DNA Binding Affinity and Sequence Permutation Preference of the Telomere Protein from Euplotes crassus †

ARTICLE *in* BIOCHEMISTRY · AUGUST 2006

Impact Factor: 3.02 · DOI: 10.1021/bi060388w · Source: PubMed

5 AUTHORS, INCLUDING:

Martin P Horvath
University of Utah
26 PUBLICATIONS **530** CITATIONS

Published in final edited form as:
Biochemistry. 2006 July 18; 45(28): 8628–8638. doi:10.1021/bi060388w.

DNA-binding affinity and sequence permutation preference of the telomere protein from *Euplotes crassus*

Takahito Suzuki, **Margaret McKenzie**, **Elizabeth Ott**, **Olesya Ilkun**, and **Martin P. Horvath***
Department of Biology, University of Utah, 257 S 1400 E, Salt Lake City Utah 84112-0840, USA

Abstract

Telomere end binding proteins from diverse organisms use various forms of an ancient protein structure to recognize and bind with single strand DNA found at the ends of telomeres. To further understand the biochemistry and evolution of these proteins we have characterized the DNA-binding properties of the telomere end binding protein from *Euplotes crassus* (EcTEBP). EcTEBP and its predicted amino-terminal DNA-binding domain, EcTEBP-N, were expressed in *E. coli* and purified. Each protein formed stoichiometric (1:1) complexes with single strand DNA oligos derived from the precisely defined d(TTTTGGGGTTTTGG) sequence found at DNA termini in *Euplotes*. Dissociation constants for DNA•EcTEBP and DNA•EcTEBP-N were comparable, with $K_{D\text{-}DNA}$ = 38 ± 2 nM for the full-length protein and $K_{D\text{-}DNA}$ = 60 ± 4 nM for the N-terminal domain, indicating that the N-terminal domain retains high affinity for DNA even in the absence of potentially stabilizing moieties located in the C-terminal domain. Rate constants for DNA association and DNA dissociation corroborated a slightly improved DNA binding performance for the full-length protein (k_a = 45 ± 4 μM^{-1} s^{-1}, k_d = 0.10 ± 0.02 s^{-1}) relative to the N-terminal domain (k_a = 18 ± 1 μM^{-1} s^{-1}, k_d = 0.15 ± 0.01 s^{-1}). Equilibrium dissociation constants measured for sequence permutations of the telomere repeat spanned a 55 – 1400 nM range, with EcTEBP and EcTEBP-N binding most tightly to d(TTGGGGTTTTGG) — the sequence corresponding with that of mature DNA termini. Additionally, competition experiments showed that EcTEBP recognizes and binds the telomere-derived 14-nucleotide DNA in preference to shorter 5′ -truncation variants. Compared with multi-subunit complexes assembled with telomere single strand DNA from *Oxytricha nova*, our results highlight the relative simplicity of the *Euplotes crassus* system where a telomere end binding protein has biochemical properties indicating one protein subunit caps the single strand DNA.

Keywords

DNA-protein complexes; telomere structure and function; telomere biology; telomere evolution; *Moneuplotes crassus*; *Sterkiella nova*

INTRODUCTION

Chromosomes in eukaryotes are typically capped by specialized nucleoprotein structures called telomeres that protect DNA ends and distinguish natural ends from breaks caused by DNA damage. Telomere DNA consists of multiple tandem repeats of a short sequence such as TTGGGG in *Tetrahymena thermophila* (1), TTTTGGGG in spirotrichs (2), and TTAGGG in vertebrates, molds and fungi (3-8) (reviewed in ref (9)). The T and G rich strand terminates with a 3′ OH group and is longer than its complimentary A and C rich strand (2,10,11). Consequently, a conserved feature of telomere ends is a stretch of single strand DNA.

*Corresponding author: Martin P. Horvath Assistant Professor of Biology University of Utah 257 S 1400 E Salt Lake City UT 84112-0840 USA e-mail: horvath@biology.utah.edu phone: +1 801-587-7854 fax: +1 801-581-2174

Telomere single strand DNA binds with telomere end binding proteins such as CDC13 in budding yeast (12), POT1 found in a broad spectrum of organisms including fission yeast, plants, and vertebrates (13), and telomere binding proteins from ciliates (14-16). Deletion of the gene encoding POT1 in fission yeast (13) and RNA interference studies in *Stylonychia lemnae* (17) showed that these telomere specific proteins are essential for cell viability and normal telomere function. Related by weak similarity in amino acid sequence, members of this family each contain one or more copies of a protein folding motif called the oligonucleotide/ oligosaccharide/ oligopeptide binding fold (OB fold) that was first described by Murzin for non-telomeric proteins including the single strand DNA binding protein from filamentous phage (18,19). X-ray and NMR structures of the telomere-associated OB fold-containing proteins in complex with telomere-derived DNA fragments revealed the use of OB folds in establishing sequence-specific contacts with DNA (20-23). In the case of the alpha telomere protein from *Oxytricha nova* (*Sterkiella nova*), an OB fold is also used to establish protein-protein interactions with a second protein subunit required for high-affinity DNA binding (20,24,25).

Ciliated protozoa such as *Oxytricha nova* and *Euplotes crassus* (*Moneuplotes crassus*) have a remarkable genome structure with large and transcriptionally silent diploid chromosomes located in germ-line micronuclei, and small gene-sized pieces of DNA that are amplified to high copy numbers in macronuclei (26-28). Each end of the macronuclear DNA is capped by a precisely defined short telomere repeat sequence and a 3′ -terminal single strand DNA, also precisely defined in length and sequence, that binds with one or two telomere end binding proteins. In *Oxytricha nova*, the single strand DNA forms a ternary complex with one alpha subunit and one beta subunit (20,24,29,30). This DNA can also bind one or two alpha subunits in the absence of beta (24,31,32). In *Euplotes crassus*, the single strand DNA forms a complex with a telomere end binding protein (33,34) and may also interact with a second telomere protein homologue during DNA synthesis (35,36).

Figure 1 shows a phylogenetic relationship deduced from amino acid sequences of the ciliate telomere end binding proteins and telomere binding protein homologues. The dendrogram agrees with those inferred from more complete analyses of actin and ribosomal RNA genes (37-39) in that the *Euplotes crassus* lineage is deeply branched relative to that of the *Oxytricha* and *Stylonychia* species. Given the large evolutionary distance separating *Oxytricha nova* and *Euplotes crasssus*, comparison of their telomere end binding proteins should enable a deeper understanding of the evolution of this important protein and provide insight to common themes and diversity for the biochemistry of protecting telomere single strand DNA.

Differences between *Oxytricha nova* and *Euplotes crassus* indicate their telomeres, although related, may adopt substantially different structures. Both of these ciliated protozoa use a d (T_4G_4) telomere repeat sequence, but DNA length is species specific. Double strand telomere DNA consists of 20 basepairs in *Oxytricha* and 28 basepairs in *Euplotes* (2). Single strand, 3′ -terminal telomere DNA comprises 16 nucleotides, two full repeats, in *Oxytricha* and 14 nucleotides, d(TTTTGGGGTTTTGG), in *Euplotes* (2).

Possibly related with these DNA length differences, telomere protein composition is also distinct. A homologue for the *Oxytricha nova* beta protein has not been found in *Euplotes crassus* (40) or, for that matter, in many other organisms including *Tetrahymena thermophila*, yeast, plants and vertebrates, even though proteins with sequence and structural similarity to the DNA-binding portions of the *Oxytricha nova* alpha protein appear to be widespread (13,40). The current view holds that alpha and beta telomere proteins work cooperatively to cap telomeres in *Oxytricha* and *Stylonychia* species (17,20,24,25) while an alpha-like protein works more independently to protect telomere single strand DNA in *Euplotes* (33). Of course a completely unrelated protein that serves a function similar to that

provided by beta may exist in *Euplotes*, and this possibility of a beta analogue is indicated with the annotation (*b*) in Figure 1.

As with nucleoprotein complexes in *Oxytricha nova* (14,16), telomere-associated protein in *Euplotes crassus* protects DNA from nucleases and remains attached to DNA even at high salt concentrations (33). *Euplotes crassus* genes encoding a basic 446-aa telomere protein and an acidic 460-aa telomere protein homologue were identified on the basis of sequence similarity with a gene for the *Oxytricha nova* alpha telomere protein (40). The 460-aa telomere protein homologue (EcRTP in Figure 1) is transiently expressed coincident with DNA synthesis and is thought to be involved with telomere replication (35). The 446-aa protein studied here (EcTEBP in Figure 1) associates tightly with telomere DNA, and previous characterization of its DNA-binding properties suggested that this protein plays an important role in establishing the precise length of telomere DNA in *Euplotes crassus* (34).

We have cloned expression vectors encoding the *Euplotes crassus* telomere end binding protein (EcTEBP) and its predicted N-terminal DNA-binding domain (EcTEBP-N) that enabled purification of milligram quantities of both full-length and truncated forms of the protein. The DNA-binding properties of recombinant EcTEBP and EcTEBP-N were, to some degree, similar to those observed previously for protein isolated directly from *Euplotes crassus* (34). Because of improved yield and solubility of the bacterially expressed proteins, we could for the first time determine binding stoichiometries, equilibrium dissociation constants and kinetic rate constants and further evaluate how DNA length and sequence register relate to DNA-binding affinity. The DNA-binding properties measured here are consistent with the idea proposed earlier (34) that the telomere end binding protein plays a critical role in determining an exact length and sequence for the single strand extension terminating DNA molecules in *Euplotes* (41,42). Compared with multi-protein complexes assembled with telomere single strand DNA in *Oxytricha nova*, our results highlight the relative simplicity of the *Euplotes crassus* system where a single telomere end binding protein caps the single strand DNA.

MATERIALS AND METHODS

Amino acid sequence alignment and phylogenetic analysis

Amino acid sequences were aligned using the multiple-sequence option of *Clustal X* (43). The crystal structure of the *Oxytricha nova* alpha telomere protein (20,44) provided a context relevant for interpretation of this sequence alignment and guided modest repositioning of three gaps so as to relocate these at connecting loops as opposed to core regions of the protein. The aligned sequences may be inspected in Figure S1 provided as supplementary data. The tree shown in Figure 1 was tested with bootstrap simulations using maximum likelihood, neighbor joining, and most parsimonious analyses implemented in *PHYLIP* (45).

Cloning, expression, and purification of proteins

DNAs encoding the full-length $EcTEBP_{446}$ (446 aa) and two forms of its N-terminal domain, $EcTEBP\text{-}N_{288}$ (288 aa) and $EcTEBP\text{-}N_{300}$ (300 aa) were prepared by PCR, digested with restriction enzymes and ligated with pET9a (Novagen) vector DNA. The resulting expression vectors were propagated in DH5α, and the inserted DNAs were verified by sequencing.

