Published in final edited form as:

J Mol Biol. 2005 July 29; 350(5): 938–952. doi:10.1016/j.jmb.2005.05.040.

Binding Linkage in a Telomere DNA–Protein Complex at the Ends of *Oxytricha nova* Chromosomes

Pawel Buczek, Rochelle S. Orr, Sean R. Pyper, Mili Shum, Emily Kimmel Irene Ota, Shawn E. Gerum, and Martin P. Horvath*

Department of Biology, University of Utah, 257 S, 1400 E, Salt Lake City, UT, 84112-0840, USA

Abstract

Alpha and beta protein subunits of the telomere end binding protein from *Oxytricha nova* (OnTEBP) combine with telomere single strand DNA to form a protective cap at the ends of chromosomes. We tested how protein–protein interactions seen in the co-crystal structure relate to DNA binding through use of fusion proteins engineered as different combinations of domains and subunits derived from OnTEBP. Joining alpha and beta resulted in a protein that bound single strand telomere DNA with high affinity ($K_{D-DNA}=1.4$ nM). Another fusion protein, constructed without the C-terminal protein–protein interaction domain of alpha, bound DNA with 200-fold diminished affinity ($K_{D-DNA}=290$ nM) even though the DNA-binding domains of alpha and beta were joined through a peptide linker. Adding back the alpha C-terminal domain as a separate protein restored high-affinity DNA binding. The binding behaviors of these fusion proteins and the native protein subunits are consistent with cooperative linkage between protein–association and DNA-binding equilibria. Linking DNA–protein stability to protein–protein contacts at a remote site may provide a trigger point for DNA–protein disassembly during telomere replication when the single strand telomere DNA must exchange between a very stable OnTEBP complex and telomerase.

Keywords

telomere-binding protein; protein-DNA interactions; protein-protein interactions; binding linkage; protein engineering

Introduction

Telomeres are specialized DNA–protein structures that protect the ends of chromosomes and distinguish natural chromosome termini from unnatural breaks produced by DNA damage.¹ The substrate, intermediates, and product of telomerase-mediated telomere DNA elongation involve a single strand DNA extension that protrudes past the region of duplex telomere DNA,² and such single strand telomere DNA has been found in protozoa, yeast, and vertebrates.^{3–5} Telomere DNA length and protein composition vary during the cell cycle and lifespan of a cell, and telomere structures may involve double strand DNA binding proteins,^{6–12} single strand DNA binding proteins,^{13–16} DNA–DNA associations involving G-quartet stabilized quadruplexes,^{17–19} and lariat DNA structures called t-loops.^{20–22} Alterations in telomere structure are associated with distinct cellular programs including senescence, apoptosis, and unlimited proliferation,^{23–26} indicating that telomeres play an important role in the processes of aging and cancer.^{27,28}

*Corresponding author: horvath@biology.utah.edu.

The telomeres from *Oxytricha nova* macronuclei are relatively simple, consisting of two and a half repeats of the duplex sequence d(TTTTGGGG)/d(CCCCAAAA), a 16 nucleotide, 3'-terminal single strand DNA extension of sequence d(TTTTGGGGTTTGGGG), and a protein cap comprising two subunits that bind to the single strand DNA^{3,13,14} (see Figure 1). The protein subunits and telomere DNA-protein complex have been well characterized.^{29–34} The 56 kDa (495 aa) alpha subunit consists of two separable domains: a 36 kDa N-terminal DNA-binding domain and a 20 kDa C-terminal domain important for interaction with the beta subunit.³¹ The 41 kDa (385 aa) beta subunit does not bind with telomere DNA on its own but enhances and modifies the DNA binding activity of alpha subunit²⁹ and, together with alpha, makes direct DNA contacts.^{32,35,36} The N-terminal 260 amino acid residues of beta define a 28 kDa protease-resistant domain that functions equivalently to the full-length beta subunit in protection of single strand DNA.³¹

Crystal structures of DNA-protein complexes potentially encountered during telomere end cap assembly include alpha complexed with telomere DNA³⁴ and alpha N-terminal domain complexed with telomere DNA.³³ In each of these DNA-complexes that lack the beta subunit, 3'-terminal GGGG nucleotides bind to the alpha N-terminal domain at a site termed the “alpha” site.^{33,34} In a co-crystal structure of the full *O. nova* alpha-beta-DNA complex,³² telomere DNA has slipped by one repeat unit so that the alpha site remains occupied, albeit with internal GGGG nucleotides, and 3'-terminal GGGG nucleotides nestle into a new binding site created by the juxtaposition of alpha and beta, accordingly termed the “alpha-beta” site.

The alpha-beta-DNA crystal structure additionally reveals a curiously extended peptide loop contributed by the beta subunit, which clasps onto the alpha subunit C-terminal domain (Figure 1). Despite a large surface area with a high degree of apparently stabilizing hydrophobic contacts at the alpha-beta interface, the two protein subunits associate in significant amounts only at very high concentrations of protein (P.B., unpublished results) or in the presence of DNA, a property identified as “DNA-dependent” subunit association.³⁰

In order to further understand how protein-protein interactions between alpha and beta relate to DNA binding, we designed and purified three novel fusion proteins engineered as different combinations of domains derived from the native alpha and beta subunits. DNA binding performance measured for each protein demonstrates that protein-protein interaction is critical for stabilizing the DNA-protein complex. In this system, the beta subunit acts as an allosteric effector that enhances an initially moderate DNA-protein stability ($K_{D-DNA} \sim 200$ nM) to achieve high DNA-protein stability ($K_{D-DNA} \sim 1$ nM). A peptide linker tethering the DNA binding elements from each subunit was a poor substitute for alpha-beta interactions. The energetic benefit of adding beta to the complex is only realized if authentic protein-protein interactions are restored. Evolution may have adopted the subunit and domain architecture seen in the OnTEBP-DNA complex in order to facilitate timely exchange of telomere single strand DNA from a protective protein-capped state to a telomerase-associated state required for telomere DNA replication.

Results

Design and expression of fusion proteins

To explore how domain architecture relates to DNA binding function, fusion proteins were engineered as new combinations of domains found in the alpha and beta subunits of the native *O. nova* telomere end binding protein (OnTEBP). The lengths of amino acid linkers connecting domains and portions of protein structure were gauged on the basis of measured distances in the alpha-beta-DNA co-crystal structure.³² Genes encoding these fusion proteins as well as domains derived from the OnTEBP alpha subunit were cloned into bacterial expression

systems. The purity and integrity of expressed proteins were verified by DNA sequencing, SDS-PAGE, and electrospray ionization mass spectroscopy.

