THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 279, No. 13, Issue of March 26, pp. 13241–13248, 2004

Printed in U.S.A.

DNA Binding Features of Human POT1

A NONAMER 5′-TAGGGTTAG-3′ MINIMAL BINDING SITE, SEQUENCE SPECIFICITY, AND INTERNAL BINDING TO MULTIMERIC SITES*

Received for publication, November 10, 2003, and in revised form, January 7, 2004
Published, JBC Papers in Press, January 7, 2004, DOI 10.1074/jbc.M312309200

Diego Loayza‡§, Heather Parsons‡, Jill Donigian, Kristina Hoke¶, and Titia de Lange‖

From the Laboratory for Cell Biology and Genetics, The Rockefeller University, New York, New York 10021

The human telomeric protein POT1 is known to bind single-stranded telomeric DNA *in vitro* and to participate in the regulation of telomere maintenance by telomerase *in vivo*. We examined the *in vitro* DNA binding features of POT1. We report that deleting the oligosaccharide/oligonucleotide-binding fold of POT1 abrogates its DNA binding activity. The minimal binding site (MBS) for POT1 was found to be the telomeric nonamer 5′-TAGGGTTAG-3′, and the optimal substrate is $[TTAGGG]_{n\ (n \geq 2)}$. POT1 displays exceptional sequence specificity when binding to MBS, tolerating changes only at position 7 (T7A). Whereas POT1 binding to MBS or $[TTAGGG]_2$ was enhanced by the proximity of a 3′ end, POT1 was able to bind to a $[TTAGGG]_5$ array when positioned internally. These data indicate that POT1 has a strong sequence preference for the human telomeric repeat tract and predict that POT1 can bind both the 3′ telomeric overhang and the displaced TTAGGG repeats at the base of the t-loop.

Human telomeres are composed of 6–10 kbp of double-stranded TTAGGG repeats that end in a single-stranded overhang of several hundred nucleotides. The 3′ end of the telomeric overhang is the substrate of telomerase, the cellular reverse transcriptase that synthesizes telomeric repeats. The telomere can adopt a "t-loop" configuration, in which the 3′ overhang strand-invades the duplex (1). The strand invasion results in a displacement loop (D-loop) of single-stranded TTAGGG repeats; these single-stranded repeats are not at a 3′ end. The actual 3′ end of the chromosome is predicted to be base-paired to the C-strand and therefore inaccessible to telomerase when telomeres are in the t-loop configuration.

Proteins that bind to the double-stranded portion of human telomeres have been studied extensively (reviewed in Ref. 4). In particular, TRF1 binds as a dimer through a MYB-type helix-loop-helix domain. The MYB domains of TRF1 dimers bind to two double-stranded 5′-YTAGGGTTR-3′ half-sites independent of spacing or orientation (2, 3). Other proteins such as tankyrase 1 and 2, Tin2, PINX1, and POT1 are recruited to telomeres by TRF1 (4). The TRF1 complex is involved in the regulation of telomere length by cis-inhibition of telomerase. In telomerase-positive cells, the overexpression of TRF1 leads to telomere shortening, and the expression of a dominant negative form leads to telomere elongation (5). The relationship between TRF1 and telomerase regulation at the 3′ telomere terminus is poorly understood.

The human single-stranded telomeric DNA-binding protein POT1 was identified based on its sequence similarity to the *Oxytricha nova* TEBPα, known to be associated to the 16-base single-stranded telomeric extension in this organism (6). Orthologs in *Schizosaccharomyces pombe*, *Arabidopsis*, mouse, human, and other eukaryotes have been identified by homology to the N-terminal OB[1] (oligosaccharide/oligonucleotide-binding) fold, a structural domain involved in DNA binding (7). The crystal structure of TEBPα revealed three OB folds, two involved in DNA binding and the third one necessary for a protein interaction with TEBPβ (8). TEBPα, -β, and the DNA form a ternary complex in which the α/β dimer forms a cage around the DNA, which itself is folded into a hairpin structure (8). No ortholog of TEBPβ has been found in mammals.

In *S. pombe*, Pot1 (spPot1) is essential for the protection of chromosome ends (6). Deletion of the *pot1+* gene leads to rapid telomere degradation. Survivors of this telomere loss have circularized chromosomes that lack all telomeric DNA. Binding studies have determined that spPot1 binds to a sequence representing fission yeast telomeric DNA (repeats of GGTTACA) but not to the human telomeric sequence (TTAGGG). The minimal binding site for the isolated DNA binding domain consists of six nucleotides, GGTTAC, with a binding constant of 83 nM (9). Cooperative binding occurs on oligonucleotides containing multimeric sites, and spPot1 displays a preference for sequences close to a 3′ end. Taken together these data are consistent with a model in which the protein initially binds the very 3′ end of chromosomes and subsequently coats the entire telomeric overhang.

Human POT1 has two important domains required for its function: an N-terminal OB fold, predicted to be necessary for DNA binding, and a protein interaction domain, mediating association with the TRF1 complex (10). Thus, POT1 is a good candidate for providing the link between TRF1 on the duplex telomeric portion and the 3′ overhang where telomerase acts. It is still unclear whether DNA binding is a primary event in POT1 targeting to telomeres or if it occurs after recruitment by TRF1.

POT1 can be detected by immunofluorescence at telomeres, and can associate with telomeres through its interaction with

* This work is supported in part by Grants CA076027 and T32 CA09673–26 from the NCI, National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ These authors contributed equally to this study.

§ Recipient of an Ann Siegel postdoctoral fellowship from the American Cancer Society (PF-00-135-01).

¶ Supported by National Institutes of Health Medical Scientist Training Program Grant GM07739 to the Cornell/Rockefeller/Sloan-Kettering Tri-Institutional M.D.-Ph.D. Program.

‖ To whom correspondence should be addressed: Laboratory for Cell Biology and Genetics, The Rockefeller University, 1230 York Ave, New York, NY 10021. Tel.: 212-327-8146; Fax: 212-327-7147; E-mail: delange@mail.rockefeller.edu.

[1] The abbreviations used are: OB fold, oligosaccharide/oligonucleotide-binding fold; MBS, minimal binding site; GST, glutathione *S*-transferase; RPA, replication protein A; nt, nucleotide(s).