For expression of protein, BL21(DE3) bacteria harboring the pLysS plasmid were transformed with expression vector DNA, and transformants were selected using 17 μg/mL chloramphenicol and 40 μg/mL kanamycin. Single colony inoculums were grown until cloudy and then subcultured into three 750-mL volumes of either LB or 2xYT media supplemented with 5 mM glucose, 30 mM potassium phosphate pH 7.8 and antibiotics. These 750-mL cultures were contained in 2-L baffle flasks and grown at 37 °C while shaking at 200 rpm.

When the optical density at 600 nm reached 1.4, cultures were slowly cooled to 16 °C before inducing with 0.5 mM IPTG. After 20 hours of continued incubation at 16 °C, bacterial cells were harvested and combined by centrifugation. Cell extracts were prepared by one freeze-thaw cycle followed by sonication.

Proteins were purified using the same procedures described for the *Oxytricha nova* telomere end binding protein alpha subunit (20) and engineered alpha-beta fusion proteins (25). Briefly this involved ammonium sulfate precipitation, two sequential cation-exchange chromatography steps (SP-Sepharose and Mono-S) followed by a size-exlusion chromatography step (Superdex-200). All purification steps were carried out at 4 °C. Enriched fractions were identified by SDS-polyacrylamide electrophoresis. Protein solutions were concentrated to about 10 mg/ml and stored on ice. Detergents were avoided at all steps during cell extraction and protein purification.

DNA oligonucleotides

Cy5-labeled, biotin-tagged, and unmodified DNA oligonucleotides were synthesized by the University of Utah DNA-Protein Core research facility. Cy5 and biotin modifications were attached to the 5′ end of oligonucleotides so that these telomere-derived fragments would retain a 3′ OH group as found in authentic single strand telomere DNA. Oligonucleotides were purified by reverse-phase HPLC in both a trityl group-protected form and following removal of trityl groups. For HPLC purification, acetonitrile gradients were developed over a butyl 10 × 250 mm semi-preparative column (Vydac) with buffer A containing 20 mM triethyl amine ammonium acetate pH 6 (TEAA) and buffer B containing 20% (v/v) 20 mM TEAA pH 6 and 80% (v/v) acetonitrile. Lyophilized DNA was washed repeatedly with water and re-lyophilized to remove trace HPLC solvents. Pure DNA was suspended in solutions containing 5 mM Tris pH 6 and 0.5 mM EDTA. DNA concentrations were determined from absorbance measurements using extinction coefficients at 260 nm calculated by the dinucleotide method (46) or the extinction coefficient of the Cy5 group at 648 nm, ε_{648} = 250,000 M^{-1} cm^{-1}.

Mobility shift assay

Binding reactions were prepared by combining 4 pmol Cy5-d(TTTTGGGGTTTTGG) with different amounts of EcTEBP or EcTEBP-N_{300} ranging from 0 to 500 pmol in 20 μL volumes and allowing the samples to equilibrate for 1 hour at room temperature. The resulting free DNA and DNA-protein species were separated by electrophoresis in an agarose 1×TAE gel for 25 minutes at room temperature, 60 mA and 70 V. Tests with acrylamide gels proved this gel material to be unsuitable since acrylamide quenched the Cy5 fluorescence signal. Fluorescently labeled DNA was detected with a Typhoon fluorescence imager (Amersham).

For pairwise competition experiments, conditions were slightly different. In these binding reactions protein and unmodified DNAs were pre-incubated for 90 min prior to adding Cy5-modified DNA. Reactions continued to equilibrate for 4 hours before analysis by electrophoresis. During the final equilibration phase the concentration of each macromolecule was 2 μM EcTEBP, 2 μM Cy5-d(TTTTGGGGTTTTGG), and differing levels of unmodified DNA which ranged from 0.5 μM — 200 μM. At the higher DNA and protein concentrations employed for these pairwise competition experiments, the complex was less susceptible to disassociation during electrophoresis as indicated by control reactions with no competitor DNA.

Equilibrium analytical ultracentrifugation

EcTEBP and its complex with Cy5-d(TTTTGGGGTTTTGG) were exchanged into binding buffer (0.15 M sodium chloride and 0.05 M Tris pH 7.5) by gel filtration chromatography. Samples were loaded into cells fitted with quartz or sapphire windows and centrifuged in a

Beckman Optima XL-I analytical ultracentrifuge. Samples were maintained at 20 °C during the course of the experiment. Radial absorbance scans were obtained using incident light at 280 nm, 0.001 cm step increments, at rotation speeds ranging from 20,000 – 25,000 rpm. Each scan was the average of 10 replicas. In calculating model curves, the partial specific volume of each protein or protein-DNA complex was set at v = 0.737 based on estimates from amino acid composition (20). The density of binding buffer (ρ) was 1.0044 g/mL as determined experimentally.

Data from multiple absorbance scans were simultaneously analyzed by a program written in *C* (M.P.H.) that employs a nonlinear least-squares algorithm to optimize parameters for molecular weight, M, and A_o in Equation (1),

$$A_{calculated} = A_{offset} + A_o \exp\left[M(1 - v\rho)\,\omega^2 \left(r^2 - r_o^2\right) / (2 \quad R \quad T)\right] \tag{1}$$

where R is 8.3145×10^7 erg mol^{-1} K^{-1}. The A*offset* parameter accounted for small (~5%) absorbance discrepancies in the reference buffer and dilution buffer. Removal of this parameter did not change the main conclusion of the analysis that the 1:1 EcTEBP-DNA complex does not oligomerize.

Uncertainties in the values of fitted parameters were obtained by a Monte Carlo bootstrap method (47). Measured data were sampled with replacement to generate several simulation data sets where any given data point from the original experiment was represented 0, 1, 2, and very rarely 3 times. Fitted parameter values were determined for each simulation data set in order to test how sensitive parameter values are with respect to removal or duplication of randomly selected data points. The sample standard deviation of these parameter values is reported as the uncertainty. This method gives a realistic estimate of uncertainty, comparable to the variance obtained in those cases where it is feasible to repeat an experiment several times in real life.

Fluorescence anisotropy titrations

DNA and protein were diluted in solutions of binding buffer containing 0.05 M Tris adjusted to pH 7.5 and 0.15 M sodium chloride or lithium chloride. Sodium ions stabilize G-quartet DNA structures that readily form for G-rich DNA. Use of lithium chloride in place of sodium chloride for some of the binding studies ensured that results would not be confounded by intra-molecular or inter-molecular DNA association reactions. In the case of DNAs with a low tendency to form G-quartet structures such as d(TTTTGGGGTTTTGG), binding constants were the same when measured in sodium or lithium-containing solutions.

Fluorescence anisotropy data were obtained at an ambient temperature of approximately 23 °C. Fluorescence anisotropy titrations for 14-nucleotide Cy5-d(TTTTGGGGTTTTGG) DNA were performed as previously described (25) using a dedicated single-channel spectrofluorimeter fitted with polarizing filters for incident and emitted light and a 20-nm bandpass filter to minimize stray scattered light. DNA binding reactions for 12-nucleotide telomere repeat permutation DNAs were prepared in triplicate and fluorescence anisotropy values were measured using a SpectraMax micro-titer plate-reader (Molecular Devices) with incident light wavelength of 625 nm and emitted light collected after a 680-nm high-pass filter. The micro-titer plate-reader format was less sensitive but enabled rapid measurement of all data so as to ensure comparable conditions. Signal-to-noise was enhanced through x10-fold signal averaging of multiple readings.

Anisotropy measurements were analyzed with use of a program written in *C* (M.P.H.) to optimize parameters describing the dissociation constant ($K_{D\text{-}DNA}$) and anisotropy characterizing free DNA and the DNA-protein species. The amounts of free protein, free DNA,

and DNA-protein complex that satisfied Equations (2) – (4) were found through numerical methods, an approach equivalent to solving the quadratic equation describing a two-state partition function but more easily extended to other more complex systems (25). Uncertainty in $K_{\text{D-DNA}}$ was determined by a Monte Carlo bootstrap method as described above for the analysis of AUC data.

$$K_{\text{D-DNA}} = [\text{DNA}_{\text{free}}]\,[\text{protein}_{\text{free}}] / \quad [\text{DNA} - \text{protein}] \tag{2}$$

$$[\text{DNA}_{\text{total}}] = [\text{DNA}_{\text{free}}] + [\text{DNA} - \text{protein}] \tag{3}$$

$$[\text{protein}_{\text{total}}] = [\text{protein}_{\text{free}}] + [\text{DNA} - \text{protein}] \tag{4}$$

Kinetics of DNA association and dissociation

DNA binding and dissociation kinetics were analyzed with a BIACORE system in the Center for Biomolecular Interaction Analysis. Biotin-d(TACATTTTGGGGTTTTGG) DNA was captured onto a CM4 sensor chip pre-immobilized with low-density levels of streptavidin. A total of seven response units (RU) were associated with the sensor chip. The 5′ -terminal d (TACA) nucleotides were necessary to obtain a single kinetic phase for DNA association and likely presented the telomere DNA in a manner that avoided steric effects close to the surface of the chip. EcTEBP$_{446}$ and EcTEBP-N$_{1\text{-}288}$ proteins were diluted to starting concentrations of 100 nM and further diluted to obtain protein solutions spanning 0.4 – 100 nM. Each protein solution was tested for DNA binding and release in triplicate. The dilution and running buffer was phosphate buffered saline adjusted to pH 7.4 with added 0.005% p20 (Biacore AB, Uppsala Sweden) and 2 mg/ml bovine serum albumin. The flow rate was 90 μL/min. Signals returned to baseline during the dissociation phase so a regeneration wash was omitted. Data were collected at 25 °C. Response curves were analyzed with a two-compartment binding model comprising the pair of coupled differential equations shown in Equations (5a) and (5b).

$$\begin{aligned} &d[\text{PTN}_{\text{inner}}]/dt= \\ &-k_a[\text{PTN}_{\text{inner}}]([\text{DNA}_{\text{total}}] - [\text{DNA}\cdot\text{PTN}]) + k_d[\text{DNA}\cdot\text{PTN}] + k_{\text{mass}}([\text{PTN}_{\text{total}}] - [\text{PTN}_{\text{inner}}]) \end{aligned} \tag{5a}$$

$$d[\text{DNA}\cdot\text{PTN}]/dt = k_a[\text{PTN}_{\text{inner}}]([\text{DNA}_{\text{total}}] - [\text{DNA}\cdot\text{PTN}]) - k_d[\text{DNA}\cdot\text{PTN}] \tag{5b}$$

This two-compartment model provides a relatively simple way of including diffusion and flow transport effects. In this model, the rate constant k_{mass} describes movement of analyte from an “outer” compartment to an “inner” compartment (48). In Equations (5a) and (5b), protein analyte contained within the inner compartment, PTN$_{\text{inner}}$, binds DNA sites to form a DNA·PTN complex in a step described by rate constant k_a (M^{-1} s^{-1}). The DNA·PTN complex dissociates in a step described by rate constant k_d (s^{-1}). These coupled differential equations were integrated by numerical methods, and values for k_a , k_d, and k_{mass} were adjusted so as to minimize the sum of least-squares residuals comparing the resulting computed curves with measured response curves .

