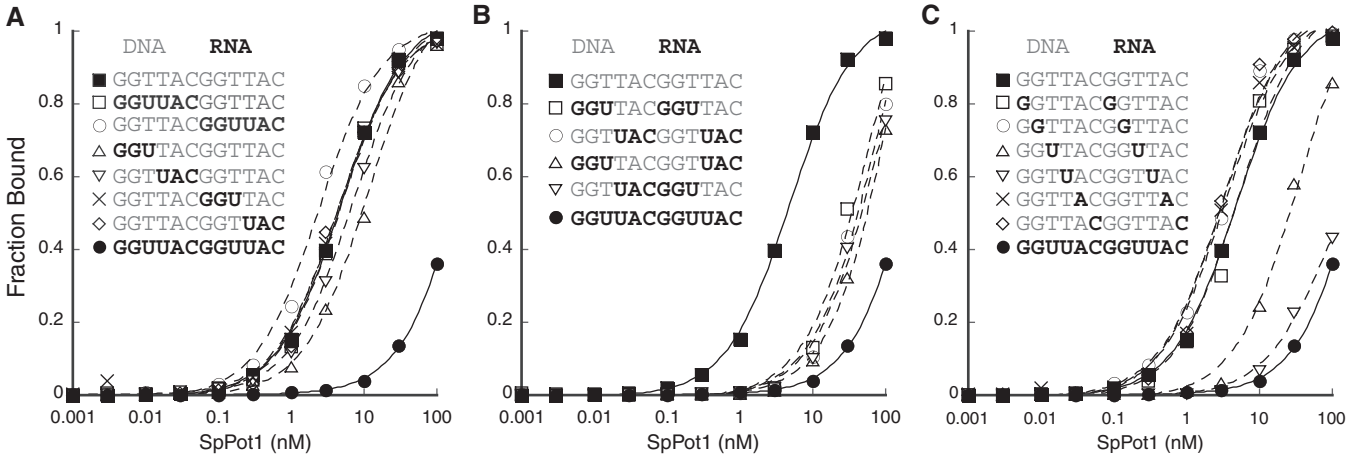


Figure 2. SpPot1–DNA dodecanucleotide interaction is affected by ribo-substitution only when thymidines in both GGTTAC repeats are replaced by uridines. Filter-binding profiles of SpPot1 with ^{32}P-labeled ssDNA–RNA of indicated sequence (ribonucleotides are depicted in black and deoxyribonucleotides in gray) were done in duplicate, and the mean of the fraction of SpPot1-bound ssDNA–RNA was plotted against SpPot1 concentration. The mean K_D's and standard errors of the duplicates are detailed in Table 1. Ribonucleotide substitutions are present in either one of the two GGTTAC repeats (**A**), or in both repeats (**B** and **C**).