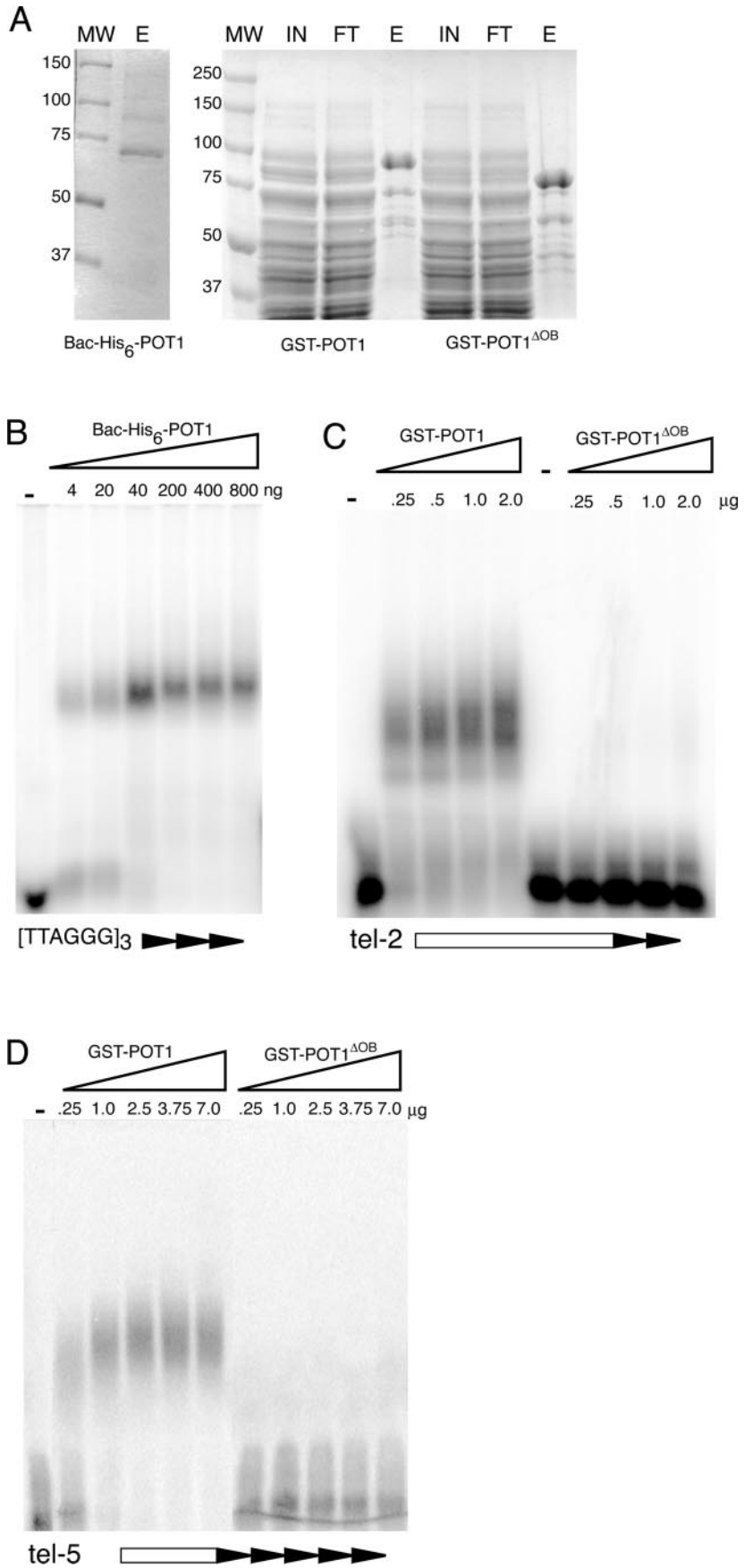


FIG. 1. **The OB fold of POT1 is necessary for DNA binding.** *A*, purification of Bac-His_6-POT1 from Sf21 cells (*left*) and GST-POT1 (*middle*) and GST-$POT1^{\Delta OB}$ (*right*) from *E. coli*. The elution (*E*), input (*IN*), and flow-through (*FT*) fractions were separated by SDS-PAGE and stained with Coomassie Blue. *B*, direct binding of Bac-His_6-POT1 to an oligonucleotide probe with three single-stranded telomeric repeats without random sequence. Here and below, each TTAGGG repeat is represented by an *arrowhead*. *C* and *D*, deletion of the N-terminal OB fold abrogates binding. The $POT1^{\Delta OB}$ truncation removes 126 amino acids from the N terminus. The probes in *C* and *D* contained 2 or 5 copies of TTAGGG at the 3′ end of a 50-mer (tel-2 and tel-5, respectively; see Table I).

the TRF1 complex (10). The exact role of the DNA binding activity of POT1 is unknown, but the expression of a mutant form of POT1 missing the OB fold ($POT1^{\Delta OB}$) leads to extensive telomere elongation in telomerase-positive cells (10). The simplest model is that POT1 lies downstream of the TRF1 complex in the telomere length regulation pathway. POT1 is proposed to

inhibit telomerase *in cis* at chromosome ends in response to the length of the duplex portion and as such could play a role in the counting mechanism performed by the TRF1 complex. POT1 has also been implicated as a positive factor for telomere length control. In some settings, overexpression of full-length POT1 (with short N-terminal and C-terminal extensions) was found to lead to telomere elongation (11). It is possible that POT1 has a dual mode of function at telomeres; one would be inhibitory to telomerase, and the other promoting telomere elongation. Such a dual role in telomere length regulation has been proposed for the distantly related protein Cdc13p in *Saccharomyces cerevisiae* (12).

The main form of POT1 detected by Western blotting is the full-length protein (10). However, POT1 transcripts are subject to alternative splicing, possibly leading to the expression of C-terminal truncations in certain cell types (7). *In vitro*, one of the variants (variant 2), corresponding to the N-terminal 38 kDa of the protein, displays an 8-fold higher binding affinity than full-length POT1. It is not known whether the alternatively spliced transcripts yield stable proteins or what the role of these variants in telomere biology might be.