RESULTS

Purification of EcTEBP and EcTEBP-N

The full-length telomere end binding protein from *Euplotes crassus*, EcTEBP$_{446}$, and two versions of its predicted N-terminal domain, EcTEBP-N$_{300}$ and EcTEBP-N$_{288}$*, were expressed under control of a T7 promoter in *E. coli*. To obtain protein in a folded and soluble

*We will distinguish the full-length 446-aa protein from its C-terminal truncated forms by indicating the carboxy-terminal amino acid position in subscript.

form required a carefully controlled culture temperature of 16 °C during protein induction. Each protein was purified essentially to homogeneity using ammonium sulfate precipitation, ion-exchange, and size-exclusion chromatography. During chromatography, temperature continued to be an important parameter, and it was necessary to carry out the ion-exchange steps at +4 °C in order to obtain a reasonable yield of pure protein. Starting with 2.25 liters of bacterial culture, 10 mg of pure EcTEBP and 18 mg of pure EcTEBP-N were obtained. Figure 2 shows a Coomassie-stained SDS-polyacrylamide gel analyzing ~20 μg of pure $EcTEBP_{446}$ and EcTEBP-N288 and indicating the high quality of each preparation. Each protein had the expected apparent molecular weight. Moreover, electrospray ionization mass spectrometry confirmed that each protein had been expressed and purified in an intact form with the initial f-Met residue removed (data not shown).

DNA-protein complex molecularity

To determine the stoichiometry and molecularity for binding DNA, the DNA-protein complex was characterized with several methods including agarose gel electrophoresis, size-exclusion chromatography, and analytical equilibrium ultracentrifugation (Figure 3). Pioneering work by Klobutcher and co-workers (2) established that telomeres from mature macronuclei of *Euplotes crassus* have a 14-nucleotide, 3′ -terminal, single strand DNA extension of sequence TTTTGGGGTTTTGG. Accordingly, we selected our telomere-derived DNA fragments to contain this sequence and placed modifying groups such as the Cy5 fluorophore at the 5′ end so as to retain the 3′ OH group of authentic telomeres.

As monitored by agarose gel electrophoresis, titration of fluorescently labeled Cy5-d (TTTTGGGGTTTTGG) with increasing amounts of EcTEBP or EcTEBP-N shifted the position of the band corresponding with free DNA to a low mobility form, indicating formation of a DNA-protein complex (Figure 3A). The concentration of DNA in this experiment was close enough to the dissociation constant that it is reasonable to expect titration end-points at a slight excess of protein relative to DNA. As evident from smeared bands at titration mid-points, there is a tendency for protein and DNA to separate, possibly due to diffusion under the influence of an electric field.

When analyzed by size-exclusion chromatography, the apparent molecular weights measured for the free protein and its DNA-protein complex were consistent with monomeric EcTEBP and a DNA-protein complex of one-to-one molecularity (Figure 3B). Absorbance values measured at 260, 280 and 648 nm also indicated equimolar amounts of protein and Cy5-labeled oligonucleotide for the DNA-protein species eluting as a single peak in the size-exclusion chromatography experiment. Integrated absorbance intensities were conserved, demonstrating that the complex was more stable than suggested by its behavior in agarose gels and that there was no nuclease activity in our protein preparations.

Representative data for multiple analytical ultracentrifugation experiments are shown in Figure 3C. Over a 5 – 22 μM concentration range, recombinant *Euplotes crassus* telomere binding protein remained monomeric and the molecular size measured for the DNA-EcTEBP complex was as expected for a one-to-one molecular complex (Table I). These results together with size estimates from the gel-filtration experiments and observation of two DNA species in titrations monitored by agarose gel electrophoresis lead us to the conclusion that one molecule of Euplotes telomere binding protein binds with a molecule of TTTTGGGGTTTTGG DNA.

DNA binding performance compared for EcTEBP and EcTEBP-N

To evaluate and compare the DNA binding performance of EcTEBP and its N-terminal domain, dissociation constants were determined from binding titrations. In binding titrations the relative amounts of free and complexed forms of fluorescently labeled Cy5-d(TTTTGGGGTTTTGG)

DNA were inferred from fluorescence anisotropy measures without separating these species. Consequently, dissociation constants could be determined under true equilibrium conditions with this method. We realize that the Cy5-moeity may influence binding; however, two observations indicate that this effect is small: 1. Unmodified DNA and Cy5-labelled DNA behave as expected in competition experiments, and 2. In the related telomere DNA-protein system from *Oxytricha nova*, dissociation constants measured for unmodified DNA by isothermal titration calorimetry (32) and for Cy5-labeled DNA by fluorescence anisotropy titrations (25) are highly similar.

Figure 4 shows binding isotherms for $EcTEBP_{446}$ and EcTEBP-N_{288}. Fluorescence anisotropy *versus* protein concentration data could be satisfactorily modeled with simple one-site binding reactions. Titrations performed with DNA concentration in substantial excess relative to the expected dissociation constant indicated that recombinant EcTEBP and EcTEBP-N proteins were essentially 100% active for binding DNA (data not shown). The dissociation constant measured for DNA·EcTEBP was $K_{D\text{-}DNA} = 38 \pm 2$ nM, slightly lower than that measured for DNA·EcTEBP-N_{288}, $K_{D\text{-}DNA} = 60 \pm 4$ nM. These results demonstrate that the predicted N-terminal domain of EcTEBP retains high DNA-binding activity and suggest that the C-terminal domain may serve to augment DNA affinity to a modest degree.

To further characterize DNA binding performance of EcTEBP and EcTEBP-N, DNA association and dissociation kinetics were measured using surface plasmon resonance (BIACORE) to detect binding and release of biotin-d(TACATTTTGGGGTTTTGG) DNA attached to sensor chips (Figure 5). Association reactions followed single-phase bimolecular behavior and dissociation reactions followed single-phase unimolecular behavior. Rate constants for DNA association and dissociation are reported in Table II. The faster association and slower dissociation rates measured for EcTEBP relative to its N-terminal DNA binding domain, EcTEBP-N_{288} corroborated a somewhat enhanced DNA binding performance for the full-length protein.

DNA association was remarkably fast, approaching diffusion-limited rates. Slower DNA-association kinetics were observed for the *Oxytricha nova* alpha telomere protein (49). Because of this high rate of DNA-association, careful consideration of mass-transfer effects was necessary. Our initial trials using densely coated sensor chips were clearly mass-transfer limited.

During optimization of the experiment, we decreased the binding site density and presented the DNA binding site farther from the chip surface with use of the biotin-d(TACA)-linker. Under these conditions mass-transfer still exerts an influence on binding response curves; however, by including a kinetic parameter for mass-transfer (k_{mass} in Table II) intrinsic association and dissociation rate constants could be determined with confidence.

The ratio of k_d/k_a is somewhat lower than $K_{D\text{-}DNA}$ determined from equilibrium titrations. Buffer conditions required for well-behaved binding and surface regeneration in the BIACORE experiments were substantially different from those employed in anisotropy-monitored binding titrations. Solution composition differences together with the biotin-d(TACA)-linker employed in the kinetic experiments may account for the different estimates for stability of the DNA complex. Because of these differences, the more revealing comparison is between the full-length protein and its truncated N-terminal domain. Kinetic measurements and the equilibrium binding data are consistent with subtle but detectable enhancement of DNA binding performance attributable to the C-terminal domain of the *Euplotes crassus* telomere end binding protein.

Sequence permutation and DNA length preference

During telomere DNA synthesis, intermediates built from d(T_4G_4) repeats but differing in the identity of 3′ -terminal nucleotides are encountered (42). Mature DNA molecules, however, have an exact sequence always terminating with d(G_2) (2). To determine the telomere sequence permutation preferred by the telomere end binding protein in *Euplotes crassus*, binding isotherms were obtained for the eight possible 12-nucleotide oligos that retain a characteristic d(T_4G_4) repeat sequence and which start and stop at different points along that sequence (Table III). Initial trials demonstrated that a 12-nucleotide length was sufficient for efficient protein binding. By using this slightly shorter DNA length, we minimized the number of DNAs with two full d(GGGG) tracts. One of these 12-nucleotide DNAs does contain two d(GGGG) tracts and is, therefore, more likely to form G-quartet stabilized dimers than DNAs with only partial sets of G-tracts. G-quartet formation requires cations such as Na^+, K^+ or NH_4^+, which coordinate O^6 atoms of the G bases at a central cavity (44,50,51). Lithium ions are smaller and more difficult to dehydrate and, therefore, do not promote assembly of G-quartet structures. To minimize complications arising from structured DNAs, these experiments were conducted in lithium chloride-containing solutions and care was taken to keep sodium and potassium ion concentrations as low as practical.

All DNAs tested formed DNA-protein complexes, and, in each case, anisotropy measures could be modeled with a simple one-site binding reaction. Table III reports the dissociation constants determined for each telomere sequence permutation. A striking trend with respect to sequence register is apparent, with sequence permutation (*e*), which terminates with two Gs, binding most tightly and stepwise increases in $K_{D\text{-}DNA}$ observed with each nucleotide displacement from this naturally occurring sequence. The sequence preference measured here demonstrates that EcTEBP can distinguish among telomere extension intermediates.

DNA molecules in mature *Euplotes crassus* macronuclei terminate with exactly 14 nucleotides of single strand telomere repeat DNA. To test whether all nucleotides are needed for a stable DNA-EcTEBP complex, the relationship between DNA length and complex stability was examined through pairwise competition experiments. Fluorescently labeled Cy5-d (TTTTGGGGTTTTGG) DNA and EcTEBP were mixed with increasing amounts of unlabeled, 5′ -truncation variants 12-, 10- or 8-nucleotides in length, and the resulting proportion of fluorescence signal remaining in a DNA-protein complex was evaluated by agarose gel electrophoresis. The 14 and 12-nucleotide unmodified oligos competed efficiently for protein binding sites; approximately 50% of the fluorescence signal was released with equimolar unlabeled DNA (Figure 6). In contrast, the 8-nucleotide oligo competed poorly for protein binding sites with 100-fold excess of unlabeled oligo required for partial release of fluorescence signal. The 10-nucleotide unmodified oligo fared nearly as well as the 14 and 12-nucleotide oligos, although there is a noticeable difference in the amount of fluorescence signal released in the 1:1 competition reaction comparing these DNAs (Figure 6). These results indicated that 10 nucleotides may be the minimal DNA length required for a stable DNA-protein complex but also suggested that complex stability continues to increase with DNA length beyond these 10 nucleotides.

To further test DNA length preference, EcTEBP was incubated with a pooled mixture of d (TTTTGGGGTTTTGG), d(TTGGGGTTTTGG), d(GGGGTTTTGG), and d(GGTTTTGG) DNAs. In these reactions, each unlabeled DNA was in slight excess of total protein. Protein-bound DNA was isolated by size-exclusion chromatography followed by proteolysis. Analysis of the resulting selected DNA by Mono Q ion-exchange chromatography (Figure 7) showed that the telomere protein preferentially bound the longest DNA corresponding with full-length single strand telomere DNA found at the ends of mature DNA molecules in *Euplotes crassus*. Examination of the DNA fractions that did not bind protein confirmed the presence

of 12, 10 and 8-nucleotide oligos, indicating that absence of these DNAs in the protein-bound fraction was not the consequence of degradation or some other artifact.