Table 1. Dissociation constants for SpPot1–ssDNA/RNA complexes

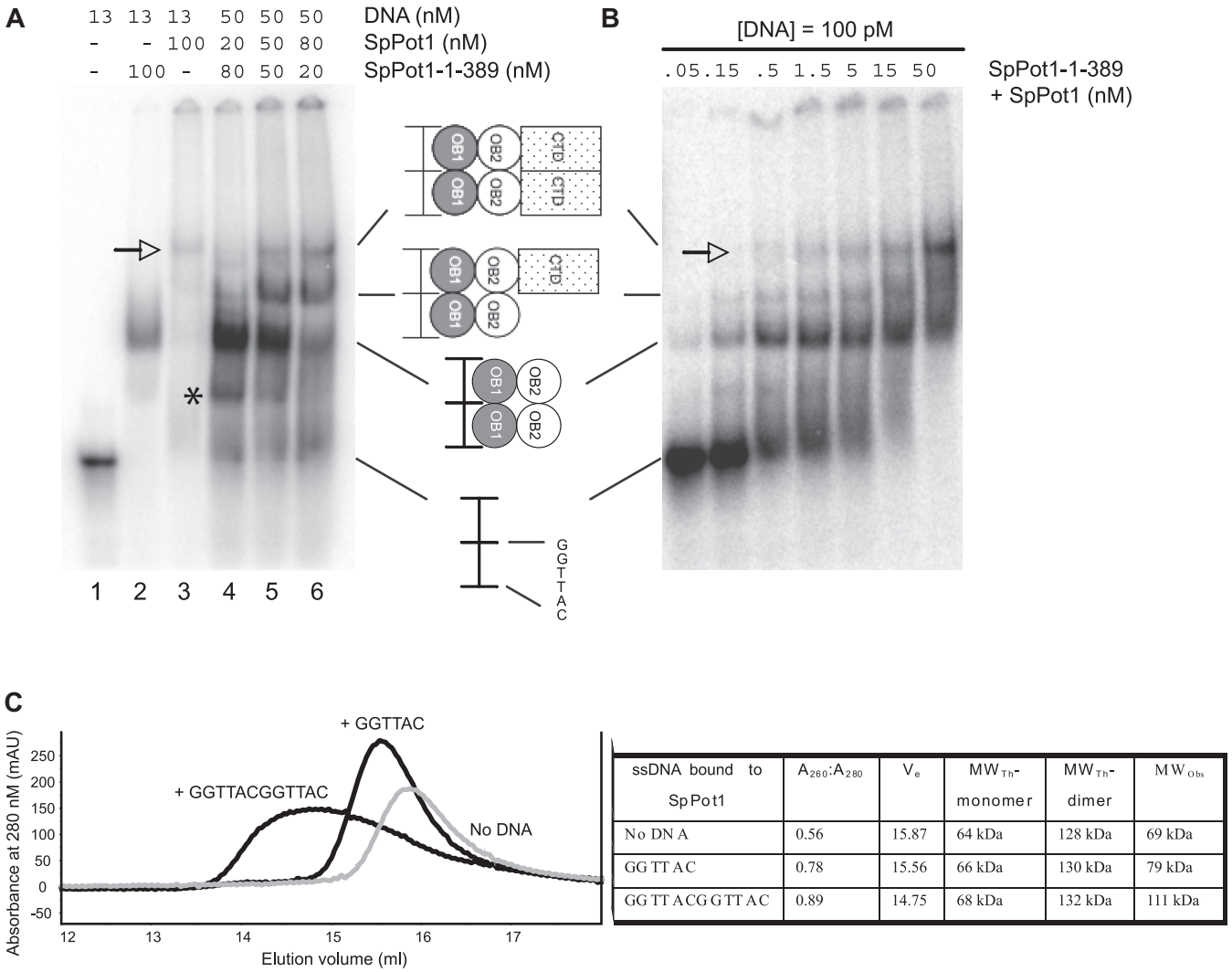
Sequence name	Oligonucleotide sequence	K_D with SpPot1 (nM)	K_D with SpPot1 + Tpz1-1–234 (nM)
d12	GGTTACGGTTAC	5.0 ± 1.3	0.22 ± 0.03
r3d9	**GGU**TACGGTTAC	11.6 ± 3.3	1.7 ± 0.08
d3r3d6	GGT**UAC**GGTTAC	7.0 ± 0.8	0.92 ± 0.12
d6r3d3	GGTTAC**GGU**TAC	4.4 ± 0.4	0.22 ± 0.03
d9r3	GGTTACGGT**UAC**	4.4 ± 0.5	0.21 ± 0.03
r6d6	**GGUUAC**GGTTAC	5.1 ± 1.6	0.90 ± 0.22
d6r6	GGTTAC**GGUUAC**	2.5 ± 0.8	0.26 ± 0.01
r3d3r3d3	**GGU**TAC**GGU**TAC	62.2 ± 9.7	68.4 ± 17.5
d3r3d3r3	GGT**UAC**GGT**UAC**	79.1 ± 8.9	~117
r3d6r3	**GGU**TACGGT**UAC**	~151	94.9 ± 3.5
d3r6d3	GGT**UACGGU**TAC	~117	~107
r1r7	**G**GTTAC**G**GTTAC	5.1 ± 0.01	0.31 ± 0.02
r2r8	G**G**TTACG**G**TTAC	3.2 ± 0.5	0.19 ± 0.03
r3r9	GG**U**TACGG**U**TAC	34.8 ± 3.8	21.3 ± 3.7
r4r10	GGT**U**ACGGT**U**AC	~100	70.3 ± 12.4
r5r11	GGTT**A**CGGTT**A**C	3.34 ± 0.4	0.31 ± 0.05
r6r12	GGTTA**C**GGTTA**C**	3.2 ± 0.6	0.30 ± 0.05
r3r4r9r10	GG**UU**ACGG**UU**AC	~238	n.d.
d12dA	GGTTACAGGTTAC	4.3 ± 0.6	n.d.
d12dAdGdG	GGTTACAGGGGTTAC	4.9 ± 0.7	n.d.
d12rA	GGTTAC**A**GGTTAC	3.4 ± 0.3	n.d.
d12rArGrG	GGTTAC**AGG**GGTTAC	4.0 ± 0.5	n.d.
r12	**GGUUACGGUUAC**	~324	~1000

Ribonucleotides are shown in black and deoxyribonucleotides in gray. K_D values denote the mean of two experiments ± the standard error. n.d. denotes K_Ds not determined.

spacer region (Table 1). Thus, RNA discrimination does not arise from ribonucleotides in the spacer region.

SpPot1 dimerizes on telomeric ssDNA

Our filter-binding data suggested that two SpPot1 molecules bind to a telomeric dodecamer using their respective OB1 domains. To test this directly, we designed an EMSA with T_{18}GGTTACGGTTAC employing a mixture of two SpPot1 variants: full-length protein (OB1–OB2–CTD) and SpPot1-1–389 (OB1–OB2), the DNA binding domain of SpPot1 (10). Both SpPot1 versions gave discrete gel-shifts when tested in isolation, with the migration consistent qualitatively with the size of the proteins (lanes 2 and 3 in Figure 3A). If these species represent monomers bound to DNA, then one would predict that the gel-shift pattern of a mixture of these proteins would be the summation of the patterns seen for each protein alone, specifically, a two-band pattern. In contrast, if the major species in lanes 2 and 3 of Figure 3A, represent two molecules each of the SpPot1 proteins bound to DNA, then with the protein mixture one would expect to see a



ssDNA bound to SpPot1	A_{260}:A_{280}	V_e	MW_{Th}-monomer	MW_{Th}-dimer	MW_{Obs}
No DNA	0.56	15.87	64 kDa	128 kDa	69 kDa
GGTTAC	0.78	15.56	66 kDa	130 kDa	79 kDa
GGTTACGGTTAC	0.89	14.75	68 kDa	132 kDa	111 kDa