Figure 2 shows how fusion proteins (i–iii in Figure 2) and domains (vi and vii) are related to the native alpha (p56a, iv in Figure 2) and beta (p28b, v in Figure 2) subunits of the *O. nova* telomere end binding protein. Systematic short names for each protein indicate the molecular mass in units of kDa and subunits or domains included. The p85ab fusion protein (85.1 kDa, 762 aa) connects the carboxy terminus of alpha with the amino terminus of beta using an eight amino acid peptide linker (aa sequence: SGANSGSG). The C-terminal residue of alpha is well determined by electron density in the crystal structure of the alpha-beta-DNA complex.³² The first eight amino acid residues of the beta subunit are not seen in the crystal structure and likely populate a range of conformations; therefore, the flexible linker connecting alpha and beta subunits in p85ab includes a total of 16 amino acid residues. The DNA encoding this linker contained an EcoRI restriction site to facilitate cloning steps. Otherwise amino acids were selected to be hydrophilic with low restrictions on dihedral angles.

The p85ab fusion protein terminates with residue 260 of beta and thus includes the 28 kDa core domain of this subunit. The N-terminal 260 amino acid residues of beta are necessary and sufficient to confer *in vivo*-like binding properties to the alpha-beta-DNA complex.³¹ The remaining 125 amino acid residues of the full-length beta subunit are not structured as judged by protease mapping (S. C. Schultz, unpublished results), resonance Raman spectroscopy,³⁷ and X-ray crystallography (M.P.H., unpublished results). Through the remainder of this text we will use the name beta when referring to the N-terminal portion of the full-length beta subunit, since this part of the protein retains binding properties of the full-length subunit, including contacts with DNA and protein–protein interactions.^{31,32}

While the p85ab fusion protein contains full-length alpha linked with beta, the other two fusion proteins are each missing the alpha C-terminal domain that otherwise participates in protein–protein interactions with beta. The p66aNb fusion protein (65.9 kDa, 593 aa, ii in Figure 2) connects the N-terminal DNA-binding domain of alpha (aa 1–326α) with the amino terminus of beta using the same eight amino acid linker as used for the p85ab protein. The p61aNbΔ fusion protein (61.3 kDa, 546 aa, iii in Figure 2) connects the N-terminal domain of alpha (aa 1–327α) with the amino terminus of beta using a four amino acid linker (aa sequence: GSAG). The principle difference distinguishing p66aNb from p61aNbΔ is the presence or absence, respectively, of a peptide loop found in beta (aa 156–200β) that interacts with the alpha C-terminal domain. In p61aNbΔ these 45 amino acid residues have been deleted and replaced with a single Asn residue. Including Asn at this position of p61aNbΔ and slight differences in the peptide linker joining the alpha N-terminal domain with beta are details intended to facilitate cloning of DNA encoding p61aNbΔ. In the absence of other proteins, p61aNbΔ and p66aNb bind telomere DNA with similar equilibrium binding constants; therefore, subtle alterations in the linker have, at most, a modest effect on binding properties.

The p37aN protein (36.8 kDa, vi in Figure 2) corresponds to the N-terminal domain of alpha (aa 1–326α) and is similar to the minimal DNA-binding domain identified through deletion mutations of alpha.³¹ The p19aC protein (19.2 kDa, 170 aa, vii in Figure 2) corresponds to the C-terminal domain of alpha (aa 327–495α) and is similar to a previously identified protein domain that permits beta to associate with the alpha–DNA complex.³¹ Amino acid residues 314–327α of the alpha subunit define a natural linker connecting the N-terminal DNA-binding domain with the C-terminal protein-interaction domain.^{31,32} Variants of these alpha-derived domains with slightly different boundary points (e.g. aa 1–314α for alpha-N and aa 316–495α for alpha-C) behaved similarly, meaning that the amino acids contained within the natural linker do not significantly determine binding properties of these domains.

Domain mapping by limited proteolysis

To test whether the alpha and beta derived domains retain a folded structure in the context of these fusion proteins, we challenged each protein with catalytic amounts of chymotrypsin. The N and C-terminal domains of alpha and the 28 kDa core domain of beta correspond to protease-resistant units of protein structure (S. C. Schultz, unpublished results). Protease-catalyzed hydrolysis requires flexible peptide regions flanking the scissile bond, so the appearance of new cleavage sites should be a sensitive indicator of local protein unfolding or misfolding.^{38–40}

Time-course reactions monitoring chymotrypsin treatment of OnTEBP-derived fusion proteins and subunits are shown in Figure 3. Under these conditions, misfolded proteins are completely processed into small molecular mass fragments within two minutes (not shown). The alpha subunit of OnTEBP has one principle site that is rapidly cleaved by chymotrypsin yielding two protease-resistant fragments corresponding to alpha N-terminal (* in Figure 3(a)) and C-terminal domains (*a* in Figure 3(a)). The 28 kDa core domain of beta is, itself, a chymotrypsin-resistant domain (*b* in Figure 3(b)). Each fusion protein retains the single chymotrypsin site present in alpha and no new sites are apparent. Protease-resistant fragments of p85ab correspond to the same alpha N-terminal domain seen with the native alpha subunit (* in Figure 3(c)) and the alpha C-terminal domain linked with beta (*c* in Figure 3(c)). p61aNbΔ digestion yields alpha N-terminal domain (* in Figure 3(d)) and the beta subunit with peptide loop deleted (*d* in Figure 3(d)). Mass spectrometry analysis confirmed that the same chymotrypsin recognition site, Leu324α, is hydrolyzed in p56a, p85ab, p66aNb, and p61aNbΔ (Table 1). Accessible chymotrypsin recognition sites are thus conserved, indicating that alpha and beta-derived domains are similarly folded in the context of these fusion proteins.

Molecularity of DNA–protein complexes

To detect the DNA in DNA–protein complexes, fluorescently-labeled Cy5-d (GGGTTTTGGGG) was used as the single strand telomere DNA extension found at the very ends of the *O. nova* chromosome. In complex with the OnTEBP heterodimer, the unlabeled form of this 11 nt single strand DNA crystallizes with the same space group and unit cell dimensions as found for crystals of the 12 nt d(GGGGTTTTGGGG) DNA-OnTEBP telomere end complex (M.P.H., unpublished result). In the co-crystal structure,³² the DNA adopts a unique fold with the 3' hydroxyl group completely buried but with no clear electron density for the 5'-terminal nucleotide. Accordingly, the fluorophore was attached to the presumably more flexible 5' end of the DNA in order to avoid steric clashes. The 11 nt length of this telomere DNA was chosen to minimize the chance of forming G-quartet stabilized DNA dimers. Such DNA structures are incompatible with binding to the alpha-beta site of OnTEBP yet readily form for *O. nova* telomere DNA sequences that have two complete GGGG repeats. By shortening the 5'-most G-tract to three Gs, symmetric dimers would contain at most two G-quartets instead of the four G-quartets found for *O. nova* telomere DNA in solution⁴¹ and in the crystalline state.^{42,43}

Initial characterization of the alpha-beta–DNA complex showed that telomere DNA combines with alpha and beta subunits with 1:1:1 molecularity.^{30,32} However, in a co-crystal structure of alpha complexed with single strand telomere DNA, alpha–DNA complexes have dimerized along a crystallographic 2-fold axis.³⁴ The dimer interface seen in the alpha–DNA complex is occluded by beta in the alpha-beta–DNA structure.³⁴ By including a portion of beta in each fusion protein, we anticipated that DNA–protein complexes would comprise one fusion protein and one molecule of DNA with no higher-order oligomers.