DISCUSSION

To better understand telomere end structure and evolution, we sought to more completely characterize the DNA-binding properties of the *Euplotes crassus* telomere end binding protein, EcTEBP, and its predicted N-terminal DNA binding domain, EcTEBP-N. Recombinantly expressed EcTEBP and EcTEBP-N could be purified to near homogeneity (Figure 2). Each protein formed a complex with the 14-nucleotide single strand DNA derived from naturally occurring telomeres in *Euplotes crassus* (Figure 3). Molecular size measurements were consistent with monomeric protein and a DNA-complex containing one protein and one DNA molecule (Table I). Efficient binding appeared to require 10 nucleotides, yet all 14 nucleotides interact with protein as the full-length DNA out-competed shorter versions in a binding selection experiment (Figure 7). An important conclusion from these experiments is that the d (TTTTGGGGTTTTGG) single strand DNA binds just one telomere protein in *Euplotes crassus*. Molecularity of the telomere nucleoprotein complex is thus simpler in *Euplotes crassus* compared with that of *Oxytricha nova* where d(TTTTGGGGTTTTGGGG) single strand DNA binds multiple proteins (24,30), either two alpha proteins (29,32) or an alpha-beta heterodimer (20,30).

Comparison with previous characterization of EcTEBP biochemistry

The results reported in this work agree in large part with binding properties observed for protein purified directly from *Euplotes crassus* (34); however, noteworthy distinctions are apparent. EcTEBP purified from its natural source bound T and G-rich single strand DNA, but binding of the 14-nucleotide d(TTTTGGGGTTTTGG) DNA required attachment of this site to a longer piece of either single-stranded or double-stranded DNA (34). Additionally, while the full-length protein isolated from *Euplotes crassus* showed differential binding of telomere repeat DNAs ending with $d(G_2)$ and $d(T_4)$, the N-terminal domain prepared by trypsin digestion appeared to have diminished sequence specificity and formed complexes with both DNA sequences (34). In our experiments, short oligos formed DNA-protein complexes as determined by several methods, and the N-terminal domain, EcTEBP-N, recapitulated DNA-binding properties measured for the full-length protein, including telomere sequence permutation preference (see Table III).

It seems that the most likely factor accounting for these differences is that we have used a recombinant expression system that allowed for high yield and quality of the purified protein. Protein purified from *Euplotes crassus* was susceptible to aggregation especially as sensitized by detergents used in extraction steps (34). Solubility of recombinantly expressed protein was also a concern, overcome ultimately by keeping the preparation cold and avoiding excursions to room temperature during induction, extraction and purification steps. We note that the recombinant protein was essentially 100% active for DNA binding as determined by stoichiometric fluorescence anisotropy titrations.

Another source of differences relates to solution conditions, especially the ionic strength of binding experiments. Salt concentration was kept to an absolute minimum in the earlier work so as to keep the G-rich DNA in a largely extended conformation (34). In our studies, we either used short oligos with at most one full d(GGGG) tract or replaced sodium ions with lithium ions, and in this way we could perform binding experiments at close to physiological ionic strength without complications arising from structured DNA. Despite these experimental differences, both the current and previous work reach the same important conclusion that the telomere end binding protein has biochemical properties well suited to recognize, bind and protect the single strand DNA termini of telomeres in *Euplotes crassus* (34). The advantages

of our recombinant expression system allowed for quantitative measures of DNA binding not previously possible, which is significant because we can now make new comparisons with the telomere binding systems of other organisms.

Role of EcTEBP in determining telomere DNA length and structure

Telomere addition is closely coupled with chromosome fragmentation during macronuclear development (27,28,52). Telomere DNA isolated from anlagen — a term applied to developing macronuclei — were longer and more heterogeneous in length and sequence register than the precisely defined telomeres of mature macronuclei (42). Many anlagen telomeres terminated with $d(T_2)$ or $d(T_3)$ (42). Interestingly, our *in vitro* binding studies showed that while the telomere protein interacts most strongly with the telomere repeat permutation terminating with $d(G_2)$, the protein will also bind with alternative permutations including those corresponding with *de novo* telomere extension intermediates (Table III). Its biochemical behavior suggests the *Euplotes crassus* telomere protein participates in a mechanism that determines the final length and sequence register of single strand telomere DNA. By making weaker interactions with telomere repeat DNA encountered during telomere addition, the protein could facilitate DNA exchange of these unfinished molecules with endonucleases thought to process telomeres. By making a significantly stronger interaction with single strand DNA corresponding to mature telomeres and thereby physically excluding nuclease access, the protein could then positively identify the finished molecules, in effect a molecular coronation by binding.

Another possible role for EcTEBP relates to folded structures accessible to telomere DNA. G-rich single strand DNA readily adopts G-quartet stabilized structures *in vitro* (53-56); however, G-quartet DNA has not generally been found inside of cells. If unchecked, these folded DNA structures would pose a formidable challenge for telomere DNA synthesis since nucleotidyl transfer catalyzed by both telomerase and DNA polymerase requires that the 3′ -terminal nucleotides adopt an open and extended structure (57,58). Folded DNA structures are resolved upon binding with telomere end binding proteins *in vitro*, and the resulting DNA-protein complexes can either inhibit telomerase or serve as especially efficient substrates for telomerase depending on how much of the 3′ terminus remains exposed (31,59-61). Recent RNA interference experiments in *Stylonychia lemnae* suggest that telomere end binding proteins carefully regulate the formation of these folded DNA structures *in vivo* as well (17).

The use of multiple subunits to achieve high binding affinity in *Oxytricha nova* may relate to the high stability of folded G-quartet stabilized DNA structures accessible to the 16-nucleotide $d(T_4G_4T_4G_4)$ DNA. In *Euplotes*, the final d(GG) tract is shorter by two nucleotides, and therefore is less likely to form stable G-quartet structures. The very tight binding, $K_{D\text{-}DNA}$ = 1.4 ± 0.2 nM (ΔG = -12 kcal/mol), observed for an alpha-beta fusion protein complexed with *Oxytricha nova* single strand DNA (25) reflects significantly greater DNA-binding energy than measured here for the *Euplotes crasssus* DNA-protein complex, $K_{D\text{-}DNA}$ = 38 ± 2 nM (ΔG = -10 kcal/mol). These proteins were each expressed and purified using a similar bacterial expression system, and their characteristic dissociation constants were determined under similar conditions with the same method. Consequently, a difference in binding affinity of ΔΔG = -2 kcal/mol likely reflects real differences in functional requirements for the OnTEBP alpha-beta protein in *Oxytricha nova* and EcTEBP in *Euplotes crassus*. The DNA-protein complex in *Euplotes crassus* may be weaker since less energy is required to resolve or prevent the formation of folded DNA structures in this organism as a consequence of having only two Gs in the 3′ -terminal telomere repeat.

Role of the C-terminal domain in telomere protection

One motivation for this work was to potentially uncover a biochemical role for the C-terminal domain in the *Euplotes crassus* system. Possible roles are provided by analogy with the *Oxytricha nova* system where the C-terminal domain of alpha mediates extensive protein-protein interactions with the beta protein (20,24). Working under the assumption that there is no beta homologue in *Euplotes crassus*, we reasoned that the C-terminal domain may establish homotropic cooperative interactions between multiple telomere protein subunits or may make direct contact with some form of telomere DNA. Molecular size determination of the DNA-protein complex (Table I) and also the observed protein concentration dependence of DNA-binding equilibria (Figure 4) and DNA-binding association reactions (Figure 5 and Table II) make multimeric *Euplotes* telomere protein intermediates appear unlikely, at least for these short DNA oligonucleotides. The similar binding performance of recombinantly expressed EcTEBP and its N-terminal domain (Figure 4) appears to also rule out a direct role for the EcTEBP C-terminal domain in DNA-binding, although structural studies will be required to confirm this conclusion.

The C-terminal domain is a well-conserved entity for telomere proteins and telomere protein homologues in ciliated protozoa. Pairwise sequence comparisons emphasized conserved elements among N-terminal portions of these proteins (40), but when considered as a group, highly conserved sites are distributed uniformly between N and C-terminal domains (supplementary data, Figure S1). If the C-terminal domain of the *Euplotes* protein is dispensable for telomere DNA-binding, why is it so well preserved in this organism? The C-terminal domain could be interacting with other conserved proteins to facilitate events critical for telomere biology. Precedence for this suggestion comes from the *Saccharomyces cerevisiae* telomere system where the telomere protein CDC13 uses one portion to recognize single strand DNA (21,62,63) and other parts to recruit telomerase (12,64-66). The C-terminal domain of the *Euplotes crassus* telomere protein could be acting similarly to direct telomerase to telomere DNA in a regulated manner timed with DNA synthesis, potentially acting in concert with the *Euplotes crassus* telomere replication protein (35,36).

Binding energy is more focused at the 3′ -terminus

While results from the DNA-length experiments suggest that the entire length of telomere single strand DNA is important for DNA-protein complex stability, comparisons of the dissociation constant measured for Cy5-d(TTTTGGGGTTTTGG) (Figure 4) and those measured for the 12-nucleotide sequence permutations (Table III) indicate that the 3′ -terminal nucleotides are more critical for complex stability than the 3′-distal nucleotides. Sequence permutations (*e*) and (*g*) are related to the telomere-derived 14-nucleotide DNA by removal of two nucleotides from 5′ or 3′ -termini, respectively. Sequence permutation (*e*) has a DNA·EcTEBP$_{446}$ dissociation constant, $K_{\text{D-DNA}} = 55 \pm 3$ nM, which is nearly equivalent to that measured for the full-length DNA, $K_{\text{D-DNA}} = 38 \pm 2$ nM. When complex stability is determined with the N-terminal domain, the DNA·EcTEBP-N$_{288}$ dissociation constants are, within experimental uncertainty, the same for sequence permutation (*e*) and the full-length DNA. These comparisons indicate that the 5′ -terminal nucleotides missing in sequence permutation (*e*) contribute very little to overall binding energy. Sequence permutation (*g*), on the other hand, has a significantly higher dissociation constant compared with the full-length 14-nucleotide DNA binding site for both DNA·EcTEBP$_{446}$ and DNA·EcTEBP-N$_{288}$, demonstrating that the removed 3′ -terminal nucleotides contribute at least —1.2 kcal/mol to binding free energy.