Figure 3. Two SpPot1 molecules bind ssDNA containing two telomeric repeats. (**A**) EMSA analysis of a mixing experiment involving indicated concentrations of SpPot1 and the SpPot1-1–389 construct, and 50 nM ($>K_D$) T_{18}GGTTACGGTTAC. The arrow indicates the mobility of SpPot1–DNA alone. The asterisk represents a species that probably arises from partial dissociation of the SpPot1-1–389 dimer–DNA complex, because it is prominent only in lanes with high concentration of the dimer–DNA complex. (**B**) Titration of a 1:1 mixture of SpPot1 and SpPot1-1–389 in the presence of 100 pM ($<K_D$) T_{18}GGTTACGGTTAC was analyzed by EMSA. Schematic representations of the various SpPot1–DNA complexes formed during the assay are shown between (A and B). (**C**) Elution profiles of SpPot1, SpPot1 + GGTTAC and SpPot1 + GGTTACGGTTAC from size-exclusion chromatography. The UV absorbance at 280 nm for the indicated samples is plotted against elution volume on the left. The table on the right summarizes the size-exclusion data analysis. A_{260}:A_{280} ratios were calculated at the peak elution volume (Ve). MW_{Th}–monomer and MW_{Th}–dimer are the theoretical molecular weights of complexes containing one and two SpPot1 molecules. MW_{Obs} is the molecular weight calculated based on Ve.

three-band pattern that will include a new species corresponding to one SpPot1 and one SpPot1-1–389 bound to DNA. The mobility of this heterodimer–DNA complex would be expected to be intermediate to that seen for the long and short SpPot1 variants bound to DNA in isolation. In agreement with the 'two SpPot1s per DNA' hypothesis, a new species with a distinct intermediate mobility was observed when SpPot1 and SpPot1-1–389 mixtures were tested (Figure 3A, lanes 4–6).

To address whether the species with intermediate mobility could be a result of non-specific SpPot1 aggregation at high DNA and protein concentrations, we conducted an EMSA of a 1:1 mixture of the two SpPot1 variants at low (100 pM) T_{18}GGTTACGGTTAC concentration (Figure 3B). With increasing concentrations of the protein mixture, the lowest band appeared first, which is consistent with a sub-nanomolar K_D for SpPot1-1–389 for ssDNA (10) compared to SpPot1 (Table 1). With further increase of concentration, all three species (including the species with intermediate mobility) were observed, suggesting that the dimerization of SpPot1 on the DNA is not limited to conditions involving high protein/DNA concentration. The observation that SpPot1 and SpPot1-1–389 (OB1–OB2) form a heterodimer on DNA by gel-shift analysis suggests that OB1–OB2 is sufficient for dimerization on DNA, although we cannot rule out