To determine the molecularity of DNA complexes formed by our fusion proteins, we analyzed DNA-p61aNbΔ and DNA-p85ab complexes by gel filtration chromatography and equilibrium

sedimentation (Figure 4). Figure 4(a) shows that DNA–fusion protein complexes travel through the gel filtration column at a rate expected for a 1:1 DNA–protein species. Calibration curves established for the gel filtration column indicate that 2:2 DNA–protein species, if present, would be resolved from 1:1 counterparts. Figure 4(b) shows representative radial absorbance scans measured for DNA–p61aNbΔ in an ultracentrifuge at three angular velocities. Similar results were obtained using twofold higher and 0.5-fold lower concentrations and for the DNA–p85ab complex. Curve fitting of the data points is consistent with DNA–protein species of 1:1 molecularity (Figure 4(b), continuous curves). Curves calculated assuming a mixture of 90% 1:1 and 10% 2:2 species deviate significantly from the measured data (Figure 4(b), broken curves), meaning that higher-order DNA–protein complexes, if they exist, are present in low abundance. Equilibrium sedimentation analysis of protein alone, without DNA, was also consistent with monomeric fusion protein in the 2–10 μM range.

DNA-binding performance

DNA binding was monitored as a function of fusion protein concentration using agarose gel electrophoresis mobility to distinguish between free DNA and protein-bound species (Figure 5(a)). The fluorophore labeled Cy5-d(GGGTTTGGGG) single strand DNA carries a negative charge as a consequence of 11 phosphate groups, while the p85ab and p61aNbΔ fusion proteins carry excess positive charge; therefore, the free DNA travels towards the anode, while the DNA–protein species travel towards the cathode. In this experiment the DNA concentration is well above the dissociation constant, conditions suitable for determining the fraction of protein active for DNA binding. The endpoint of each titration seen in Figure 5(a) corresponds to 1:1 DNA:protein stoichiometry, indicating 100% DNA binding activity.

To obtain a measure of DNA binding affinity, anisotropy of the Cy5 fluorophore-labeled telomere DNA was monitored as a function of fusion protein concentration (Figure 5(b)). The free DNA species, predominant at low protein concentration, tumbles more quickly and exhibits a lower anisotropy than does the DNA–protein complex, predominant at high protein concentration. Since the free and protein-bound species are in solution together during anisotropy measurements, this technique enables determination of binding constants under true equilibrium conditions.

Figure 5(b) compares binding isotherms for the three fusion proteins, p85ab (square symbols), p66aNb (filled circles), and p61aNbΔ (open circles). Dissociation constants and corresponding free energies of DNA binding are reported in Table 2. Although each fusion protein contains the DNA-binding portions of native alpha and beta subunits, there is a dramatic difference in DNA affinity. The full length alpha-beta fusion protein (p85ab) binds DNA with a K_{D-DNA} of 1.4 nM ($\Delta G_{DNA} = -11.9 \text{ kcal mol}^{-1}$) while the more minimal fusion proteins that lack an alpha C-terminal domain (p66aNb and p61aNbΔ) bind DNA with a K_{D-DNA} of 220–290 nM ($\Delta G_{DNA} = -8.9 \text{ kcal mol}^{-1}$). In terms of free energy of DNA binding, there is a +3 kcal mol⁻¹ destabilization of the DNA complex as a consequence of removing the alpha-derived domain involved in protein–protein association. In fact, K_{D-DNA} measured for p66aNb or p61aNbΔ fusion protein is close to the dissociation constant measured for alpha N-terminal domain (p37aN) or alpha (p56a) alone in the absence of beta (Table 2).

Considering beta provides several key DNA contacts in the OnTEBP–DNA complex, it is puzzling that we measure no energetic benefit for including beta in the alpha N-terminal domain fusion proteins. One might suspect that DNA fails to interact with beta-derived portions of structure in the context of p66aNb or p61aNbΔ. However, a recently obtained co-crystal structure of p61aNbΔ complexed with DNA provides clear-cut evidence for beta–DNA interactions, at least in the crystalline state (M.P.H., unpublished results). Crystals of p61aNbΔ complexed with DNA grow in a $P2_12_12_1$ space group and diffract X-rays to the 2 Å resolution limit. Preliminary $|F_o| - |F_c|$ electron density maps calculated with DNA omitted from

the $|F_c|$ term show that beta makes the same DNA contacts as seen for the native alpha-beta-DNA complex that crystallizes in a different $P6_122$ space group.³² Of course, energy terms introduced by crystal packing contacts have likely selected one particular DNA-protein complex from a mixture in solution. The new crystal structure nevertheless shows that DNA can productively occupy the alpha-beta site in the context of an alpha N-terminal domain-beta fusion protein.

Binding linkage

The difference in DNA binding affinity measured for p85ab relative to DNA affinity measured for p61aNbΔ or p66aNb may relate to missing protein-protein interactions or, alternatively, these fusion proteins may simply bind more weakly because of strain or steric clashes introduced by the unnatural manner in which DNA-binding domains were connected. To rule out the latter possibility, we repeated the p66aNb-DNA titrations at several concentrations of p19aC, the protein domain missing in these fusion proteins. Fang and co-workers had previously demonstrated that the alpha C-terminal domain could be added as a separate protein together with alpha N-terminal domain and beta to reconstitute the OnTEBP-DNA complex found *in vivo*.³¹ We, therefore, reasoned that adding p19aC to the p66aNb fusion protein should complete protein-protein interactions and restore high-affinity DNA binding.

Figure 6(a) shows DNA binding titration curves obtained for p66aNb alone and together with p19aC. The apparent bimolecular association constant obtained if each isotherm is analyzed individually approaches a limiting value as the concentration of p19aC increases from 0 μM to 32 μM (Figure 6(d)). The binding data could be satisfactorily described by a linkage cycle defined by three equilibrium constants (equation (1)). In this model, p66aNb (A in equation (1)) has an intrinsic affinity for DNA measured as $\Delta G_1 = -8.9 \text{ kcal mol}^{-1}$ (Table 3). p66aNb also interacts with the alpha C-terminal domain (B in equation (1)) with an affinity characterized by $\Delta G_2 = -7.1 \text{ kcal mol}^{-1}$. The DNA and protein-binding sites on A (p66aNb) interact cooperatively so that DNA affinity increases by $\Delta\Delta G = -2.7 \text{ kcal mol}^{-1}$ if B (p19aC) saturates the protein-binding site. Since the binding equilibria are linked into a closed cycle, DNA binding with A likewise enhances interaction of A with B by this same $\Delta\Delta G$ value. The overall stability of the full A·B·DNA complex is given by

$$\Delta G_3 = \Delta G_1 + \Delta G_2 + \Delta\Delta G = -18.6 \text{ kcal mol}^{-1}$$