A key question remains regarding the environment and structure of the 3′ -terminal nucleotides in the DNA-EcTEBP complex. Our data are consistent with binding contacts between protein and the 3′ -terminal nucleotides. It is not clear, however, whether the DNA adopts a loop

structure with the 3′-OH group completely buried as found in the crystal structure of a *Oxytricha nova* alpha-beta-DNA complex (20), or whether the final nucleotides instead interact with a more open pocket like that of the *Oxytricha nova* alpha N-terminal domain which retains a 3′-OH group preference but which can nonetheless accommodate additional nucleotides (32,49, 67). A *Euplotes crassus* telomere DNA-protein complex structure, to be determined in the future, may confirm one of these possibilities or may reveal a distinct mechanism for capping telomere ends.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENT

We acknowledge cDNA generously provided by C. Price, help with analytical ultracentrifugation experiments from L. Joss, BIACORE kinetic analysis by D. Myszka at the University of Utah's Center for Biomolecular Interaction, use of a spectrofluorimeter in the lab of D. Blumenthal, and temporary access to a micro-titer plate-reader with fluorescence anisotropy capability generously loaned by Molecular Devices. The NIH funded this work with a grant to M.P.H. (R01 GM067994). The University of Utah DNA/Peptide Research Core facility receives support from the National Cancer Institute (5P30 CA42014).

REFERENCES

1. Blackburn EH, Gall JG. A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in *Tetrahymena*. J Mol Biol 1978;120:33–53. [PubMed: 642006]
2. Klobutcher LA, Swanton MT, Donini P, Prescott DM. All gene-sized DNA molecules in four species of hypotrichs have the same terminal sequence and an unusual 3′ terminus. Proc Natl Acad Sci USA 1981;78:3015–9. [PubMed: 6265931]
3. Forney J, Henderson ER, Blackburn EH. Identification of the telomeric sequence of the acellular slime molds *Didymium iridis* and *Physarum polycephalum*. Nucleic Acids Res 1987;15:9143–52. [PubMed: 3684591]
4. Schechtman MG. Characterization of telomere DNA from *Neurospora crassa*. Gene 1990;88:159–65. [PubMed: 1971801]
5. Javerzat JP, Bhattacherjee V, Barreau C. Isolation of telomeric DNA from the filamentous fungus *Podospora anserina* and construction of a self-replicating linear plasmid showing high transformation frequency. Nucleic Acids Res 1993;21:497–504. [PubMed: 8441663]
6. Edman JC. Isolation of telomerelike sequences from *Cryptococcus neoformans* and their use in high-efficiency transformation. Mol Cell Biol 1992;12:2777–83. [PubMed: 1588969]
7. Coleman MJ, McHale MT, Arnau J, Watson A, Oliver RP. Cloning and characterisation of telomeric DNA from *Cladosporium fulvum*. Gene 1993;132:67–73. [PubMed: 8104848]
8. Meyne J, Ratliff RL, Moyzis RK. Conservation of the human telomere sequence (TTAGGG)n among vertebrates. Proc Natl Acad Sci USA 1989;86:7049–53. [PubMed: 2780561]
9. Zakian VA. Telomeres: beginning to understand the end. Science 1995;270:1601–7. [PubMed: 7502069]
10. Henderson ER, Blackburn EH. An overhanging 3′ terminus is a conserved feature of telomeres. Mol Cell Biol 1989;9:345–8. [PubMed: 2927395]
11. McElligott R, Wellinger RJ. The terminal DNA structure of mammalian chromosomes. Embo J 1997;16:3705–14. [PubMed: 9218811]
12. Lin JJ, Zakian VA. The Saccharomyces CDC13 protein is a single-strand TG1-3 telomeric DNA-binding protein in vitro that affects telomere behavior *in vivo*. Proc Natl Acad Sci USA 1996;93:13760–5. [PubMed: 8943008]
13. Baumann P, Cech TR. Pot1, the putative telomere end-binding protein in fission yeast and humans. Science 2001;292:1171–5. [PubMed: 11349150]
14. Gottschling DE, Zakian VA. Telomere proteins: specific recognition and protection of the natural termini of *Oxytricha macronuclear* DNA. Cell 1986;47:195–205. [PubMed: 3094961]

15. Price CM, Cech TR. Properties of the telomeric DNA-binding protein from *Oxytricha nova*. Biochemistry 1989;28:769–74. [PubMed: 2713343]
16. Price CM, Cech TR. Telomeric DNA-protein interactions of *Oxytricha* macronuclear DNA. Genes Dev 1987;1:783–93. [PubMed: 3123321]
17. Paeschke K, Simonsson T, Postberg J, Rhodes D, Lipps HJ. Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo. Nat Struct Mol Biol 2005;12:847–54. [PubMed: 16142245]
18. Murzin AG. OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. Embo J 1993;12:861–7. [PubMed: 8458342]
19. Theobald DL, Cervantes RB, Lundblad V, Wuttke DS. Homology among telomeric end-protection proteins. Structure (Camb) 2003;11:1049–50. [PubMed: 12962623]
20. Horvath MP, Schweiker VL, Bevilacqua JM, Ruggles JA, Schultz SC. Crystal structure of the *Oxytricha nova* telomere end binding protein complexed with single strand DNA. Cell 1998;95:963–74. [PubMed: 9875850]
21. Mitton-Fry RM, Anderson EM, Theobald DL, Glustrom LW, Wuttke DS. Structural basis for telomeric single-stranded DNA recognition by yeast Cdc13. J Mol Biol 2004;338:241–55. [PubMed: 15066429]
22. Lei M, Podell ER, Baumann P, Cech TR. DNA self-recognition in the structure of Pot1 bound to telomeric single-stranded DNA. Nature 2003;426:198–203. [PubMed: 14614509]
23. Lei M, Podell ER, Cech TR. Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection. Nat Struct Mol Biol 2004;11:1223–9. [PubMed: 15558049]
24. Fang G, Gray JT, Cech TR. *Oxytricha* telomere-binding protein: separable DNA-binding and dimerization domains of the alpha-subunit. Genes Dev 1993;7:870–82. [PubMed: 8491383]
25. Buczek P, Orr RS, Pyper SR, Shum M, Kimmel E, Ota I, Gerum SE, Horvath MP. Binding linkage in a telomere DNA-protein complex at the ends of *Oxytricha nova* chromosomes. J Mol Biol 2005;350:938–52. [PubMed: 15967465]
26. Prescott DM, Murti KG. Chromosome structure in ciliated protozoans. Cold Spring Harb Symp Quant Biol 1973;38:609–18. [PubMed: 4208796]
27. Klobutcher LA, Herrick G. Developmental genome reorganization in ciliated protozoa: the transposon link. Prog Nucleic Acid Res Mol Biol 1997;56:1–62. [PubMed: 9187050]
28. Jahn CL, Klobutcher LA. Genome remodeling in ciliated protozoa. Annu Rev Microbiol 2002;56:489–520. [PubMed: 12142486]
29. Gray JT, Celander DW, Price CM, Cech TR. Cloning and expression of genes for the *Oxytricha* telomere-binding protein: specific subunit interactions in the telomeric complex. Cell 1991;67:807–14. [PubMed: 1840510]
30. Fang G, Cech TR. *Oxytricha* telomere-binding protein: DNA-dependent dimerization of the alpha and beta subunits. Proc Natl Acad Sci USA 1993;90:6056–60. [PubMed: 8327484]
31. Froelich-Ammon SJ, Dickinson BA, Bevilacqua JM, Schultz SC, Cech TR. Modulation of telomerase activity by telomere DNA-binding proteins in *Oxytricha*. Genes Dev 1998;12:1504–14. [PubMed: 9585510]
32. Buczek P, Horvath MP. Thermodynamic characterization of binding *Oxytricha nova* single strand telomere DNA with the alpha protein N-terminal domain. J. Mol. Biol. 2006, (in press, accepted 17-Feb-2006, doi:10.1016/j.jmb.2006.02.082).
33. Price CM. Telomere structure in *Euplotes crassus:* characterization of DNA-protein interactions and isolation of a telomere-binding protein. Mol Cell Biol 1990;10:3421–31. [PubMed: 2355912]
34. Price CM, Skopp R, Krueger J, Williams D. DNA recognition and binding by the *Euplotes telomere* protein. Biochemistry 1992;31:10835–43. [PubMed: 1420196]
35. Skopp R, Wang W, Price C. rTP: a candidate telomere protein that is associated with DNA replication. Chromosoma 1996;105:82–91. [PubMed: 8753697]
36. Carlson DL, Skopp R, Price CM. DNA-Binding properties of the replication telomere protein. Biochemistry 1997;36:15900–8. [PubMed: 9398323]

37. Sogin ML, Swanton MT, Gunderson JH, Elwood HJ. Sequence of the small subunit ribosomal RNA gene from the *hypotrichous ciliate Euplotes aediculatus*. J Protozool 1986;33:26–9. [PubMed: 3007752]
38. Sogin ML, Elwood HJ. Primary structure of the Paramecium tetraurelia small-subunit rRNA coding region: phylogenetic relationships within the *Ciliophora*. J Mol Evol 1986;23:53–60. [PubMed: 3084799]
39. Croft KE, Dalby AB, Hogan DJ, Orr KE, Hewitt EA, Africa RJ, DuBois ML, Prescott DM. Macronuclear molecules encoding actins in spirotrichs. J Mol Evol 2003;56:341–50. [PubMed: 12612837]
40. Wang W, Skopp R, Scofield M, Price C. *Euplotes crassus* has genes encoding telomere-binding proteins and telomere-binding protein homologs. Nucleic Acids Res 1992;20:6621–9. [PubMed: 1480483]
41. Vermeesch JR, Williams D, Price CM. Telomere processing in *Euplotes*. Nucleic Acids Res 1993;21:5366–71. [PubMed: 8265350]
42. Vermeesch JR, Price CM. Telomeric DNA sequence and structure following de novo telomere synthesis in *Euplotes crassus*. Mol Cell Biol 1994;14:554–66. [PubMed: 8264623]
43. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 1997;25:4876–82. [PubMed: 9396791]
44. Horvath MP, Schultz SC. DNA G-quartets in a 1.86 A resolution structure of an *Oxytricha nova* telomeric protein-DNA complex. J Mol Biol 2001;310:367–77. [PubMed: 11428895]
45. Felsenstein, J. 2004.
46. Cantor CR, Warshaw MM, Shapiro H. Oligonucleotide interactions. 3. Circular dichroism studies of the conformation of deoxyoligonucleotides. Biopolymers 1970;9:1059–77. [PubMed: 5449435]
47. Press, WH.; Teukolsky, SA.; Vetterling, WT.; Flannery, BP. Numerical recipes in C. Cambridge University Press; Cambridge (UK): 1992. Confidence limits on estimated model parameters; p. 689-699.
48. Myszka DG, X. H, Dembo M, Morton TA, Goldstein B. Extending the range of rate constants available from BIACORE: interpreting mass transport-influenced binding data. Biophys J 1998;75:583–94. [PubMed: 9675161]
49. Classen S, Lyons D, Cech TR, Schultz SC. Sequence-specific and 3′-end selective single-strand DNA binding by the *Oxytricha nova* telomere end binding protein alpha subunit. Biochemistry 2003;42:9269–77. [PubMed: 12899613]
50. Phillips K, Dauter Z, Murchie AI, Lilley DM, Luisi B. The crystal structure of a parallel-stranded guanine tetraplex at 0.95 Å resolution. J. Mol. Biol 1997;273:171–182. [PubMed: 9367755]
51. Hud NV, Schultze P, Sklenar V, Feigon J. Binding sites and dynamics of ammonium ions in a telomere repeat DNA quadruplex. J Mol Biol 1999;285:233–43. [PubMed: 9878402]
52. Klobutcher LA. Characterization of *in vivo* developmental chromosome fragmentation intermediates in *E. crassus*. Mol Cell 1999;4:695–704. [PubMed: 10619017]
53. Sen D, Gilbert W. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. Nature 1988;334:364–6. [PubMed: 3393228]
54. Sundquist WI, Klug A. Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. Nature 1989;342:825–9. [PubMed: 2601741]
55. Smith FW, Feigon J. Quadruplex structure of *Oxytricha* telomeric DNA oligonucleotides. Nature 1992;356:164–8. [PubMed: 1545871]
56. Parkinson GN, Lee MP, Neidle S. Crystal structure of parallel quadruplexes from human telomeric DNA. Nature 2002;417:876–80. [PubMed: 12050675]
57. Zahler AM, Prescott DM. Telomere terminal transferase activity in the *hypotrichous ciliate Oxytricha* nova and a model for replication of the ends of linear DNA molecules. Nucleic Acids Res 1988;16:6953–72. [PubMed: 3136437]
58. Zahler AM, Williamson JR, Cech TR, Prescott DM. Inhibition of telomerase by G-quartet DNA structures. Nature 1991;350:718–20. [PubMed: 2023635]