Human POT1 binds to single-stranded human telomeric TTAGGG repeats but not to double-stranded telomeric DNA nor to the C-rich telomeric repeat strand (6). POT1 only binds efficiently to the human telomeric sequence and not to the *S. pombe* telomeric DNA nor to the *O. nova* TTTTGGGG sequence (6). To date, the minimal binding site, sequence specificity, and 3′ end dependence of POT1 have not been studied in detail. Here, we address these and other aspects of the DNA binding features of POT1.

EXPERIMENTAL PROCEDURES

Purification of Escherichia coli or Baculovirus-expressed POT1—The cloning of the POT1 and POT1$^{\Delta OB}$ cDNAs was described previously (10). The full-length POT1 cDNA was cloned as a BamHI-XhoI fragment in FastBac HTb (Clontech), adding a His_6 tag to the N terminus and the transfections, and virus amplifications and protein production were performed as described in the manufacturer's protocols. The protein was purified out of 100 ml of Sf21 cells 48 h after infection (m.o.i. = 5). For protein purification from *E. coli*, the POT1 and POT1$^{\Delta OB}$ cDNAs were cloned in the BamHI and XhoI sites of pGEX-4T2 (Amersham Biosciences), resulting in N-terminal GST fusions. GST fusion proteins were purified on glutathione beads as directed by the manufacturer. After purification, the protein was dialyzed against 20 mM Hepes, pH 7.9, 500 mM KCl (150 mM for GST fusions), and 20% glycerol, flash-frozen in liquid nitrogen, and stored in aliquots at −80 °C. The binding affinity of POT1 declined 3–5-fold over a period of 2 weeks and was stable afterward.

Oligonucleotides and Probe Labeling—All oligonucleotides were obtained from Genelink as gel-purified 50-mers with indicated random and telomeric sequences (see Table I). Oligonucleotides were labeled at the 5′ end with T4 polynucleotide kinase (New England Biolabs) and [γ-^{32}P]ATP (3000 Ci/mmol, PerkinElmer Life Sciences) and purified through a Sephadex G50 column in 10 mM Tris-Hcl, pH 7.5, 100 mM NaCl, 1 mM EDTA, 0.1% SDS. The labeled oligos were extracted with phenol/chloroform/isoamyl alcohol, precipitated with 0.2 M NaAc, pH 5.5, and 2 volumes of EtOH (−20 °C overnight) and dissolved in 10 mM Tris, pH 8.0.

Band Shift Assays—The binding reactions were performed in 20 μl of the following buffer: 20 mM glycine-NaOH, pH 9.0, 0.1 mM dithiothreitol, 2% glycerol, 50 ng of β-casein, 0.5 μg of sonicated and denatured *E. coli* DNA (mean size ~400 nt), and 0.25 nM probe. The protein was added last, and the binding reaction was incubated for 30 min at room temperature. For direct binding, a range of 0.1–5 μg of protein was used. For competition assays, 0.5–1 μg of protein was added to the reactions containing up to 100-fold molar excess of unlabeled competitor oligonucleotide. Electrophoresis was performed in 0.6% agarose gels run in 0.1× Tris borate EDTA. The gels were run for 40 min at 160 V, dried on Whatman DE81 paper at 80 °C and exposed on phosphorimaging screens. Quantitation was performed using the ImageQuant software.

TABLE I
Sequences of POT1 binding substrates

Name	Sequence (5′ to 3′)[a]
tel-1	A-B-C-D-E-F-**(TTAGGG)**
tel-2	A-B-C-D-E-**(TTAGGG)**$_2$
tel-3	A-B-C-D-**(TTAGGG)**$_3$
tel-4	A-B-C-**(TTAGGG)**$_4$
tel-5	A-B-**(TTAGGG)**$_5$
int-5	A-**(TTAGGG)**$_5$-B
int-2	A-K-**(TTAGGG)**$_2$-L-B
int-MBS	A-H-**(TAGGGTTAG)**-J-B
5′-5	**(TTAGGG)**$_5$-A-B
MBS	A-B-C-D-E-G-**TAGGGTTAG**
MBS+1	A-B-C-D-E-C**CTTAGGGTTAG**
Permut1	A-B-C-D-E-G-**AGGGTTAGG**
Permut2	A-B-C-D-E-G-**GGGTTAGGG**
Permut3	A-B-C-D-E-G-**GGTTAGGGT**
Permut4	A-B-C-D-E-G-**GTTAGGGTT**
Permut5	A-B-C-D-E-G-**TTAGGGTTA**
MBS-T1A	A-B-C-D-E-G-**CAGGGTTAG**
MBS-A2T	A-B-C-D-E-G-**TTGGGTTAG**
MBS-G3C	A-B-C-D-E-G-**TACGGTTAG**
MBS-G4C	A-B-C-D-E-G-**TAGCGTTAG**
MBS-G5C	A-B-C-D-E-G-**TAGGCTTAG**
MBS-T6A	A-B-C-D-E-G-**TAGGGATAG**
MBS-T7A	A-B-C-D-E-G-**TAGGGTAAG**
MBS-A8T	A-B-C-D-E-G-**TAGGGTTTG**
MBS-G9C	A-B-C-D-E-G-**TAGGGTTAC**
MBS-G9A	A-B-C-D-E-G-**TAGGGTTAA**

[a] Non-telomeric portions of the oligonucleotides are: A, GCAAGCTTTA; B, CCGATACAGC; C, CGCTCA; D, GACTGA; E, CTCGAA; F, GAACTC; G, CAC; H, GAACTCGATCC; J, CACTGCAACT; K, GAACTCGAC; L, CACTGCAAC.

RESULTS

Purification of POT1 and Band Shift Assay Conditions—POT1 was expressed in insect cells (His-tagged) or in *E. coli* (GST-tagged) and purified on affinity matrix (Fig. 1, *A* and *B*). Band shift assays were performed with end-labeled single-stranded DNA probes (Table I). All binding substrates were kept at a constant length of 50 nt by changing the length of the random sequence at the 5′ end of the probes to compensate for variations in the length of the telomeric sequences. POT1 did not bind to any of the random subsequences (see below). Various conditions for POT1 binding to TTAGGG repeat probes were tested, and optimal binding occurred at pH 9.0 in the presence of casein and 0.5 μg of single-stranded sheared *E. coli* DNA competitor (data not shown). Reactions were incubated for 30 min at room temperature and fractionated on 0.6% agarose gels in 0.1× Tris borate EDTA (see "Experimental Procedures").