(Table 3), in excellent agreement with values measured by isothermal titration calorimetry (P.B., unpublished results):

(1)

Adding alpha C-terminal domain to the DNA–complex, either in *trans* as described for the p66aNb-p19aC-DNA binding system or in *cis* as part of the p85ab alpha-beta fusion protein, restores high-affinity interaction with DNA (Figures 5(b) and 6(d)), presumably by reconstituting protein–protein interactions. Adding exogenous alpha C-terminal domain (p19aC) had no effect on the DNA binding performance of p61aNbΔ (Figure 6(c)), consistent with the idea that protein–protein interactions that enhance DNA affinity involve the extended peptide loop of beta, a unit of structure deleted in p61aNbΔ. The minimal p61aNbΔ fusion protein binds DNA with $\Delta G_{\text{DNA}} = -8.9 \text{ kcal mol}^{-1}$ (Table 2), a value that corresponds to A interacting with DNA ($\Delta G_1 = -8.9 \text{ kcal mol}^{-1}$; equation (1) and Table 3), without any additional linkage through protein–protein interactions. The full alpha-beta p85ab fusion protein binds DNA with $\Delta G_{\text{DNA}} = -11.9 \text{ kcal mol}^{-1}$ (Table 2), and this value corresponds closely to the energy of a pre-formed A·B complex binding with DNA ($\Delta G_1 + \Delta \Delta G = -11.6$; equation (1) and Table 3). p61aNbΔ and p85ab fusion proteins thus represent limiting cases of the OnTEBP binding system (Figure 6(d)).

Previous studies showed that alpha and beta act cooperatively to bind telomere DNA.³¹ To see how binding linkage for the p66aNb-p19aC–DNA binding system compares with the behavior of native subunits, we evaluated the binding of alpha to single strand telomere DNA at several concentrations of beta spanning the range 0–20 μM (Figure 6(b) and (d)). Equation (1) predicted the response of the alpha-beta-DNA binding system to changes in alpha and beta concentration, and thermodynamic parameters measuring binding linkage for native alpha and beta subunits were, within experimental error, the same as binding linkage parameters determined for alpha C-terminal domain (p19aC) and alpha N-terminal domain-beta fusion protein (p66aNb). Apparently units of protein structure in the *O. nova* telomere end binding

protein are modular and retain their respective binding properties regardless of the exact manner in which they are connected.

Discussion

An initial motivation for creating telomere end binding fusion proteins was to facilitate analysis of protein–DNA interactions by converting the ternary *O. nova* alpha–beta–DNA binding system into a simpler two-component system. Additionally, alternate designs of the original telomere DNA–protein architecture would test the role played by various units of protein structure. Here we focused on the protein–protein interactions seen in the alpha–beta–DNA crystal structure^{32, 43} mediated by the C-terminal domain of alpha and an extended peptide loop of beta. We originally hypothesized that these protein–protein interactions merely hold the DNA-binding portions of each subunit in close proximity, a job that could be played equally well by a peptide linker joining the N-terminal domain of alpha with beta.

In each fusion protein, domains were well folded, as indicated by resistance to protease digestion (Figure 3). As expected, each fusion protein was active in binding single strand telomere DNA and formed a complex consisting of one DNA molecule and one protein (Figures 4 and 5). The dissociation constant measured for the p85ab fusion protein that connects full length alpha with beta ($K_{D-DNA} = 1.4$ nM) is in accord with the dissociation constant previously determined for native alpha and beta subunits³¹ ($K_{D-DNA} = 2$ nM²), indicating that the peptide linker joining alpha with beta has not introduced strain or steric clashes that hinder DNA binding. Telomere DNA co-crystallizes with p85ab in the same space group and with the same unit cell dimensions as does telomere DNA bound by individual alpha and beta subunits (M.P.H., unpublished result), further indication that the p85ab fusion protein forms a telomere DNA complex that closely represents the native alpha–beta–DNA complex.

The fusion proteins p66aNb and p61aNb Δ , constructed with alpha N-terminal domain connected to beta, bind single strand DNA with lower affinity compared with the p85ab fusion protein (Figure 5(b) and Table 2), forcing us to reconsider our original hypothesis that protein–protein interactions can be replaced with a peptide linker. High affinity DNA binding was restored, in the case of p66aNb, by adding alpha C-terminal domain (p19aC) as a separate protein. Early characterization of the *O. nova* telomere binding protein demonstrated that DNA-binding and protein-association domains of alpha could be separated and added back together to reconstitute the telomere DNA–protein complex seen *in vivo*.³¹ Here we reproduce the phenomena of “protein complementation in *trans*”,³¹ measure thermodynamic parameters describing how protein–protein association is linked to DNA–protein complex stability (equation (1) and Table 3), and show that binding linkage requires the 45 amino acid residue protein–interaction peptide loop of beta (Figure 6(c)).

The alpha C-terminal domain, p19aC, acts cooperatively with p66aNb to enhance stability of the DNA–protein complex in two ways: (1) by binding with p66aNb ($\Delta G_2 = -7.1$ kcal mol⁻¹), and (2) by enhancing the intrinsic DNA-binding energy of p66aNb ($\Delta\Delta G = -2.7$ kcal mol⁻¹). Beta shows nearly identical behavior, acting cooperatively with the native alpha subunit (p56a) to enhance stability of the DNA–protein complex (Figure 6 and Table 3). That the measures of binding linkage describing these two A, B, DNA binding systems is within experimental uncertainty the same (Table 3) is a surprising result since beta, in addition to making protein–protein interactions with alpha, also makes direct contacts with five nucleotides of the single strand telomere DNA,³² while the alpha C-terminal domain (p19aC) only makes protein–protein interactions at a site remote from the DNA interface. Apparently, the energetic benefit of adding beta to the DNA–protein complex is largely derived from cooperativity between protein–protein and DNA–protein association ($\Delta\Delta G$ of equation (1)) as opposed to direct contacts made by beta and DNA.

Protein–protein interactions could exert an allosteric effect on the DNA-binding domains of alpha and beta either to optimize DNA contacts or to reduce the entropic cost of forming a three-component DNA–protein complex. Allosteric control of protein function is commonly thought of in terms of small molecules interacting with multi-meric proteins with symmetric disposition of subunits (e.g. protons modulating oxygen binding capacity of hemoglobin,^{44, 45} and ATP/CTP influencing catalytic efficiency of aspartate trans-carbamylase⁴⁶). In the *O. nova* telomere end complex, large units of protein structure act on a heterodimer to influence DNA–protein stability.