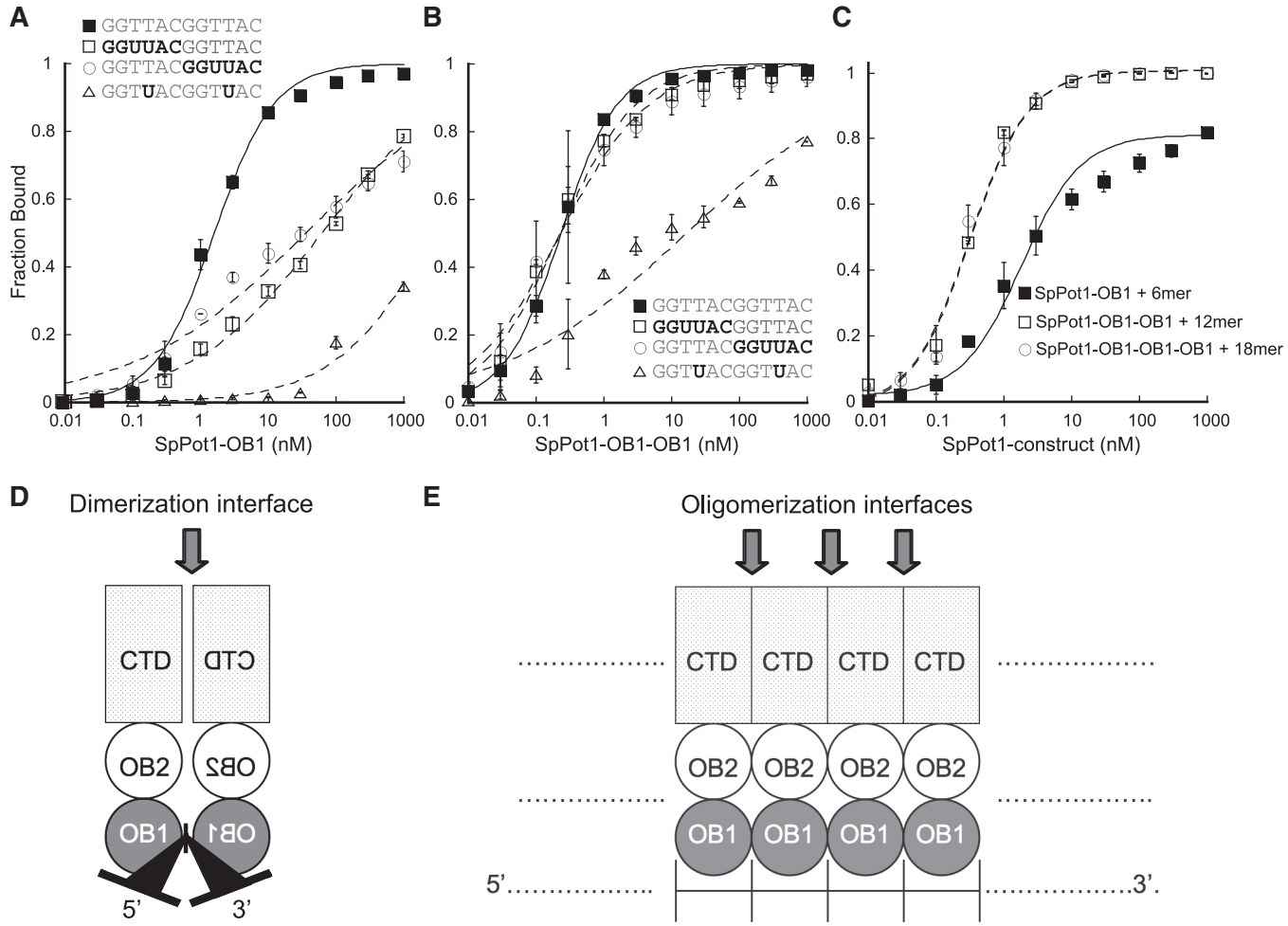


Figure 4. OB1–OB1 tandem fusion protein exhibits full-length SpPot1-like DNA binding and RNA discrimination characteristics. Filter-binding profiles of OB1 (**A**) or OB1–OB1 (**B**) with oligonucleotides of indicated sequence are shown. Ribonucleotides are depicted in black and deoxyribonucleotides in gray. (**C**) Filter-binding profiles of OB1–GGTTAC, OB1–OB1–GGTTACGGTTAC and OB1–OB1–OB1–GGTTACGGTT ACGGTTAC are shown with error bars indicating standard deviations of two independent experiments. (**D**) Model for specific dimerization of SpPot1 on telomeric ssDNA. OB1 and OB2 represent the two N-terminal OB domains of SpPot1 and CTD represents amino acids of SpPot1 C-terminal to amino acid 389. (**E**) Model for head-to-tail oligomerization of SpPot1 on telomeric ssDNA, which was not observed.

that additional elements on the CTD of SpPot1 might further contribute to dimerization.

Is SpPot1 a dimer in the absence of DNA or is dimerization induced by DNA binding? To address this, we conducted size-exclusion chromatography of SpPot1 either in isolation, or in complex with GGTTAC or GG TTACGGTTAC (on a Superdex 200 analytical column calibrated using protein standards; see 'Materials and Methods' section and Supplementary Figure S1). The molecular weights calculated based on the elution volume for the protein/complex peak upon chromatography (see 'Materials and Methods' section for molecular weight calculation) suggest that in the absence of DNA or when bound to GGTTAC, SpPot1 is a monomer (Figure 3C), whereas a shift to a higher observed molecular weight suggesting dimerization of SpPot1 is evident in the presence of GGTTACGGTTAC (Figure 3C). A broad peak for the SpPot1-GGTTACGGTTAC complex (elution profile in Figure 3C) and the fact that the interpolated molecular weight of SpPot1- GGTTACGGTTAC is 21 kDa lower than the theoretical molecular weight of a dimer-containing complex (inset table of Figure 3C) may be explained by partial dissociation or incomplete formation of the SpPot1 dimer-DNA complex, which is also evident in EMSA experiments (asterisks in Figure 3A and Figure 5B).

The OB1–OB1 tandem fusion mimics the telomeric DNA-binding and RNA-discrimination functions of full-length SpPot1

To evaluate if OB1 mimicked full-length SpPot1 in its RNA discrimination profile on a telomeric dodecamer, we expressed and purified an active OB1 polypeptide that efficiently bound telomeric ssDNA (Figure 1A and B; see 'Materials and Methods' section) and assayed its binding to all-DNA and tri-ribo-substituted oligonucleotides by filter-binding. With OB1, reduced affinity is seen when either hexameric repeat is all-RNA (Figure 4A), which is in sharp contrast to binding data with full-length protein, where substitution of any one repeat with

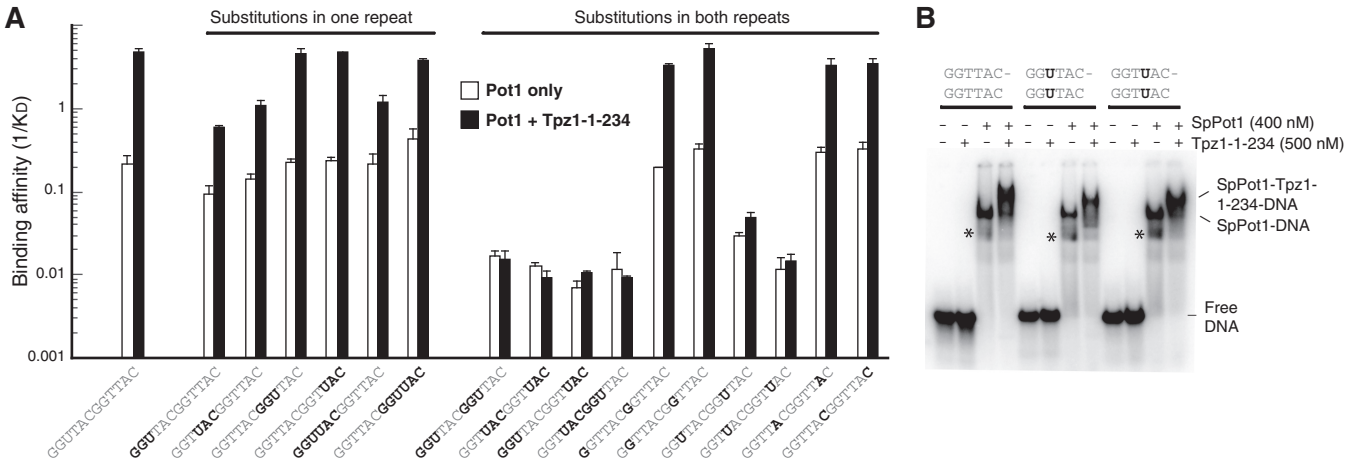


Figure 5. Tpz1 increases SpPot1's affinity for telomeric DNA and discrimination against RNA. (**A**) The inverse of K_D (measure of binding affinity) is plotted as a bar graph for dodecameric oligonucleotides containing a mixture of ribonucleotides (black letters) and deoxyribonucleotides (gray) in the telomeric sequence. The white bars and black bars indicate binding affinities with SpPot1 alone and with SpPot1 + Tpz1-1–234. Error bars indicate standard deviations of two independent experiments. (**B**) EMSA of the indicated oligonucleotides with SpPot1 in the absence or presence of Tpz1-1–234 showing that Tpz1-1–234 binds SpPot1–DNA to elicit a super-shift with all-DNA and ribo-substituted DNA dodecamers. The asterisk represents a species probably resulting from partial dissociation of the SpPot1 dimer–DNA complex.

ribonucleotides failed to elicit a binding defect (Figure 2A). The shallow slope for transitions involving hexa-ribonucleotide-substituted oligonucleotides probably reflects a mixture of one strong binding event (involving DNA hexamer) and one weak binding event (involving RNA hexamer).