59. Shippen DE, Blackburn EH, Price CM. DNA bound by the *Oxytricha* telomere protein is accessible to telomerase and other DNA polymerases. Proc Natl Acad Sci USA 1994;91:405–9. [PubMed: 7506421]

60. Lei M, Zaug AJ, Podell ER, Cech TR. Switching human telomerase on and off with hPOT1 protein in vitro. J Biol Chem 2005;280:20449–56. [PubMed: 15792951]

61. Zaug AJ, Podell ER, Cech TR. Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro. Proc Natl Acad Sci USA 2005;102:10864–9. [PubMed: 16043710]

62. Anderson EM, Halsey WA, Wuttke DS. Delineation of the high-affinity single-stranded telomeric DNA-binding domain of *Saccharomyces cerevisiae* Cdc13. Nucleic Acids Res 2002;30:4305–13. [PubMed: 12364610]

63. Mitton-Fry RM, Anderson EM, Hughes TR, Lundblad V, Wuttke DS. Conserved structure for single-stranded telomeric DNA recognition. Science 2002;296:145–7. [PubMed: 11935027]

64. Lendvay TS, Morris DK, Sah J, Balasubramanian B, Lundblad V. Senescence mutants of *Saccharomyces cerevisiae* with a defect in telomere replication identify three additional EST genes. Genetics 1996;144:1399–412. [PubMed: 8978029]

65. Evans SK, Lundblad V. Est1 and Cdc13 as comediators of telomerase access. Science 1999;286:117–20. [PubMed: 10506558]

66. Qi H, Zakian VA. The *Saccharomyces* telomere-binding protein Cdc13p interacts with both the catalytic subunit of DNA polymerase alpha and the telomerase-associated est1 protein. Genes Dev 2000;14:1777–88. [PubMed: 10898792]

67. Classen S, Ruggles JA, Schultz SC. Crystal structure of the N-terminal domain of *Oxytricha nova* telomere end-binding protein alpha subunit both uncomplexed and complexed with telomeric ssDNA. J Mol Biol 2001;314:1113–25. [PubMed: 11743727]

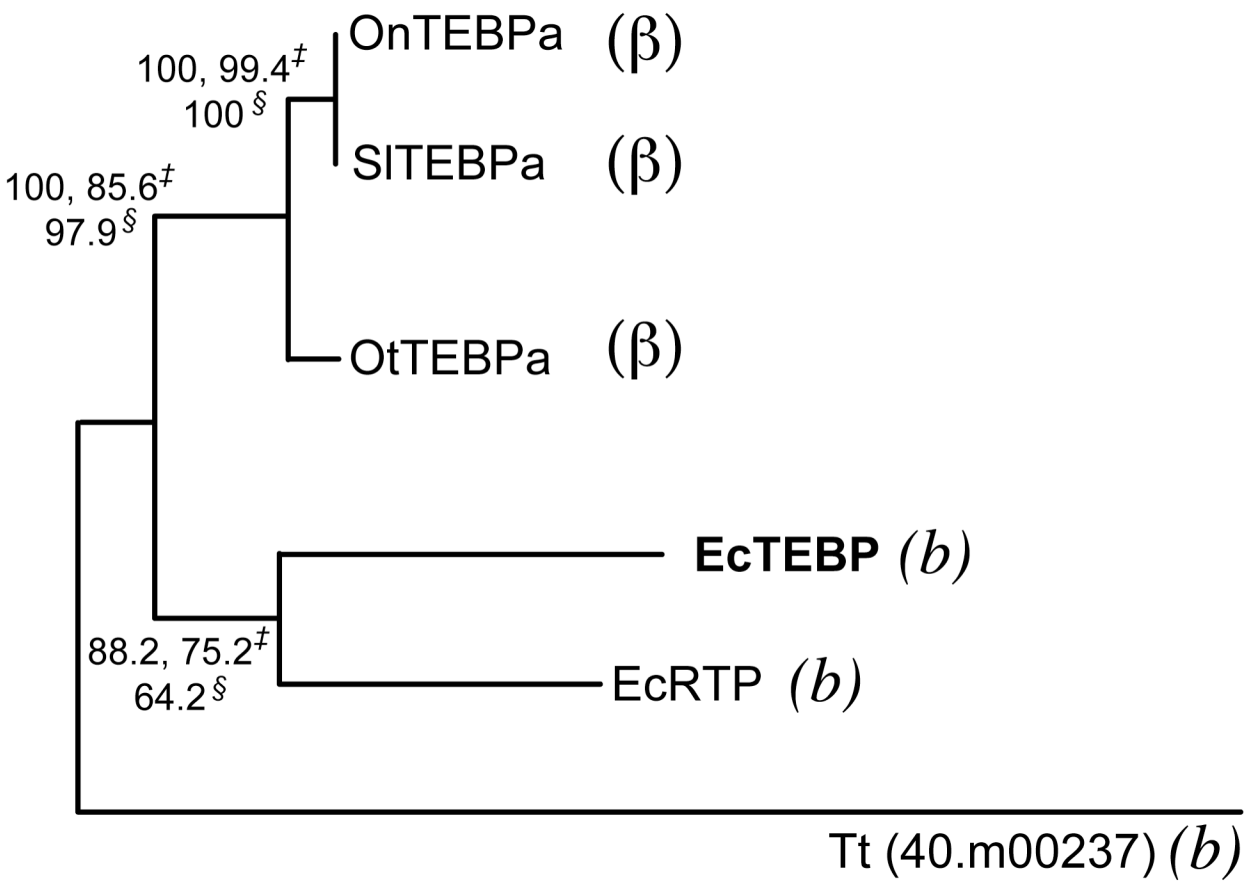

FIGURE 1.

Telomere proteins and telomere protein homologues from ciliates. A dendrogram shows phylogenetic relationships among telomere end binding proteins and homologues from ciliates. Presence of a beta telomere protein in the *Oxytricha* and *Stylonychia* species is indicated with (β). *Euplotes* and *Tetrahymena* do not have a beta homologue but may use another protein in place of beta. The possibility of a non-homologous analogue of beta is indicated with (*b*) next to EcTEBP, EcRTP, and Tt(40.m00237). Bootstrap support for branches are indicated for most parsimonious, maximum likelihood ‡, and neighbor joining analyses §.

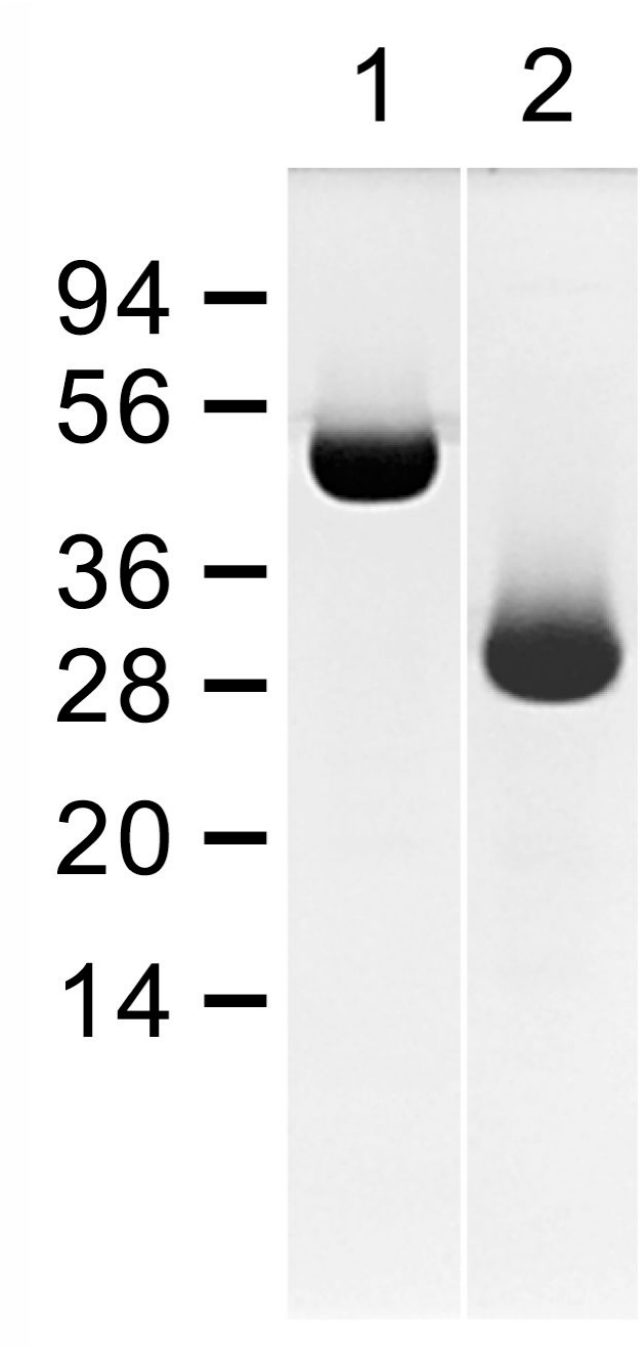

FIGURE 2.

SDS-polyacrylamide gel electrophoresis of purified *Euplotes crassus* telomere end binding protein (lane 1) and its predicted N-terminal DNA binding domain (lane 2). Approximately 20 μg of pure protein was loaded in each lane and detected by Coomassie staining. The apparent molecular weights of $EcTEBP_{446}$ and $EcTEBP\text{-}N_{288}$ were 52 ± 2 kDa and 32 ± 2 kDa, respectively, in close agreement with 51.4 kDa and 32.9 kDa — weights calculated on the basis of amino acid sequence and confirmed by mass spectrometry.