How might linkage of protein–protein interactions and DNA–protein stability be related to biological function of a telomere capping protein? The native alpha-beta–DNA complex dissociates slowly with a half-life measured in days.³⁰ Dissociation of the telomere end complex on such a time scale is incompatible with replication of DNA and elongation of telomeres during *O. nova* vegetative cell division. Indeed, telomere DNA protected in an alpha-beta–DNA complex does not serve as a substrate for telomerase *in vitro*.⁴⁷ Protein–protein contacts at a remote site may provide a trigger for facilitating telomere protein–DNA complex dissociation (Figure 7). Post-translational modification such as proteolysis (Figure 7), phosphorylation (as previously suggested⁴⁸), or other agents such as chaperones that weaken protein–protein association would, through allosteric interactions proposed here, have the consequence of weakening DNA–protein association. The resulting complex would become isoenergetic with DNA-alpha species previously discussed as potential intermediates in a step-wise telomere end complex assembly pathway.³³ Rapid inter-conversion of these intermediates could accelerate exchange of single strand DNA from a very stable protective protein complex to a telomere elongation complex with telomerase.

Cooperative interactions and binding linkage may be shared features of telomeres in diverse organisms. Proteins that have sequence similarity with the alpha N-terminal domain of *O. nova* have been identified in *Euplotes crassus*,⁴⁹ fission yeast and vertebrates.¹⁶ The related proteins from fission yeast and vertebrates, called POT1 for protection of telomeres, bind single strand telomere DNA^{50,51} with structurally conserved OB-folds^{52–54} first encountered as telomere DNA binding units in OnTEBP³² and subsequently in CDC13, an orthologous single strand telomere DNA binding protein from budding yeast.⁵⁵ OB-folds in OnTEBP provide DNA-interacting modules in the form of the alpha N-terminal domain and beta subunit as well as the protein-interaction alpha C-terminal domain.³² Although the structure and behavior of POT1 and CDC13 are distinct from OnTEBP in detail, it is nevertheless interesting to speculate that the common use of OB-folds in these single strand telomere DNA-binding proteins^{56,57} hints at a general use of binding linkage in telomere end capping systems.

While proteins related to the OnTEBP alpha subunit appear to be nearly ubiquitous, homologs of the OnTEBP beta subunit have not been found. Telomere capping proteins like POT1 may have evolved homotypic linkage properties as measured for the fission yeast POT1 *in vitro*,⁵⁰ or some other protein may substitute for beta in heterotypic interactions. A POT1 truncation mutant missing its DNA-binding OB-fold is still recruited to human telomeres,⁵⁸ even though this POT1ΔOB protein no longer binds telomere single strand DNA *in vitro*.⁵¹ The implication is that, in addition to DNA interactions, protein–protein interactions play a role in POT1 telomere localization. PTO1 is a recently identified protein that interacts directly with POT1 and also participates in a complex composed of several telomere associated proteins.⁵⁹ PTO1–POT1 interactions may regulate or facilitate POT1's function at the telomere. Intriguingly, yeast two-hybrid analysis mapped the POT1–PTO1 interacting regions to a C-terminal region of POT1 and an internal segment of PTO1,⁵⁹ much like the organization of OnTEBP protein–protein interaction elements in alpha and beta.

Association of OnTEBP alpha and beta subunits is DNA-dependent,³⁰ and we suspect that interaction networks in other telomeres are likewise shaped by telomere DNA. In most eukaryotes, the double strand telomere DNA and single strand DNA extension are longer and more variable in length (e.g. 5–10 kilobase-pairs in humans) than the precisely defined telomeres of mature macronuclear chromosomes in *Oxytricha* and *Euplotes*. These longer telomeres form a lariat structure with single strand telomere DNA invading the duplex portion as shown by electron microscopy.^{20,21} The looped structure, called a t-loop, is promoted by TRF2,⁶⁰ a protein that binds double strand telomere DNA.¹¹ Certain mutant alleles of TRF2 exhibit loss of the 3'-terminal single strand extension¹⁰ and stochastic removal of t-loop sized circular DNA through recombination,⁶¹ underscoring the importance of TRF2 for telomere end structure. While TRF2 does not bind single strand telomere DNA on its own, TRF2 does associate with single strand d(TTAGGG)₄ DNA together with POT1.⁶² POT1 has properties appropriate for binding with the strand of DNA displaced at the base of a t-loop in addition to protecting the 3'-terminal single strand telomere DNA extension.⁵¹ POT1 and TRF2 thus may be coordinating with each other to construct, maintain, and disassemble the special t-loop DNA structure.⁶² Binding linkage networks similar to the one described for OnTEBP subunits could provide a molecular mechanism that enables POT1, TRF2, and other telomere proteins to exchange among functionally important telomere end structures.

Methods

Cloning, expression, and purification of proteins

DNA encoding the p85ab fusion protein was prepared by PCR in two parts. The alpha subunit-encoding part was generated by amplifying an alpha gene template with oligos TAAGAAGGAGATATACATATGTCCACTGCCGCTAAG and GAACCTGAATTCGCACCGGAGTA GATGAGCTTGGTGTC. The beta subunit-encoding part was generated by amplifying a beta gene template with oligos GTGCGAATTCAGGTTCTGGTTCCAAAGGCG CATCTGC and GCAGCCTGATCATTAGAGCTTCTTTCTTTGG. These oligos removed the natural stop and start codons and added the nucleotides needed to encode the eight amino acid residue linker joining alpha and beta subunits. Each oligo also added a specific restriction site to each end of the PCR product. Restriction sites in each oligo are underlined. The resulting PCR products were treated with restriction enzymes (NEB): NdeI/EcoRI in the case of the alpha-encoding fragment and EcoRI/BclI in the case of the beta-encoding fragment. These two pieces of DNA were ligated with pET24a (Novagen) plasmid DNA that had been prepared with NdeI and BamHI restriction enzymes.

DNA encoding the p66aNb fusion protein was similarly prepared and cloned into pET9a (Novagen) plasmid DNA prepared with NdeI and BamHI restriction enzymes. The alpha N-terminal domain-encoding DNA fragment was generated by amplifying an alpha gene template with oligos GGAGATATACATATGTCCACTGCCGCTAAG and GAACCTGAATTCGCACCGGACTTCTTGAGAGAGGCGACC. The beta encoding DNA fragment was generated by amplifying a beta gene template with oligos GTGCGAATTCAGGTTCTGGTTCCAAAGGCGCATCTGC and GCAGCCGGATCCTTAGAGCTTCTTTCTTTGG. DNA fragments were prepared using the same combination of restriction enzymes as used in cloning the p85ab gene except that the beta-encoding fragment was treated with an EcoRI/BamHI combination.

DNA encoding the p61aNbΔ fusion protein was prepared in three parts. The alpha N-terminal domain encoding part was generated by amplifying an alpha gene template with the oligo pair TAAGAAGGAGATATACATATGTCCACTGCCGCTAAG/TGGAACCTGCAGAACCGTTCTTCTTGAGAGAGGC. The parts encoding residues 2–155β and residues 201–260β of beta were generated by separately amplifying a beta gene

template with oligo pairs ACGGTTCTGCAGGTTCCAAAGGCGCATCTGC/GTCGGAATTCGCGGTGTGTCTGAAGTG and CCGCGAATTCGACTTCTCCTTCAAG/GTTAGCAGCCTGATCATTAGAGCTTCTTTCTTTGG, respectively. These oligos encoded a four amino-acid residue linker joining the alpha N-terminal domain with beta and an Asn residue joining regions of beta flanking the extended peptide loop (residues 156–200β). The three DNA fragments were treated with appropriate pairs of restriction enzymes: NdeI/PstI for fragment I, PstI/EcoRI for fragment II, and EcoRI/BclI for fragment III prior to ligation with pET9a vector DNA prepared with NdeI/BamHI digestion.