Purified Bac-POT1 or GST-POT1 (Fig. 1*A*) was able to shift a probe containing two, three, or five TTAGGG repeats equally well (Figs. 1, *B* and *C*, and 2*A*). For most preparations of protein, the K_{app} was as low as 20 nM as determined from the concentration of protein required to bind 50% of the probe. To determine whether the N-terminal OB fold was required for the *in vitro* DNA binding activity of POT1, we tested a truncated protein (POT1$^{\Delta OB}$) lacking the first 126 amino acids of the POT1 open reading frame, including the amino acids that make up the putative OB fold. This deletion mutant had no detectable DNA binding activity under our assay conditions (Fig. 1, *C* and *D*). From protein titration experiments, we estimated that POT1$^{\Delta OB}$ was at least 30-fold less active than the full-length POT1 protein (Fig. 1*D* and data not shown), indicating that the N terminus of POT1 is essential for DNA binding *in vitro*. No binding was detected with POT1$^{\Delta OB}$ on probes presenting two (Fig. 1*C*) or five (Fig. 1*D*) TTAGGG repeats.

POT1 Binds [TTAGGG]$_{n \geq 2}$ but Not a Single TTAGGG Site at a 3′ End—In initial experiments, we found that POT1 bound probes with two or five TTAGGG repeats at a 3′ end (tel-2 or

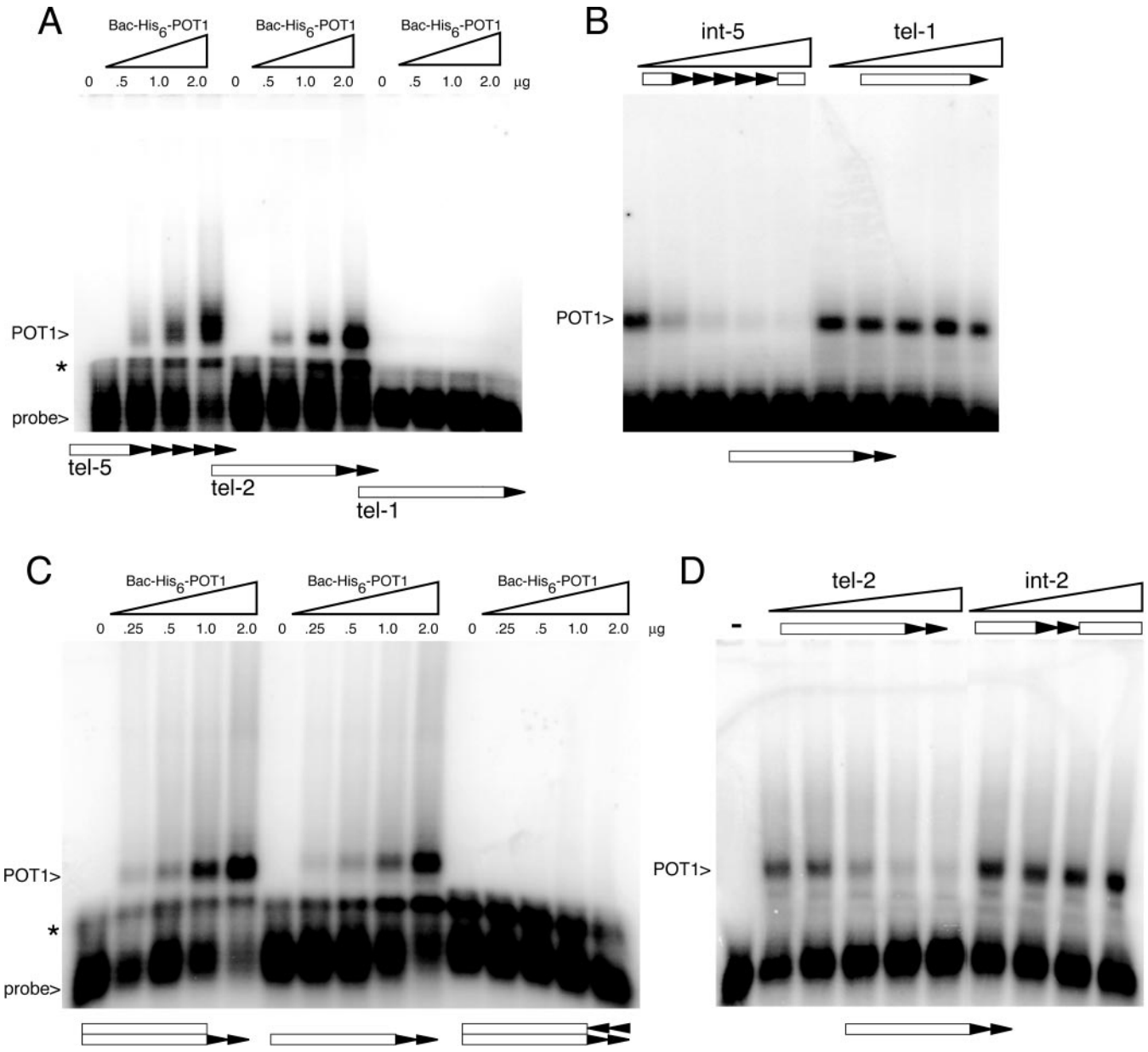


FIG. 2. **Binding of POT1 to TTAGGG sequences of various lengths.** *A*, direct binding of POT1 to five (tel-5), two (tel-2), or one (tel-1) TTAGGG repeats placed at the 3′ end. Protein amounts in each binding reaction are indicated above the lanes. *B*, competition of int-5 or tel-1 for binding of POT1 to labeled tel-2. Molar excess of competitor DNAs, 1-, 3-, 6-, and 9-fold, is shown. *C*, direct binding of POT1 to two TTAGGG repeats present at the 3′ end of a single-stranded oligonucleotide (*middle*), to a probe with double-stranded random sequence (*left*), and to a fully double-stranded probe (*right*). *D*, competition of tel-2 or int-2 for binding of POT1 to labeled tel-2. Molar excess of cold competitor DNAs, 1-, 3-, 6-, and 9-fold, is shown. For all panels, the structure of the probes is indicated below the gels, and the structure of the competitor DNAs is shown above. The positions of the POT1 complex and the probes are indicated. The *asterisk* indicates a nonspecific binding activity.

tel-5) equally well, whereas a probe with only one 3′-terminal TTAGGG sequence (tel-1) was not bound (Fig. 2). This result was obtained both by direct binding assays and by monitoring the competition of various DNAs for the binding of POT1 to tel-2 (Fig. 2, *A* and *B*; Table II). Thus, the minimal binding site for POT1 is longer than 6 nt (one repeat) and shorter than or equal to 12 nt (two repeats).