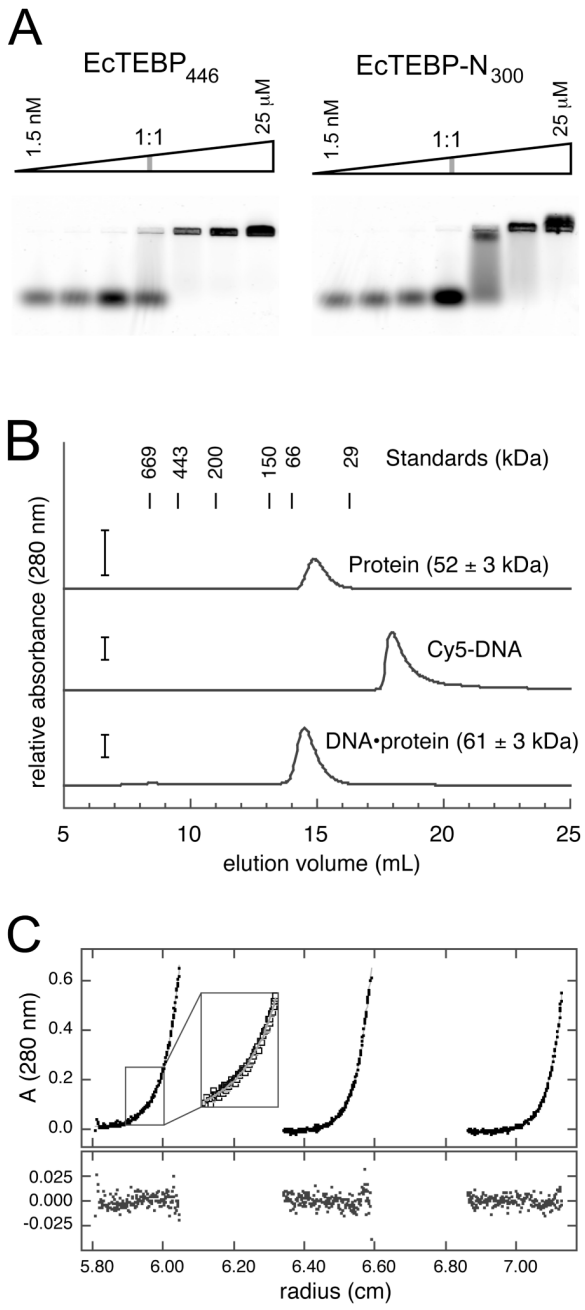

FIGURE 3.

Evaluation of DNA-protein complex formation by the *Euplotes crassus* telomere end binding protein. (A) Agarose gel electrophoresis of fluorescently labeled Cy5-d (TTTTGGGGTTTTGG) DNA indicated formation of a DNA-protein complex. Free DNA traveled towards the anode (below the gel), and the DNA-protein complex favored with increasing protein levels traveled more slowly towards the cathode (above the gel). Enhanced fluorescence for the sample just preceding complex formation is reproducible and may reflect protection from solvent quenching. (B) Gel-filtration Superdex-200 chromatography of EcTEBP, Cy5-labeled DNA, and a 1:1 mixture of protein and DNA showed elution volumes indicating monomeric protein and a DNA-protein complex with one-to-one molecularity. Note

that the protein's chromatogram is depicted with a 2-fold more sensitive relative absorbance scale. (C) Representative radial scans obtained from analytical ultracentrifugation of DNA-EcTEBP at three different concentrations. Data shown here, obtained at 24,000 rpm, were analyzed together with data at other angular velocities to determine the molecular size of the DNA-protein complex and of the protein alone (Table I). Under all conditions tested, the protein behaved as a monomer and complexes with the 14-nucleotide telomere-derived Cy5-d (TTTTGGGGTTTTGG) DNA were of one-to-one molecularity.

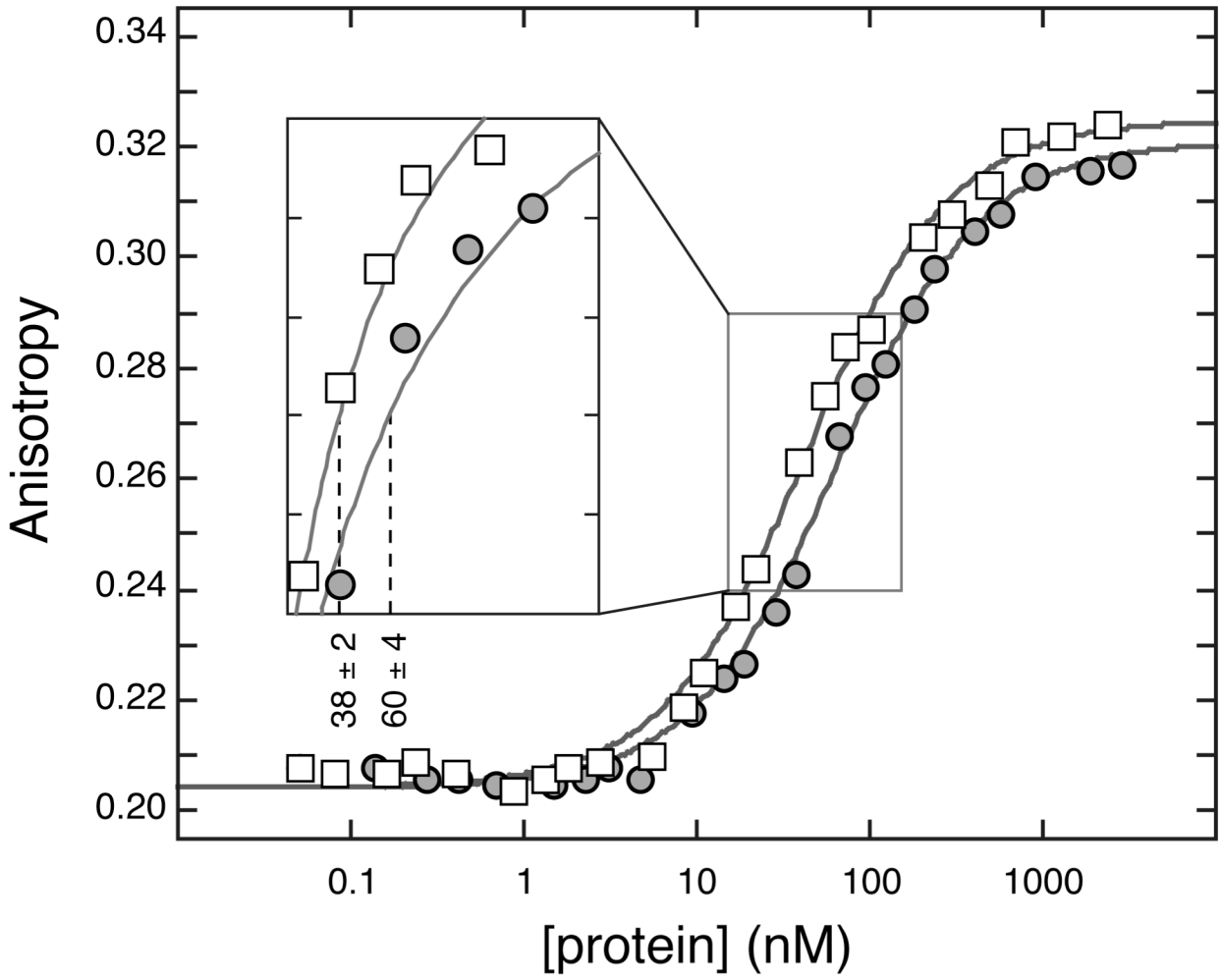

FIGURE 4.

Binding isotherms for the *Euplotes crassus* telomere end binding protein and its N-terminal domain. Fluorescence anisotropy is plotted as a function of $EcTEBP_{446}$ (squares) and EcTEBP-N_{288} (circles) concentration. Similar results were obtained with a second form of the N-terminal domain, EcTEBP-N_{300} (data not shown). Model curves represent the results of nonlinear least-squares fitting of the measured data to a simple one-site binding model. The full-length $EcTEBP_{446}$ binds Cy5-d(TTTTGGGGTTTTGG) DNA with $K_{D\text{-}DNA}$ = 38 ± 2 nM, slightly lower than $K_{D\text{-}DNA}$ = 60 ± 4 nM measured for the N-terminal DNA-binding domain.

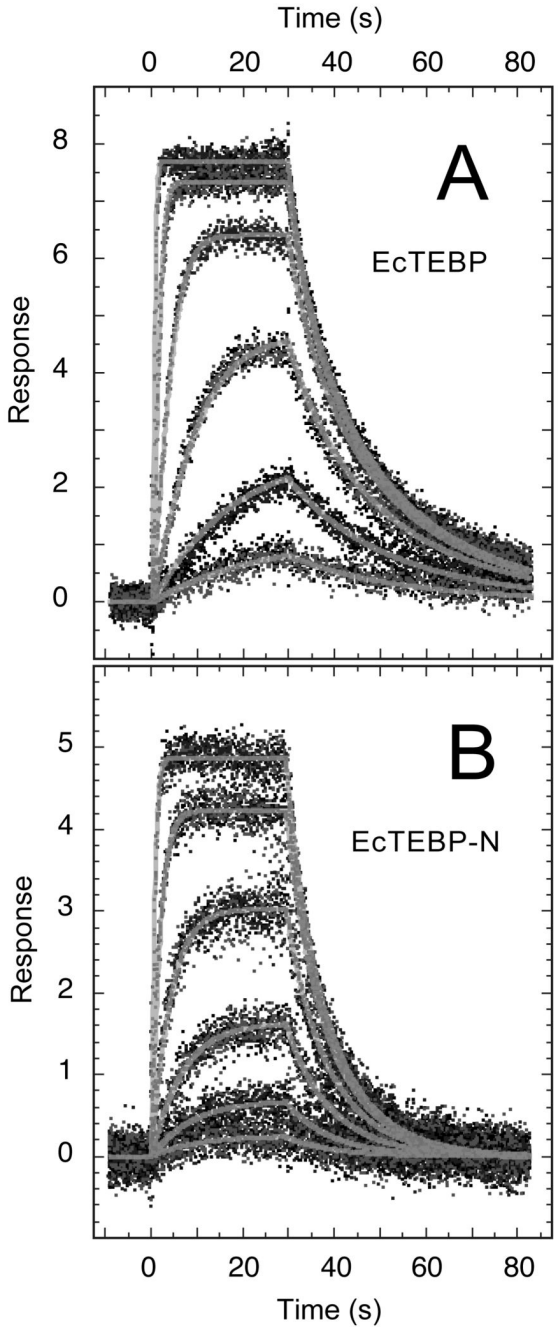

FIGURE 5.