DNAs encoding the alpha subunit-derived domains p37aN and p19aC were prepared as single PCR generated fragments using oligo pairs

TAAGAAGGAGATATACATATGAACGTCAGCCTCAACGC/

GCCGGATCCTTACTTCTTGAGAGAGGCGACC and

GGAGATATACATATGAACGTCAGCCTCAACGC/

GTTAGCAGCCGGATCCTTAGTAGATGAGCTTGGTG, respectively, to amplify an alpha gene template. These DNAs were each prepared with NdeI/BamHI digestion prior to ligation with pET24a (p37aN encoding construct) or pET9a (p19aC encoding construct).

Expression constructs were propagated as plasmid DNA in DH5α grown in LB media supplemented with 40 µg/ml of kanamycin. DNA encoding each protein was sequenced. For expression of protein, plasmid DNA expression constructs were transformed into BL21(DE3) pLys-S hosts and grown at 37 °C shaking at 200 rpm in two-liter baffles in 750 ml of 2×YT media supplemented with 5 mM glucose, 30 mM potassium phosphate (pH 7.8), 17 µg/ml of chloramphenicol, and 40 µg/ml of kanamycin. When the absorbance at 600 nm reached 1.1, cultures were chilled to room temperature (22 °C) and induced with 0.5–1.0 mM IPTG. After 6–12 hours of continued shaking at room temperature, bacteria were harvested by low-speed centrifugation, washed with lysis buffer (50 mM Hepes (pH 7.5), 50 mM NaCl, 1 mM EDTA, 0.02% (w/v) sodium azide, 2 mM DTT), and frozen at –20 °C. Crude protein extracts were prepared by thawing and suspending cells in 200 ml of lysis buffer prior to sonication (Misonix) on ice. Cell debris was separated from soluble protein by centrifugation at 13,000 rpm in a Sorvall SLA-1500 rotor. Proteins were fractionated by stepwise ammonium sulfate precipitation (35% and 75% saturating amounts of ammonium sulfate). The material precipitating with 75% saturating ammonium sulfate was resuspended in dialysis buffer (25 mM Hepes (pH 7.5), 100 mM NaCl, 0.25 mM EDTA, 0.02% sodium azide, 1 mM DTT) and dialyzed overnight against the same buffer. The dialysate was clarified by centrifugation at 18,000 rpm in a Sorvall SS34 rotor followed by filtration through low protein-binding filters of pore size 0.4 µm and 0.22 µm (Nalgene or Millipore).

Proteins were purified to homogeneity as described for native alpha and beta subunits³² through ion-exchange and size exclusion chromatography using sequentially SP-Sepharose HP, Mono-S, and Superdex-75 or Superdex-200 columns (Amersham/Pharmacia). Proteins bound to cation exchange columns equilibrated with buffer A (25 mM Hepes (pH 7.5), 100 mM NaCl, 0.1 mM EDTA, 0.02% sodium azide, 1 mM DTT) and were eluted with a 0–60% gradient replacing buffer A with buffer B (same composition as buffer A except 1 M in NaCl). Size exclusion chromatography was carried out in Superdex Buffer (5 mM Tris (pH 7.5), 50 mM NaCl, 0.1 mM EDTA, 0.02% sodium azide, 1 mM DTT) or binding buffer (50 mM Tris (pH 7.5), 150 mM NaCl, 0.02% sodium azide). Fractions enriched in telomere binding protein were identified through SDS-PAGE analysis. Pure protein was concentrated to 10–20 mg/ml and stored on ice. Protein concentration and yield was determined for each protein using absorbance at 280 nm and extinction coefficients calculated according to the number of tyrosine and tryptophan residues.⁶³

Domain mapping by chymotrypsin digestion

AEBSF, a serine protease suicide inhibitor, was obtained from Boehringer-Mannheim. Chymotrypsin (α -chymotrypsin, TLCK-treated, type VII) was obtained as a lyophilized powder from Sigma, dissolved in a buffer containing 50% (v/v) glycerol, 50 mM Tris (pH 8), 50 mM NaCl, 2 mM calcium chloride, and stored as a 2 mg/ml stock solution at -20°C in small aliquots. Working solutions of 200 $\mu\text{g/ml}$ chymotrypsin were prepared by dilution in a buffer of this composition without glycerol. Fusion proteins or native OnTEBP subunits were diluted to 1 mg/ml in Superdex Buffer. Chymotrypsin reactions (1:250 mass ratio for protease:target protein) were incubated at 25°C and stopped by adding SDS-containing loading buffer and immediate boiling prior to SDS-PAGE on a 12% (w/v) acrylamide gel.⁶⁴ Alternatively, reactions were stopped by adding 40 $\mu\text{g/ml}$ AEBSF and then samples were dialyzed against water and freeze dried prior to electrospray ionization-mass spectrometry analysis.

Gel electrophoresis mobility shift binding assay

Cy5-d(GGGTTTTGGGG) DNA was prepared and HPLC purified by the University of Utah DNA/peptide synthesis facility. Concentration of the DNA was determined by the molar extinction coefficient of Cy5 at 648 nm, $250,000\text{ M}^{-1}\text{ cm}^{-1}$. Binding reactions containing 200 pmol of Cy5-d(GGGTTTTGGGG) DNA and increasing amounts of protein (0, 20, 40, 80, 120, 160, 200, 400, 800, 1600, and 2000 pmol) were incubated for two hours at room temperature in a buffer composed of 10 mM Tris (pH 7.5) and 10 mM NaCl prior to electrophoresis in a 2% agarose, 1 $\times$ TAE gel at 70 V for 60 minutes. An image of the gel was obtained with a Typhoon 8600 fluorescence imager (Amersham).

Equilibrium analytical ultracentrifugation

DNA-protein complexes were prepared by combining 1.5 molar equivalents of Cy5-d (GGGTTTTGGGG) with fusion protein prior to gel filtration chromatography on a Superdex-75 column using binding buffer as the mobile phase (Figure 4(a)). Absorbance measures at 260 nm, 280 nm and 648 nm of the peak corresponding to a DNA-protein complex were consistent with 1:1 proportion of Cy5-labeled DNA and protein. Protein samples without added DNA were also exchanged into binding buffer by gel filtration chromatography. Matched solutions of buffer alone were prepared by collecting eluted solvent in fractions just prior to fractions containing macromolecules. Serial dilutions of protein and of DNA-protein complexes were prepared by mixing with matched buffer. Samples were loaded into cells fitted with quartz 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 and 648 nm, 0.001 cm step increments, at rotation speeds ranging from 16,000–23,000 rpm. Each scan was the average of ten replicates. Equilibrium was established by comparing scans obtained at four-hour intervals. The radial range where light-bending artifacts were not present was determined by examining radial scans using incident light at 400 nm. In calculating model curves, the partial specific volume of each protein or protein-DNA complex was set at 0.757 based on estimates from the alpha-beta-DNA crystal structure. The density of binding buffer 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 non-linear least-squares algorithm⁶⁵ to optimize parameters for molecular mass and A_0 in equation (2):

$$A_{\text{calculated}} = A_0 \exp \left[\frac{M(1 - v\rho)\omega^2(r^2 - r_0^2)}{2R \times T} \right] \quad (2)$$

where R is $8.3145 \times 10^7 \text{ erg mol}^{-1} \text{ K}^{-1}$. To calculate model curves expected assuming a certain fraction of the absorbing species are in a 2:2 complex (broken curves in Figure 4(b)), equation

(2) was modified to include additional terms with each term weighted by the corresponding relative species abundance.