Annealing the tel-2 probe to its complementary oligonucleotide, thereby rendering the telomeric part double-stranded, abolished POT1 binding (Fig. 2*C*). In contrast, a probe having only the nontelomeric portion in double-stranded form bound as well as the purely single-stranded probe (Fig. 2*C*). These findings are consistent with previously published results (6), indicating that POT1 binds only to single-stranded telomeric sequences.

We subsequently used competition for POT1 binding to tel-2 to determine the ability of POT1 to interact with various substrates. We placed the telomeric repeats internally in the oligonucleotide flanked by random sequences (see Table I). An oligonucleotide with five repeats placed internally (int-5) competed efficiently for POT1 binding to tel-2 (Fig. 2*B*). The competition efficiency, as measured by the molar excess of the competitor required for 50% reduction in POT1 binding to the labeled tel-2 probe, was similar when the five telomeric repeats were placed at the 3′ end of the binding substrate (tel-5) or internally (int-5) (Table II, 1.2- *versus* 1.7-fold).

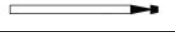
However, with substrates containing only two repeats, there was a significant difference when the telomeric sequence was placed internally; int-2, containing two internal TTAGGG repeats, did not compete for POT1 binding to the tel-2 probe (Fig. 2*D*), whereas tel-2 itself (carrying the TTAGGG repeats at the 3′ end) showed the expected competition (50% reduction in

TABLE II
Preference of POT1 for 3′-terminal position of short but not long repeal arrays

name	structure[a]	telomeric sequence	direct binding[b]	molar excess at 50% competition[c]
tel-1		TTAGGG-3′	-	>100
tel-2		$[TTAGGG]_2$-3′	+	1.5
tel-5		$[TTAGGG]_5$-3′	+	1.7
int-MBS		TAGGGTTAG	-	>100
int-2		$[TTAGGG]_2$	-	>100
int-5		$[TTAGGG]_5$	+	1.2
5′-5		5′-$[TTAGGG]_5$	+	2.9

[a] Random sequences are represented by open boxes, telomeric repeats by arrowheads.
[b] Direct binding was tested with 0.1–5 μg of POT1 protein.
[c] Competition experiments were done with tel-2 as a probe. The values are derived from 2–3 experiments.

TABLE III
POT1 binds to a single permutation of a telomeric nonamer

name	structure[a]	permutation	direct binding[b]	molar excess at 50% competition[c]
MBS+1		TTAGGGTTAG	+	n.d.
MBS		TAGGGTTAG	+	2.7
Permut1		AGGGTTAGG	-	>100
Permut2		GGGTTAGGG	-	>100
Permut3		GGTTAGGGT	-	>100
Permut4		GTTAGGGTT	-	>100
Permut5		TTAGGGTTA	-	>100

[a] Random sequences are represented by open boxes and telomeric repeats by arrowheads.
[b] Direct binding was tested with 0.1–5 μg of POT1 protein.
[c] Competition experiments were done with tel-2 as a probe. The values are derived from 2–3 experiments. n.d., not determined.

binding upon addition of 1.5-fold molar excess of cold probe) (Table II). This result was the first indication that, on short substrates, POT1 displays a preference for binding sites present at a free 3′ end (more than a 100-fold for substrates with two TTAGGG repeats) (Table II). Because no preference for the 3′ end was detected with substrates containing five TTAGGG repeats (Table II; Fig. 2, *A* and *B*), we concluded that a free 3′ end is not strictly required for POT1 binding, but the binding to short sites is facilitated when they are present at a 3′ end. Placing the TTAGGG repeats at the 5′ end did not significantly affect the binding either, as a substrate with five TTAGGG repeats placed at the 5′ end (5′-5) competed well for binding to the tel-2 probe (molar excess = 2.9) (Table II).

The Minimal POT1 Binding Site—A nonamer binding site for POT1 was identified by testing oligonucleotides with permutations of one and a half TTAGGG repeats positioned at a free 3′ end (Table III). Of all possible permutations, only one (the sequence 5′-TAGGGTTAG-3′) was capable of competing efficiently for POT1 (Fig. 3, *A–C*; Table III). The molar excess required for 50% competition was about 2-fold more than for tel-2 itself (Tables II and III). Direct binding assays confirmed the binding of POT1 to the TAGGGTTAG sequence and the lack of binding to other permutations (Fig. 3*B*). Omitting the first T of the TAGGGTTAG site reduced its ability to compete for tel-2 by more than 50-fold (see Table III). In contrast, the addition of one T residue 5′ (resulting in TTAGGGTTAG) did not improve binding significantly as observed by direct binding (data not shown). These data indicated that the sequence TAGGGTTAG represented the minimal binding site (MBS) sequence for POT1. By direct binding, there was no significant difference in affinity between tel-2 and MBS (Fig. 3*C*).

To determine whether POT1 needed a 3′ end to bind to MBS, we placed this sequence internally in a 50-nt probe flanked 5′ and 3′ by random sequences. As was the case for the probe with two internal TTAGGG repeats, POT1 did not bind to the internally placed MBS either by direct binding or by competition assay (Fig. 3*D*). Thus, the lack of binding to short sequences ($\leq$12 nt) at internal sites is not due to the particular permutation used. We concluded that the sequence TAGGGTTAG represents the minimal binding for POT1 when present at the 3′ end of the DNA.