Kinetics of DNA association and dissociation for full-length *Euplotes crassus* telomere end binding protein (A) and its N-terminal DNA-binding domain (B). Traces for each of six protein concentrations were obtained in triplicate. Model curves represent nonlinear least-squares fitting of multiple data sets. Values of kinetic rate constants are reported in Table II. The full-length $EcTEBP_{446}$ binds DNA faster and dissociates more slowly relative to the behavior of the EcTEBP-N_{288} DNA-binding domain.

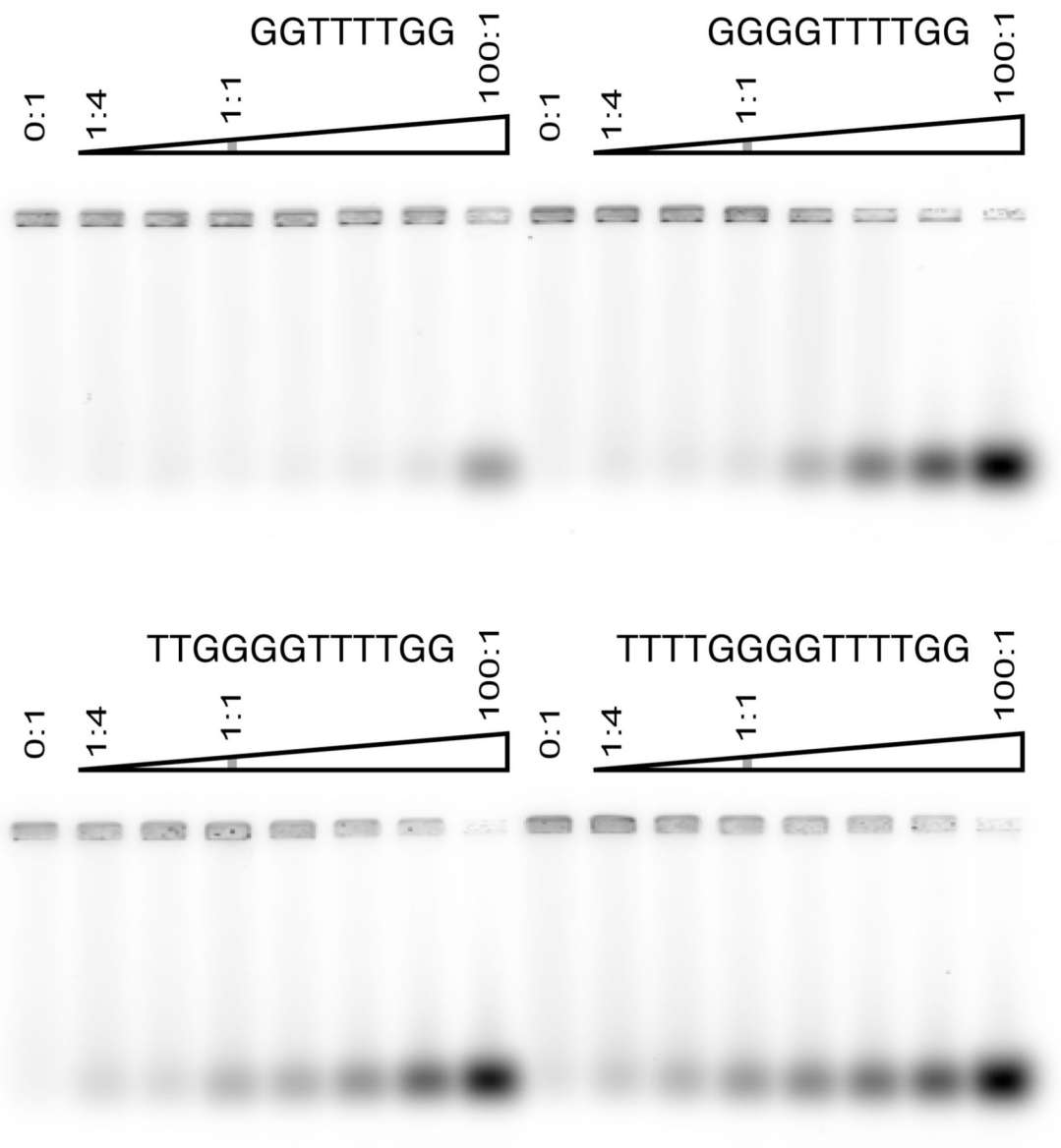

FIGURE 6.

Pairwise competition experiments. Fluorescently labeled 14-nucleotide DNA, Cy5-d (TTTTGGGGTTTTGG), competed for binding to EcTEBP with four different lengths of unmodified DNA. The concentration of EcTEBP and Cy5-labeled DNA were each 2 μM. A control sample without unmodified competitor DNA (0:1) was included in each panel and is followed, left-to-right, by increasing amounts of unmodified competitor DNA at μM concentrations of 0.5, 1, 2 (1:1), 4, 8, 18, and 200 (100:1).

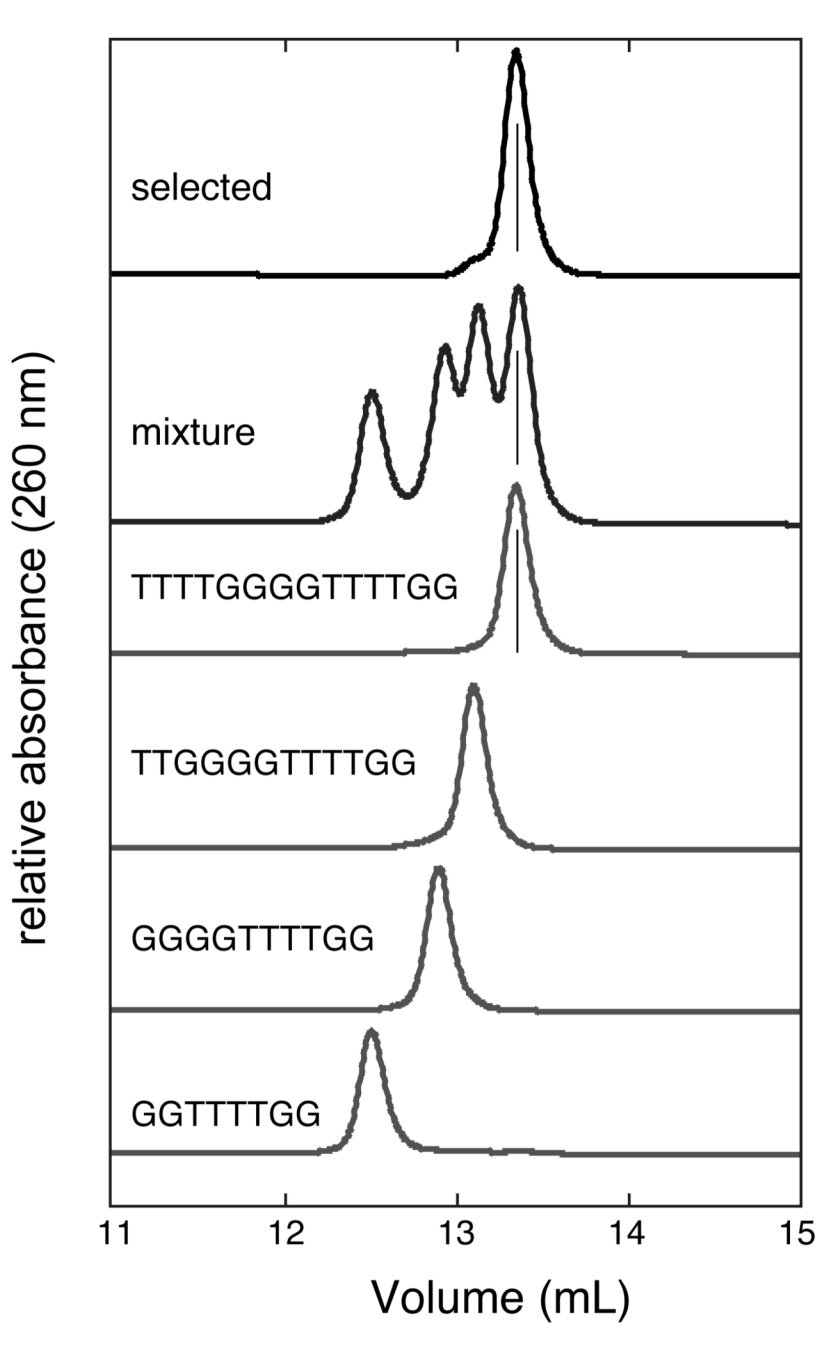

FIGURE 7.

DNA length preference of the *Euplotes crassus* telomere end binding protein. Each absorbance trace shows the lithium chloride-gradient elution of DNA from an anion-exchange Mono Q column. All four DNAs, d(GGTTTTGG), d(GGGGTTTTGG), d(TTGGGGTTTTGG), and d (TTTTGGGGTTTTGG) were combined to generate an equimolar mixture. Total DNA extracted from DNA-protein complexes (top trace, labeled “selected”) showed that the telomere end binding protein preferentially binds full-length 14-nucleotide telomeric DNA. The leading shoulder apparent in the “selected” DNA trace indicates low-level (~3%) binding of 12-nucleotide telomeric DNA.

Table I

Molecular mass (Da) measured by equilibrium sedimentation

Species	Equilibrium sedimentation	Calculated
EcTEBP	50,800 ± 400	51,413.4
EcTEBP·DNA	56,100 ± 300	56,292.8*

* Cy5-TTTTGGGGTTTTGG adds 4,879.4 Da assuming a 1:1 DNA·protein complex. Errors are the sample standard deviations measured from Monte Carlo bootstrap simulations.

Table II

Kinetic constants measured for binding and release of telomere DNA*

Protein	k_a (M^{-1} s^{-1})	k_d (s^{-1})	k_{mass} (RU M^{-1} s^{-1})	k_d / k_a(M)
$EcTEBP_{446}$	4.5 (4) × 10^7	0.10 (2)	1.8 (2) × 10^8	2.2 (5) × 10^{-9}
$EcTEBP\text{-}N_{288}$	1.8 (1) × 107	0.15 (1)	1.6 (2) × 108	8.3 (7) × 10–9

* Uncertainty in the least significant digit is reported as two standard error units in parenthesis

Table III

Dissociation constants measured for 12-nucleotide telomere repeat permutations*

Index	DNA	$K_{D\text{-}DNA}$ (nM) EcTEBP$_{446}$	$K_{D\text{-}DNA}$ (nM) EcTEBP-N$_{288}$
(*a*) ±4	GGTTTTGGGGTT	1200 ± 100	1400 ± 200
(*b*) +3	GGGTTTTGGGGT	730 ± 110	----‡
(*c*) +2	GGGGTTTTGGGG	430 ± 30	440 ± 70
(*d*) +1	TGGGGTTTTGGG	170 ± 10	----
(*e*) 0	TTGGGGTTTTGG	55 ± 3	57 ± 5
(*f*) –1	TTTGGGGTTTTG	160 ± 10	----
(*g*) –2	TTTTGGGGTTTT	400 ± 40	520 ± 80
(*h*) –3	GTTTTGGGGTTT	650 ± 80	----

*Errors are the sample standard deviations obtained from Monte Carlo bootstrap simulations.

‡Not determined.