Fluorescence anisotropy titrations

Proteins were serially diluted in binding buffer to prepare working solutions related in protein concentration by factors of 1/6. An absorbance scan of the 1/36 dilution permitted determination of the protein concentration in each working solution. Lyophilized Cy5-d (GGGTTTTGGGG) DNA was suspended in a buffer composed of 2 mM Tris (pH 8) and 0.1 mM EDTA to give an ~200 μ M DNA stock solution. On the day of an experiment the DNA was diluted in binding buffer and the DNA concentration was determined using absorbance at 648 nm. A working solution of known DNA concentration was prepared by mixing 4 μ l of the diluted DNA solution with 1.35 ml of Binding Buffer directly in a quartz fluorescence cell (Helma). Titrations were carried out at 20 °C. Beginning with the most dilute protein solution, 4 μ l aliquots were added to the DNA solution and mixed by gentle repeated pipetting. Fluorescence anisotropy measures were obtained with a spectrofluorimeter (ISS, Urbana Illinois) using incident light at 640 nm (2 mm slit width, fwhm = 16 nm). Scattered light was removed from detected fluorescence with the use of a 671 nm bandpass (full width at 1% transmittance = 20 nm) interference filter (Edmund Industrial Optics) placed between the sample and the photomultiplier. Light intensity was integrated for five seconds and typically ranged from 1000 photons to 10,000 photons. The average of three to eight anisotropy readings was recorded for each protein concentration. Titrations obtained with high concentrations of DNA and attenuated light intensities confirmed that each protein was 100% active in DNA binding.

Data from multiple titrations were simultaneously analyzed with use of a program written in C (M.P.H.) that employed a non-linear least-squares algorithm⁶⁵ to optimize parameters describing an apparent dissociation constant (K_{D-DNA}) and the characteristic anisotropy of free DNA and the DNA–protein complex. For each value of K_{D-DNA} and total protein concentration, the amounts of free DNA, DNA–protein complex, and free protein that satisfy equations (3)–(5) were found through numerical methods. Equation (3) is a definition of chemical equilibrium and equations (4) and (5) relate to conservation of mass. This approach is equivalent to solving the quadratic form of a two-state partition function but can be more easily generalized to the linked binding system of equation (1):

$$K_{D-DNA} = [\text{DNA}_{\text{free}}][\text{PTN}_{\text{free}}]/[\text{DNA} - \text{PTN}] \quad (3)$$

$$\text{DNA}_{\text{total}} = [\text{DNA}_{\text{free}}] + [\text{DNA} - \text{PTN}] \quad (4)$$

$$\text{PTN}_{\text{total}} = [\text{PTN}_{\text{free}}] + [\text{DNA} - \text{PTN}] \quad (5)$$

Parameter values describing free energy of DNA binding were invariant with respect to increases in DNA concentration in the range of 0.8 nM–3.2 nM, indicating that the data analysis was robust even though the total DNA concentration was only slightly lower than the measured dissociation constant in some cases.

To obtain an estimate of uncertainty in each parameter's value, the analysis was repeated for a number of simulated trials. In each trial, re-sampling of the original N measurements with replacement generated a new set of N data. In a given simulation, an original data point is represented zero, one, two, and very rarely three times. Uncertainties determined as the sample

standard deviation of parameter values optimized for each of 40 simulated trials by this bootstrap method were comparable to uncertainties determined by repeating the titration experiment in triplicate as carried out for the p85ab fusion protein (see Figure 5(b)).

Binding linkage analysis

To obtain linkage parameters reported in Table 3, fluorescence anisotropy was measured for increasing concentrations of p66aNb in the presence of known concentrations of p19aC and 1 nM Cy5-d(GGGTTTTGGGG). The initial anisotropy reading in the absence of added p66aNb did not vary systematically with respect to concentration of p19aC, indicating that p19aC does not bind telomere DNA on its own. A total of eight titrations were performed with the concentration of p19aC spanning 0.1 μM –32 μM . For a ninth titration no p19aC protein was added. Anisotropy measures were analyzed in terms of three dissociation constants (K_{D-1} , K_{D-2} , and K_{D-3} of equation (1)). Parameters for these three binding constants and the anisotropy values characterizing the three forms of DNA (A_{DNA} , the DNA free in solution; $A_{\text{A-DNA}}$, DNA complexed with p66aNb; and $A_{\text{A-B-DNA}}$, DNA in the ternary p66aNb–p19aC–DNA complex) were adjusted to minimize the sum of squared residuals between measured anisotropy values and values calculated according to equation (6):

$$\text{Anisotropy}_{\text{calculated}} = A_{\text{DNA}}(d) + A_{\text{A-DNA}}(e) + A_{\text{A-B-DNA}}(f) \quad (6)$$

where d is the fraction of total DNA free in solution, e is the fraction of total DNA complexed with p66aNb, and f is the fraction of total DNA in the ternary p66aNb–p19aC–DNA complex. For given total concentrations of p66aNb, p19aC, and DNA, values of d , e , and f that satisfy equations (7)–(9) were obtained through numerical methods. Equations (7)–(9) are obtained by substituting the chemical definitions of K_{D-1} , K_{D-2} , and K_{D-3} into conservation of mass equations for the binding agents, DNA, p66aNb, and p19aC:

$$1 = d + e + f \quad (7)$$

$$[\text{p66aNb}_{\text{total}}]/[\text{DNA}_{\text{total}}] = (K_{D-1}/[\text{DNA}_{\text{total}}])(e/d) + (K_{D-3}/(K_{D-2}[\text{DNA}_{\text{total}}]))(f/d) + e + f \quad (8)$$

$$[\text{p19aC}_{\text{total}}]/[\text{DNA}_{\text{total}}] = (K_{D-3}/(K_{D-1}[\text{DNA}_{\text{total}}]))(f/e) + (K_{D-3}/(K_{D-2}[\text{DNA}_{\text{total}}]))(f/d) + f \quad (9)$$

Data measured for eight titrations of p56a binding to DNA with concentration of p28b varying from 0 μM to 20 μM were similarly analyzed with p56a substituting for p66aNb (A of equation (1)), and p28b substituting for p19aC (B of equation (1)).