POT1 Binds with High Sequence Specificity—The minimal binding site provided a simplified context for the analysis of the sequence specificity of POT1. A-T, T-A, G-C, or C-G transversions were introduced at each of the nine positions in MBS, and the resulting MBS variants were tested by competition for POT1 binding to tel-2 (Fig. 4*A*, Table IV). High affinity POT1 binding occurred with only one MBS variant, the sequence 5′-TAGGGTAAG-3′, a change of T to A at position 7 (referred to as MBS-T7A) (Fig. 4*B*, Table IV). Based on competition assays, POT1 bound to MBS-T7A with an affinity about 2-fold lower than to MBS (molar excess of 4 for MBS-T7A compared with 2.7 for MBS, Table IV). By direct binding, POT1 also bound to MBS and MBS-T7A (Fig. 4*B*). The sequence TAGG**C**TTAG (MBS-G5C), containing a G to C change at position 5 (in bold), was also tolerated but with a significant loss of affinity compared with MBS (Table IV). All of the other changes greatly reduced the binding of POT1 to MBS (Fig. 4*A* and Table IV).

As a test for the relevance of the sequence specificity displayed by POT1 when binding to MBS, we tested the binding of POT1 to repeats of the sequence GGTTAC, a change that should abolish binding based on the lack of binding of POT1 to MBS-G3C and MBS-G9C (TACGGTTAG and TAGGGTTAC, respectively) (Table IV). Consistent with the results of the MBS studies, the substrate containing five tandem GGTTAC repeats could not compete effectively with tel-2 for POT1 binding (Fig. 4*C*). However, a DNA substrate with the same G-C change only in the 3′-terminal repeat (resulting in $[GGTTAG]_4$GGTTAC-3′) was an effective competitor (Fig. 4*C*). This result exemplifies the internal binding activity of POT1 and shows that changes in the telomeric repeat must occur at multiple binding sites to significantly affect binding activity. We conclude that POT1 displays high sequence specificity in the context of the minimal binding site.

DISCUSSION

The characterization of the DNA binding properties of POT1 is important for understanding its function at telomeres. In this study, we have defined the minimal binding site for human POT1 and determined the nucleotides important for binding to this sequence. The sequence 5′-TAGGGTTAG-3′ constitutes the only nonamer permutation that bound POT1 with comparable affinity to longer telomeric repeat arrays. The same sequence placed internally lost its affinity for POT1, suggesting a preference for 3′ end binding on short sites. However, human and mouse telomeres have been shown to contain long tandem arrays of TTAGGG repeats, which can contain 20–30 overlapping MBS sequences. According to our analysis, POT1 binds very well to such long TTAGGG repeat arrays even when they are not directly at a 3′ terminus. This predicts that much of the 3′ overhang and most of the telomeric D-loop could be coated by POT1.

The POT1 site is at least 9 nt in length, which is long

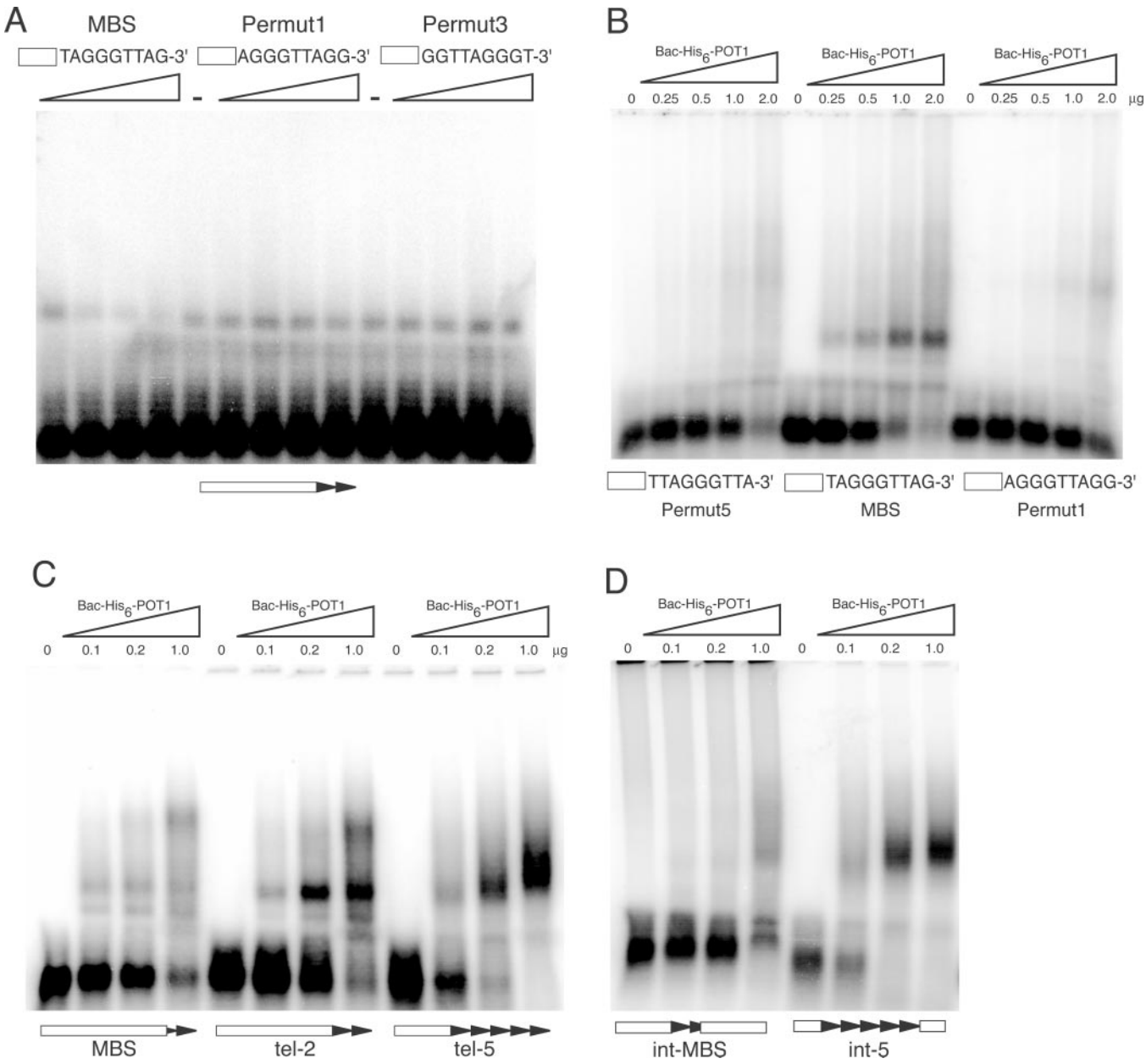


FIG. 3. **Identification of the minimal binding site.** *A*, competition of different permutations of telomeric nonamers for binding of POT1 to tel-2. Probes and competitor DNAs are indicated below and above the gel. Competitor to probe ratios were 1-, 3-, 6-, and 9-fold. *B*, direct binding assay with three nonamer permutations. *C*, comparison of direct binding of POT1 to MBS, tel2, and tel-5. *D*, direct binding of POT1 to probes with the MBS internally (*int-MBS*) or five TTAGGG repeats internally (*int-5*). In *panels B*, *C*, and *D*, the lanes with 1–2 μg of POT1 show smearing of the probe upward in the gel because of a slight salt effect.