To account for the difference in the RNA discrimination profiles between full-length SpPot1 and OB1 alone, we hypothesized that binding of the first SpPot1 to DNA creates a binding site for the second molecule (mediated by a protein–protein interface) that can mask the deleterious effects of ribo-substitution in a single repeat. Accordingly, the reason why OB1 in isolation does not recapitulate full-length-like RNA discrimination on a dodecamer is that amino acids of SpPot1 that are C-terminal to OB1 are required to reconstitute the complete putative protein–protein interface.

We therefore tested this hypothesis with an OB1–OB1 dimer construct, obtained via fusing two OB1 ORFs (see 'Materials and Methods' section), to see if it would rescue the dimerization defect of OB1 and mimic full-length-like DNA-binding and RNA-binding characteristics. The OB1-OB1 fusion was expressed in bacteria, purified to homogeneity and shown to form a discrete complex with telomeric ssDNA in an EMSA (see 'Materials and Methods' section, and Figure 1A and B). Indeed, the OB1–OB1 protein has higher DNA binding affinity for a 12-mer compared to OB1 (0.21 nM versus 1.4 nM; Supplementary Figure S2), consistent with additive binding of two OB1 domains of OB1-OB1 to DNA. Intriguingly, the RNA discrimination profile of OB1–OB1 is identical to that of full-length SpPot1, but not that of OB1 alone. Like with full-length SpPot1 (Figure 2A), with OB1–OB1 (Figure 4B) both hexa-ribonucleotide substituted oligonucleotides (that showed defects with OB1 binding; Figure 4A) showed all-DNA like binding profiles, whereas the dT4/dT10 to rU4/rU10 substituted oligonucleotide showed reduced binding. Hence, the OB1-OB1 fusion protein fully recapitulates SpPot1 DNA binding and RNA discrimination properties.

Higher affinity of OB1–OB1 versus OB1 for ssDNA strongly supports a model where a stable dimer of SpPot1 is bound to ssDNA (Figure 4D). However, it is also possible that by tethering two OB1 domains, we are increasing affinity by forcing two domains to bind DNA in a manner independent of protein–protein interaction. The latter model would predict that addition of more telomeric repeats and tethering of more OB1 domains together would result in formation of a linear polymer of SpPot1 on DNA (Figure 4E). To distinguish between the two possibilities, we designed an OB1-OB1-OB1 construct (see 'Material and Methods' section, and Figure 1A and B), which contained three OB1 ORFs fused in tandem, and compared the binding affinities of OB1–G GTTAC, OB1–OB1–GGTTACGGTTAC, and OB1–OB1–OB1-GGTTACGGTTACGGTTAC (Figure 4C). If linking of OB1 domains increases DNA affinity solely by increasing DNA binding surface on the protein, one would expect OB1–OB1–OB1 to have greater affinity for an 18-mer DNA compared to OB1–OB1 for a 12-mer DNA. In contrast, the stable dimer model would predict that OB1–OB1–OB1 would bind the 18-mer DNA using just two OB1 domains, and hence the affinity would resemble that of OB1–OB1 for a 12-mer DNA. Consistent with our stable-dimer model (Figure 4D), the binding affinity of OB1–OB1–OB1 with the 18-mer was indistinguishable from that of OB1–OB1 with the 12-mer but higher (by ~10-fold) than that of OB1 with a 6-mer, providing further evidence that SpPot1 binds telomeric ssDNA as a stable dimer (Figure 4C).

Tpz1 increases affinity of SpPot1 for telomeric DNA and discrimination of SpPot1 against RNA

Tpz1 is the binding partner of SpPot1 *in vivo* and is believed to be the *S. pombe* counterpart of mammalian

TPP1 (12). We designed the Tpz1-1–234 construct based on secondary structure predictions and results from Miyoshi *et al.* (12), who showed that the first 223 amino acids of Tpz1 suffice for SpPot1 interaction *in vivo* (Figure 1A). Like mammalian TPP1, Tpz1 by itself does not show high affinity for telomeric DNA, but forms a stable ternary complex with SpPot1–DNA (Figure 1B). To test whether Tpz1 increases affinity of SpPot1 for telomeric ssDNA like mammalian TPP1 does (11), we conducted filter-binding experiments of SpPot1 with GGTTACGGTTAC in the presence of 200 nM Tpz1-1–234 and observed a 23-fold increase in affinity compared to experiments done with SpPot1 alone (0.22 ± 0.03 nM versus 5.0 ± 1.3 nM; Figure 5A and Table 1).