considering that OB folds usually interact with short sites (2–5 nt). One possibility is that POT1 either binds as a dimer or contains multiple OB folds each recognizing part of the MBS. As OB folds lack strong primary sequence signatures, it is not excluded that POT1 contains two or even three of these motifs. Another possibility is that POT1 acts similarly to Cdc13, which interacts with an 11-nt site using a single OB fold (13, 14). In this case, the most 5′ GTGT sequence constitutes a "hot spot" of binding affinity. Additional weaker interactions, in part facilitated by a 30-amino acid loop, extend the preferred binding substrate to 11 nt (14, 15). A final possibility is that human POT1 recognizes the telomeric TAGGGTTAG site through both protein DNA interactions and DNA-DNA interactions, a mode of substrate recognition exemplified by the TEBPα/β complex (8). The DNA-DNA interactions in the POT1-DNA complex could involve stacking interactions and higher order folding of the single-stranded DNA. Thus, the MBS sequence could be a good POT1 substrate both because it provides the optimal base and backbone contacts and has the ability to adopt the proper configuration for POT1 binding. The resolution of the crystal structure of human POT1 bound to the DNA will shed light on these issues. Similarly, the role of the 3′ terminus in the formation of the POT1-DNA complex could involve protein-DNA contacts, as in the TEBPα/β complex, and/or DNA-DNA contacts, which could explain the preference of POT1 for its minimal binding site at a 3′ end.

We have shown previously that an N-terminal deletion mutant, POT1$^{\Delta OB}$, which lacks the DNA binding domain, can still localize to telomeres (10). This indicates that the DNA binding function is not necessary for targeting POT1 to telomeres *per se*. If POT1 is primarily recruited to telomeres by protein interaction through the TRF1 complex, it could associate with its cognate DNA binding site as a secondary event. Therefore, the specific positioning of POT1 at chromosome

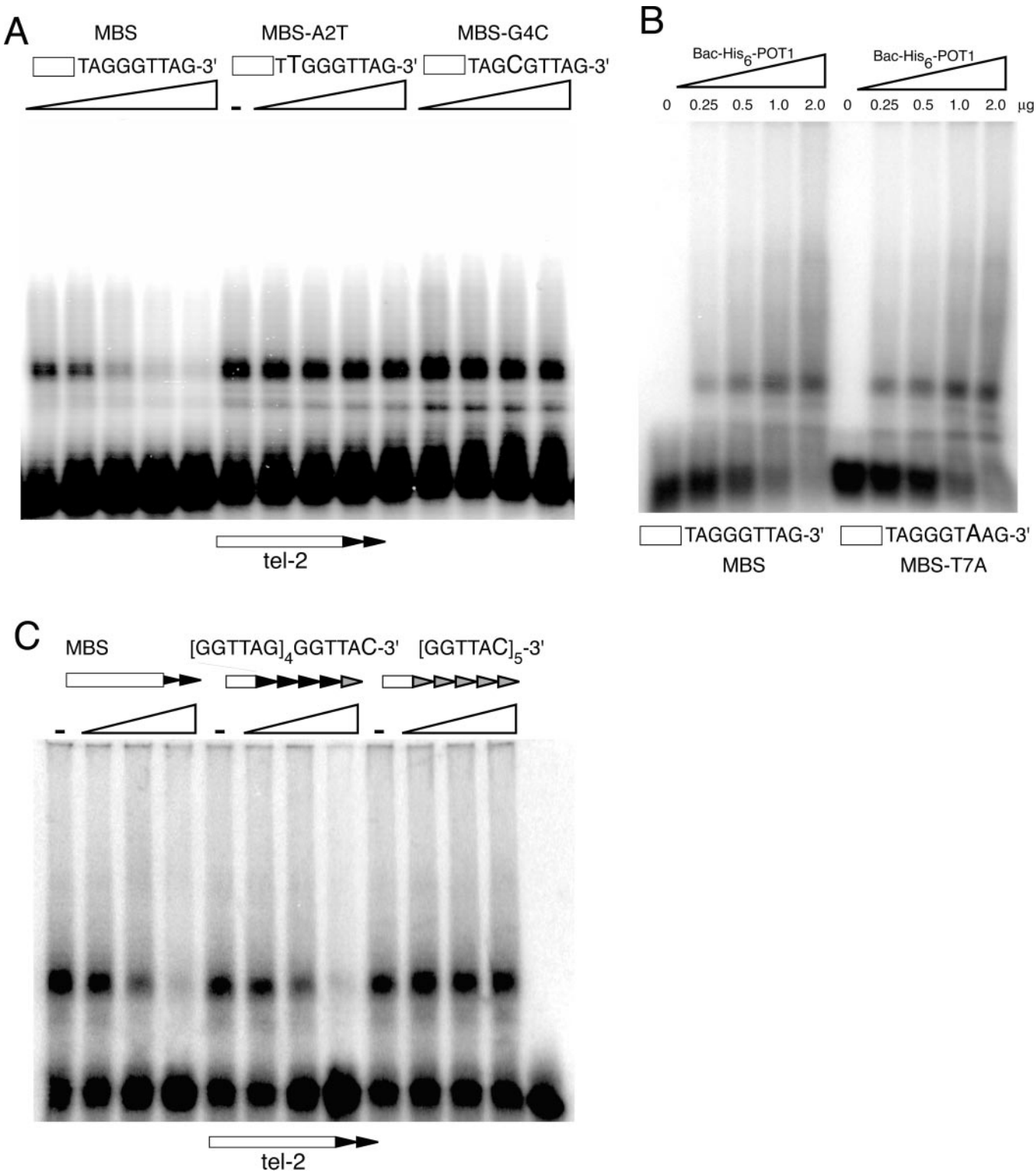


FIG. 4. **Sequence specificity of POT1.** *A*, competition of MBS variants for binding of POT1 to tel-2. Molar excess of competitor is 1-, 3-, 6-, and 9-fold. *B*, direct binding to MBS (*left*) and MBS with the T7 to A transversion (*right*). *C*, competition assay with MBS and two different substrates with a mutated $[GGTTAG]_5$ array at the 3′ end. The probe is tel-2. Molar excess of competitor is 1-, 3-, and 9-fold.

ends would rely on two highly specific events, its interaction with the TRF1 complex and its ability to engage TAGGGTTAG sites. In this context, recognition of the 3′ terminus, which is a third hallmark of the telomere, may not be an important factor.