Studies on human and mouse POT1 proteins have shown that ribo-substitutions that decrease POT1-DNA affinity also decrease TPP1's ability to stimulate the POT1–DNA interaction (19). Here, we tested various ribo-substituted oligonucleotides for stimulation of binding with SpPot1 by Tpz1-1–234 and observed that all substrates that were defective in SpPot1-binding (due to dT to rU substitutions in both repeats; short white bars in Figure 5A) showed no significant stimulation of binding when SpPot1 was supplemented with Tpz1-1–234 (compare short white bars with accompanying black bars in Figure 5A). To test the possibility that lack of a 'Tpz1 effect' on the dT to rU substituted oligonucleotides is due to inability to form SpPot1–Tpz1 complexes on these substrates, we carried out an EMSA and probed for retardation of SpPot1–DNA complexes in the presence of Tpz1 (Figure 5B). Both ribo-substituted as well as all-DNA dodecamer complexes of SpPot1 underwent a supershift when Tpz1-1–234 was included, discounting the trivial interpretation that failure of Tpz1 to enhance affinity was due to failure to bind. Thus, we conclude that the ternary Pot1–Tpz1–oligonucleotide complex still forms in the presence of rU substitutions in both repeats, but that the detailed manner in which Tpz1 binds is no longer able to stabilize Pot1–DNA.

Inspection of Table 1 revealed that discrimination against rU3/rU9- or rU4/rU10-containing oligonucleotides by SpPot1 is significantly larger in the presence of Tpz1-1–234 than in its absence. For instance, SpPot1–r4r10 showed a 20-fold (~100/5.0; Table 1) increase of K_D with respect to SpPot1–d12 in the absence of Tpz1-1–234, whereas the increase in K_D was 320-fold (70.3/0.22; Table 1) in the presence of the Tpz1 construct. These data imply that the discrimination against RNA is greatly enhanced when SpPot1 is bound by Tpz1.

DISCUSSION

Through results presented here, we have uncovered a new feature of SpPot1 binding to telomeric DNA, namely DNA-induced dimerization. Binding studies with ribo-substituted oligonucleotides revealed that both telomeric repeats in GGTTACGGTTAC contribute equally to RNA discrimination, suggesting that identical (or very similar) protein components engage the two covalently linked GGTTAC repeats. One possibility is that in SpPot1, the OB1 and OB2 domains might have identical binding modes to GGTTAC repeats, which would be in contrast to mammalian POT1s, where the two OB domains have been shown to interact differently with the two telomeric repeats. This scenario seems improbable given that there is scarce sequence similarity between OB1 and OB2 of SpPot1 and that the DNA binding characteristics of OB2 in isolation are distinct from those of OB1 or OB1–OB2 (10,15).

An alternative explanation for the functional 'symmetry' observed within GGTTACGGTTAC is that SpPot1 dimerizes on two GGTTAC repeats using the OB1 domain of each monomer. This possibility seems probable because the RNA discrimination profile of SpPOT1 with GG**TT**ACGG**TT**AC we observe here is an exact duplicate of the RNA discrimination profile of OB1 with GG**TT**AC. Also, previous studies by Lei *et al.* (13), showing that OB1 bound long stretches of telomeric ssDNA with high cooperativity, provided precedence for oligomerization of SpPot1 on DNA. To test our 'dimer hypothesis', we conducted an EMSA with a mixture of two versions of SpPot1: full-length protein and SpPot1-1–389 (OB1–OB2), which is the DNA binding domain of SpPot1 (10). With this mixture, we clearly observed formation of a new species representing SpPot1–SpPot1-1–389–GGTTACGGTTAC, providing strong evidence of dimerization of SpPot1 on DNA (Figure 3A and B).

Using size-exclusion chromatography, we determined that SpPot1 exists as a monomer in solution in the absence of DNA or when bound to a single telomeric repeat, which according to our dimer hypothesis allows for only one SpPot1 binding. In contrast, the SpPot1–GGTTACGGTTAC complex showed a significant shift towards a lower elution volume, providing strong evidence for DNA-induced dimerization of SpPot1 on a telomeric dodecanucleotide (Figure 3C). Although the binding data and size-exclusion studies are consistent with a dimer model, they do not rule out non-specific dimerization of SpPot1 on a 12-mer ssDNA. To distinguish between a specific dimer (Figure 4D) versus head-to-tail non-specific oligomerization (Figure 4E) of SpPot1 on DNA, we exploited tandem OB1 fusion proteins. Consistent with our specific dimer model, increasing the number of tethered OB1 domain from two copies to three copies with concomitant increase in target telomeric ssDNA from two repeats to three repeats did not result in an increase in binding affinity (Figure 4C).

Based on the data presented, we propose a model for SpPot1 dimerization on DNA that can explain the unique RNA discrimination profile of SpPot1 and the OB1 constructs tested here (Supplementary Figure S3). With GGTTACGGTTAC, the first SpPot1 molecule binds via the structurally characterized OB1-GGTTAC interaction (6) (grey arrow in Supplementary Figure S3A). This binding event creates a high-affinity second SpPot1-binding site possessing a protein–DNA (grey arrow) surface and a protein–protein interaction surface (black arrow) leading to SpPot1 dimer formation on DNA. When dT to rU substitutions are present in one repeat only, the first SpPot1 binding occurs at the all-DNA repeat. Although

binding of the second SpPot1 molecule is hindered by the ribo-substitution ('X' in Supplementary Figure S3B), this impediment is bypassed by the putative SpPot1–SpPot1 interaction, leading to successful dimer formation on the DNA. If rT to rU substitutions are present on both telomeric repeats, then the first step of binding is hindered, precluding dimer formation and resulting in a large observable binding defect (Supplementary Figure 3C).

Our dimer-model also rationalizes effectively the DNA and RNA binding properties of the various OB1 constructs. OB1 displays lower affinity and shallow binding transitions with oligonucleotides with ribo-substitutions in a single repeat, consistent with a mixture of a strong and weak binding event (Figure 4A). This would suggest that the two OB1 molecules are binding the two repeats independent of each other. According to our model, OB1 retains the DNA binding interface, but not the entire protein–protein interface needed to facilitate binding of the second SpPot1 OB1 molecule in rU-substituted oligonucleotides (Supplementary Figure S3D). The OB1–OB1 fusion protein retains the two ligands required for full interaction with telomeric DNA, and emulates the protein–protein interface by replacing the amino acids of SpPot1 C-terminal to OB1 (required for dimerization according to our model) with a covalent link (Supplementary Figure 3E). Hence, like full-length SpPot1, OB1–OB1 binds with equal efficacy to all-DNA oligonucleotides or those with dT to rU substitutions in one repeat only, but shows binding defects when these substitutions are present in both repeats.

What is the biological significance of SpPot1 dimerization on DNA? One difference between *S. pombe* telomeres and mammalian telomeres is the presence of spacer nucleotides separating hexameric repeats in the former but not the latter. Hence, although a 12-mer DNA consisting of two telomeric repeats can be defined as the binding unit for mammalian POT1s, the longest substrate that can be strictly defined for sequence-specific SpPot1 binding is a 6mer DNA containing one complete telomeric repeat. To compensate for the 50% loss in the number of nucleotides bound if SpPot1 only bound one telomeric repeat, the protein might have evolved to bind two telomeric repeats in close vicinity (but that are not necessarily contiguous) by using two OB1 domains that are tethered via dimerization. The point of tether would act as a pivot around which the two OB1 domains could swing around to accommodate the heterogeneous spacers between adjacent telomeric repeats. Although not seen with mammalian POT1s, dimerization has been observed for another telomeric ssDNA-binding protein, Cdc13, found in *Saccharomyces cerevisiae* (21), an organism that also possesses a heterogeneous telomere sequence.

Very recently, Altschuler *et al.* (22) reported that GGTT ACGGTTAC**GGT** (15-mer) is an efficient binding substrate for SpPot1, and is bound by SpPot1 in a mode distinct from that used to bind GGTTACGGTTAC (12-mer), the substrate used in this study. In contrast to dimerization of SpPot1 on the 12-mer shown here, gel-shift analysis and multiangle light scattering experiments by Altschuler *et al.* show that full-length SpPot1 binds the 15-mer as a monomer. Given that the 15-mer substrate has exactly one binding site each for OB1 (GGT TAC) and OB2 (GGTTACGGT), we update our model by proposing that SpPot1 binds a 15-mer as a monomer using OB1 to bind the 5′-most GGTTAC repeat and OB2 to bind the remaining GGTTACGGT nucleotides (Supplementary Figure S3F).

Our updated model, which suggests that SpPot1 binds a 12mer as a dimer using two OB1 domains, and a 15mer as a monomer using OB1 and OB2, is fully consistent with base-substitution data reported in the Altschuler *et al.* study. Specifically, Altschuler *et al.* observe that with regards to SpPot1 binding, base-substitutions in either GGTTAC of the 12-mer lead to large binding defects, whereas base-substitution in only the 5′-most GGTTAC repeat of the 15mer leads to a substantial binding defect. This observation can be easily explained using our model, keeping in mind OB1's high sequence specificity (6) and OB2's low sequence specificity (15). Since binding of SpPot1 to a 12-mer is mediated by OB1 binding to both repeats, specificity determinants are found in both repeats of this substrate. On the other hand, because the OB1 of SpPot1 binds only the 5′-most GGTTAC repeat of a 15-mer, specificity determinants are found only in this repeat of the oligonucleotide.

The protein mutational data presented by Altschuler *et al.* are also explained easily in light of our updated model. The authors report that mutations in the known DNA binding site of OB1 made in the context of SpPot1-1–389 cause a greater binding defect with the 12-mer (280- to 2700-fold) than with the 15-mer (5- to 60-fold). Since our model invokes OB1 being associated with both GGTTAC repeats of the 12-mer (12 out of 12 nt) but only with the 5′-most GGTTAC repeat in the 15-mer (6 out of 15 nt), it is not surprising that a larger decrease in binding with OB1-mutants is observed with the 12-mer versus the 15-mer.

What could be the biological significance of the co-existence of dual binding modes of SpPot1 on telomeric DNA? One possible scenario is that the 15-mer is the 'default' substrate for SpPot1 binding and that binding to the 12-mer via dimerization evolved as a strategy for SpPot1 to bind to short (~12 nt) ss G-overhangs. This could prevent chromosome end-deprotection at shorter telomere overhangs that cannot be efficiently bound by an SpPot1 monomer because of an incomplete OB2 binding site (only six out of the nine target nucleotides of OB2 are available for binding in a 12-mer). An alternative, albeit more speculative rationale for the existence of dual DNA binding modes of SpPot1 is that the monomer-to-dimer SpPot1 switch, occurring when overhang lengths shorten to 12 nt, could provide a platform for recruiting two Tpz1 molecules. These would then bind two Ccq1 (12) molecules, helping to recruit dimeric telomerase (assuming telomerase in *S. pombe* is dimeric) and selectively lengthening telomeres with short overhangs.

Finally, we demonstrated biochemically that Tpz1 is a functional homolog of mammalian TPP1, because Tpz1 forms a ternary complex with SpPot1–ssDNA, increases the affinity of SpPot1 for DNA, and also increases the discrimination of RNA by SpPot1. All of these properties

are characteristic of mammalian TPP1 proteins. The increased discrimination of SpPot1 against RNA of telomeric sequence when it is bound to Tpz1 is particularly relevant in the context of the discovery of TERRA, telomeric RNA that is transcribed in *S. pombe* (as well as in mammals and *S. cerevisiae*). In the absence of such enhanced RNA discrimination, SpPot1 might be sequestered by TERRA and would not be able to bind telomeres to perform its end-protection function.