POT1 discriminates effectively between MBS and variant sites that have single nucleotide changes. This ability of POT1 to distinguish between telomeric repeats and closely related sequences could be highly relevant to its biological function in vertebrates. Many vertebrates contain sequences closely resembling telomeric DNA at chromosome internal sites (16). In some cases, these sequences are highly repetitive, forming a major class of mini-satellite DNA (17). The sequence specificity of POT1 may be sufficient to prevent binding to these pseudotelomeric repeats. This distinction is especially important because the TRF1 complex can be loaded on the chromosome-

TABLE IV
Sequence specificity of POT1 when binding to MBS

Name	3′-Terminal sequence	Direct binding[a]	Molar excess at 50% competition[b]
MBS	TAGGGTTAG-3′	+	2.7
MBS-T1C	**C**AGGGTTAG-3′	−	>100
MBS-A2T	T**T**GGGTTAG-3′	−	>100
MBS-G3C	TA**C**GGTTAG-3′	−	>100
MBS-G4C	TAG**C**GTTAG-3′	−	>100
MBS-G5C	TAGG**C**TTAG-3′	+/−	7
MBS-T6A	TAGGG**A**TAG-3′	−	>100
MBS-T7A	TAGGGT**A**AG-3′	+	4
MBS-A8T	TAGGGTTT**T**G-3′	−	>100
MBS-G9C	TAGGGTTA**C**-3′	−	>100
MBS-G9A	TAGGGTTA**A**-3′	−	>100

[a] Direct binding was tested with 0.1–5 μg of POT1 protein.
[b] Competition experiments were done with tel-2 as a probe. The values are derived rom 2–3 experiments.

internal pseudo-telomeric repeats (18), potentially recruiting POT1 to these sites. Thus, the high sequence specificity displayed by POT1 may not be important for actual targeting and localization of the protein to telomeres but may be important for preventing POT1 from binding to nontelomeric sites in the genome during S-phase.

In this context, it would be interesting to compare the affinities of POT1 and replication protein A (RPA), the single-stranded DNA-binding protein involved in DNA replication (19), to telomeric and nontelomeric probes. It is expected that the affinity of RPA for the telomeric sequence is lower than that of POT1, which would result in the exclusion of RPA by POT1 from the telomeric overhang. On the other hand, because RPA does not display sequence specificity on single-stranded DNA, it could effectively compete off POT1 from nontelomeric sites, which may be important during DNA replication. In addition, RPA would presumably have to displace POT1 from telomeric DNA during the replication of telomeres. It is likely that specific regulatory pathways function to facilitate this exchange.

Acknowledgments—We thank the de Lange laboratory for comments on the manuscript and D. Hockemeyer for helpful discussion.

REFERENCES

1. Griffith, J. D., Comeau, L., Rosenfield, S., Stansel, R. M., Bianchi, A., Moss, H., and de Lange, T. (1999) *Cell* **97,** 503–514
2. Bianchi, A., Stansel, R. M., Fairall, L., Griffith, J. D., Rhodes, D., and de Lange, T. (1999) *EMBO J.* **18,** 5735–5744
3. Bianchi, A., Smith, S., Chong, L., Elias, P., and de Lange, T. (1997) *EMBO J.* **16,** 1785–1794
4. Smogorzewska, A., and de Lange, T. (2004) *Annu. Rev. Biochem*., in press
5. van Steensel, B., and de Lange, T. (1997) *Nature* **385,** 740–743
6. Baumann, P., and Cech, T. R. (2001) *Science* **292,** 1171–1175
7. Baumann, P., Podell, E., and Cech, T. R. (2002) *Mol. Cell. Biol.* **22,** 8079–8087
8. Horvath, M. P., Schweiker, V. L., Bevilacqua, J. M., Ruggles, J. A., and Schultz, S. C. (1998) *Cell* **95,** 963–974
9. Lei, M., Baumann, P., and Cech, T. R. (2002) *Biochemistry* **41,** 14560–14568
10. Loayza, D., and de Lange, T. (2003) *Nature* **424,** 1013–1018
11. Colgin, L. M., Baran, K., Baumann, P., Cech, T. R., and Reddel, R. R. (2003) *Curr. Biol.* **13,** 942–946
12. Chandra, A., Hughes, T. R., Nugent, C. I., and Lundblad, V. (2001) *Genes Dev.* **15,** 404–414
13. Hughes, T. R., Weilbaecher, R. G., Walterscheid, M., and Lundblad, V. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 6457–6462
14. Mitton-Fry, R. M., Anderson, E. M., Hughes, T. R., Lundblad, V., and Wuttke, D. S. (2002) *Science* **296,** 145–147
15. Anderson, E. M., Halsey, W. A., and Wuttke, D. S. (2003) *Biochemistry* **42,** 3751–3758
16. Meyne, J., Baker, R. J., Hobart, H. H., Hsu, T. C., Ryder, O. A., Ward, O. G., Wiley, J. E., Wurster-Hill, D. H., Yates, T. L., and Moyzis, R. K. (1990) *Chromosoma* (*Berl.*) **99,** 3–10
17. Southern, E. M. (1970) *Nature* **227,** 794–798
18. Smogorzewska, A., van Steensel, B., Bianchi, A., Oelmann, S., Schaefer, M. R., Schnapp, G., and de Lange, T. (2000) *Mol. Cell. Biol.* **20,** 1659–1668
19. Wold, M. S. (1997) *Annu. Rev. Biochem.* **66,** 61–92