SUPPLEMENTARY DATA

Supplementary data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Fuyuki Ishikawa (Kyoto U.) for providing the cDNA for SpPot1 and Tpz1, Deborah S. Wuttke and Sarah E. Altschuler (U. of Colorado-Boulder) for providing SpPot1-1–389 protein, for sharing unpublished results and for numerous helpful discussions.

FUNDING

National Institutes of Health (grant GM28039) to T.R.C.; Howard Hughes Medical Institute-Helen Hay Whitney Foundation fellowship to J.N. Funding for open access charge: Howard Hughes Medical Institute.

Conflict of interest statement. None declared

REFERENCES

1. Blackburn,E.H. (2001) Switching and signaling at the telomere. *Cell*, **106**, 661–673.
2. Palm,W. and De Lange,T. (2008) How shelterin protects mammalian telomeres. *Annu. Rev. Genet*., **42**, 301–334.
3. Makarov,V.L., Hirose,Y. and Langmore,J.P. (1997) Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell*, **88**, 657–666.
4. Wright,W.E., Tesmer,V.M., Huffman,K.E., Levene,S.D. and Shay,J.W. (1997) Normal human chromosomes have long G-rich telomeric overhangs at one end. *Genes Dev*., **11**, 2801–2809.
5. Baumann,P. and Cech,T.R. (2001) Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science*, **292**, 1171–1175.
6. Lei,M., Podell,E.R., Baumann,P. and Cech,T.R. (2003) DNA self-recognition in the structure of Pot1 bound to telomeric single-stranded DNA. *Nature*, **426**, 198–203.
7. Lei,M., Podell,E.R. and Cech,T.R. (2004) Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection. *Nat. Struct. Mol. Biol*., **11**, 1223–1229.
8. Wu,L., Multani,A.S., He,H., Cosme-Blanco,W., Deng,Y., Deng,J.M., Bachilo,O., Pathak,S., Tahara,H., Bailey,S.M. *et al.* (2006) Pot1 deficiency initiates DNA damage checkpoint activation and aberrant homologous recombination at telomeres. *Cell*, **126**, 49–62.
9. Loayza,D. and De Lange,T. (2003) POT1 as a terminal transducer of TRF1 telomere length control. *Nature*, **423**, 1013–1018.
10. Croy,J.E., Podell,E.R. and Wuttke,D.S. (2006) A new model for Schizosaccharomyces pombe telomere recognition: the telomeric single-stranded DNA-binding activity of Pot11-389. *J. Mol. Biol*., **361**, 80–93.
11. Wang,F., Podell,E.R., Zaug,A.J., Yang,Y., Baciu,P., Cech,T.R. and Lei,M. (2007) The POT1-TPP1 telomere complex is a telomerase processivity factor. *Nature*, **445**, 506–510.
12. Miyoshi,T., Kanoh,J., Saito,M. and Ishikawa,F. (2008) Fission yeast Pot1-Tpp1 protects telomeres and regulates telomere length. *Science*, **320**, 1341–1344.
13. Lei,M., Baumann,P. and Cech,T.R. (2002) Cooperative binding of single-stranded telomeric DNA by the Pot1 protein of Schizosaccharomyces pombe. *Biochemistry*, **41**, 14560–14568.
14. Trujillo,K.M., Bunch,J.T. and Baumann,P. (2005) Extended DNA binding site in Pot1 broadens sequence specificity to allow recognition of heterogeneous fission yeast telomeres. *J. Biol. Chem*., **280**, 9119–9128.
15. Croy,J.E., Altschuler,S.E., Grimm,N.E. and Wuttke,D.S. (2009) Nonadditivity in the recognition of single-stranded DNA by the Schizosaccharomyces pombe protection of telomeres 1 DNA-binding domain, Pot1-DBD. *Biochemistry*, **48**, 6864–6875.
16. Azzalin,C.M., Reichenbach,P., Khoriauli,L., Giulotto,E. and Lingner,J. (2007) Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science*, **318**, 798–801.
17. Luke,B. and Lingner,J. (2009) TERRA: telomeric repeat-containing RNA. *EMBO J*., **28**, 2503–2510.
18. Luke,B., Panza,A., Redon,S., Iglesias,N., Li,Z. and Lingner,J. (2008) The Rat1p 5′ to 3′ exonuclease degrades telomeric repeat-containing RNA and promotes telomere elongation in Saccharomyces cerevisiae. *Mol. Cell*, **32**, 465–477.
19. Nandakumar,J., Podell,E.R. and Cech,T.R. (2010) How telomeric protein POT1 avoids RNA to achieve specificity for single-stranded DNA. *Proc. Natl Acad. Sci. USA*, **107**, 651–656.
20. Mossessova,E. and Lima,C.D. (2000) Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Mol. Cell*, **5**, 865–876.
21. Sun,J., Yang,Y., Wan,K., Mao,N., Yu,T.Y., Lin,Y.C., DeZwaan,D.C., Freeman,B.C., Lin,J.J., Lue,N.F. *et al.* (2011) Structural bases of dimerization of yeast telomere protein Cdc13 and its interaction with the catalytic subunit of DNA polymerase alpha. *Cell Res*., **21**, 258–274.
22. Altschuler,S.E., Dickey,T.H. and Wuttke,D.S. (2011) Schizosaccharomyces pombe protection of telomeres 1 utilizes alternate binding modes to accommodate different telomeric sequences. *Biochemistry* [Epub ahead of print; August 4, 2011; doi:10.1021/bi200826a].