Acknowledgments

We thank L. Joss for help with equilibrium sedimentation experiments, D. Blumenthal for use of his spectrofluorimeter, S. Kojima for calibrating the size exclusion column, R. Schackmann for DNA synthesis, M. Hoff for DNA sequencing, and C. Nelson and P. Krishna for mass spectrometry. We appreciate critical reading of the manuscript by S. Classen, D. Blair and D. Goldenberg as well as constructive comments provided by anonymous reviewers. M.P.H. thanks M. T. K. Horvath for advice on non-linear equations. The NIH supported this work through a grant to M.P.H. (R01 GM067994). The Protein-DNA core facility at the University of Utah receives support from the National Cancer Institute (5P30 CA42014).

Abbreviations used

OnTEBP	<i>Oxytricha nova</i> telomere end binding protein
DTT	DL-dithiothreitol
aa	amino acids

References

- Blackburn EH. Structure and function of telomeres. *Nature* 1991;350:569–573. [PubMed: 1708110]
- Lingner J, Cech TR. Telomerase and chromosome end maintenance. *Curr Opin Genet Dev* 1998;8:226–232. [PubMed: 9610414]
- 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;78:3015–3019. [PubMed: 6265931]
- Henderson ER, Blackburn EH. An overhanging 3' terminus is a conserved feature of telomeres. *Mol Cell Biol* 1989;9:345–348. [PubMed: 2927395]
- McElligott R, Wellinger RJ. The terminal DNA structure of mammalian chromosomes. *EMBO J* 1997;16:3705–3714. [PubMed: 9218811]
- Kyrion G, Liu K, Liu C, Lustig AJ. RAP1 and telomere structure regulate telomere position effects in *Saccharomyces cerevisiae*. *Genes Dev* 1993;7:1146–1159. [PubMed: 8319907]
- Krauskopf A, Blackburn EH. Rap1 protein regulates telomere turnover in yeast. *Proc Natl Acad Sci USA* 1998;95:12486–12491. [PubMed: 9770512]
- Wahlin J, Cohn M. *Saccharomyces cerevisiae* RAP1 binds to telomeric sequences with spatial flexibility. *Nucl Acids Res* 2000;28:2292–2301. [PubMed: 10871358]
- Li B, Oestreich S, de Lange T. Identification of human Rap1: implications for telomere evolution. *Cell* 2000;101:471–483. [PubMed: 10850490]
- van Steensel B, Smogorzewska A, de Lange T. TRF2 protects human telomeres from end-to-end fusions. *Cell* 1998;92:401–413. [PubMed: 9476899]
- Broccoli D, Smogorzewska A, Chong L, de Lange T. Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2. *Nature Genet* 1997;17:231–235. [PubMed: 9326950]
- Bianchi A, Smith S, Chong L, Elias P, de Lange T. TRF1 is a dimer and bends telomeric DNA. *EMBO J* 1997;16:1785–1794. [PubMed: 9130722]
- 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]
- Price CM, Cech TR. Telomeric DNA-protein interactions of *Oxytricha* macronuclear DNA. *Genes Dev* 1987;1:783–793. [PubMed: 3123321]
- 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;93:13760–13765. [PubMed: 8943008]
- Baumann P, Cech TR. Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science* 2001;292:1171–1175. [PubMed: 11349150]
- Sundquist WI, Klug A. Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. *Nature* 1989;342:825–829. [PubMed: 2601741]
- Parkinson GN, Lee MP, Neidle S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* 2002;417:876–880. [PubMed: 12050675]
- Neidle S, Parkinson GN. The structure of telomeric DNA. *Curr Opin Struct Biol* 2003;13:275–283. [PubMed: 12831878]
- Griffith JD, Comeau L, Rosenfield S, Stansel RM, Bianchi A, Moss H, de Lange T. Mammalian telomeres end in a large duplex loop. *Cell* 1999;97:503–514. [PubMed: 10338214]
- Nikitina T, Woodcock CL. Closed chromatin loops at the ends of chromosomes. *J Cell Biol* 2004;166:161–165. [PubMed: 15249582]

22. de Lange T. T-loops and the origin of telomeres. *Nature Rev Mol Cell Biol* 2004;5:323–329. [PubMed: 15071557]
23. Counter CM, Avilion AA, LeFeuvre CE, Stewart NG, Greider CW, Harley CB, Bacchetti S. Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. *EMBO J* 1992;11:1921–1929. [PubMed: 1582420]
24. Bodnar AG, Ouellette M, Frolkis M, Holt SE, Chiu CP, Morin GB, et al. Extension of lifespan by introduction of telomerase into normal human cells. *Science* 1998;279:349–352. [PubMed: 9454332]
25. Karlseder J, Smogorzewska A, de Lange T. Senescence induced by altered telomere state, not telomere loss. *Science* 2002;295:2446–2449. [PubMed: 11923537]
26. Blackburn EH. Telomere states and cell fates. *Nature* 2000;408:53–56. [PubMed: 11081503]
27. Counter CM. The roles of telomeres and telomerase in cell life span. *Mutat Res* 1996;366:45–63. [PubMed: 8921986]
28. Harley CB, Vaziri H, Counter CM, Allsopp RC. The telomere hypothesis of cellular aging. *Expt Gerontol* 1992;27:375–382.
29. Gray JT, Celandier 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;67:807–814. [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;90:6056–6060. [PubMed: 8327484]
31. Fang G, Gray JT, Cech TR. *Oxytricha* telomere-binding protein: separable DNA-binding and dimerization domains of the alpha-subunit. *Genes Dev* 1993;7:870–882. [PubMed: 8491383]
32. 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;95:963–974. [PubMed: 9875850]
33. 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;314:1113–1125. [PubMed: 11743727]
34. Peersen OB, Ruggles JA, Schultz SC. Dimeric structure of the *Oxytricha nova* telomere end-binding protein alpha-subunit bound to ssDNA. *Nature Struct Biol* 2002;9:182–187. [PubMed: 11836536]
35. Willis MC, Hicke BJ, Uhlenbeck OC, Cech TR, Koch TH. Photocrosslinking of 5-iodouracil-substituted RNA and DNA to proteins. *Science* 1993;262:1255–1257. [PubMed: 7694369]
36. Hicke BJ, Willis MC, Koch TH, Cech TR. Telomeric protein-DNA point contacts identified by photocross-linking using 5-bromodeoxyuridine. *Biochemistry* 1994;33:3364–3373. [PubMed: 8136374]
37. Laporte L, Stultz J, Thomas GJ Jr. Solution conformations and interactions of alpha and beta subunits of the *Oxytricha nova* telomere binding protein: investigation by Raman spectroscopy. *Biochemistry* 1997;36:8053–8059. [PubMed: 9201953]
38. Fontana A, Fassina G, Vita C, Dalzoppo D, Zamaï M, Zambonin M. Correlation between sites of limited proteolysis and segmental mobility in thermolysin. *Biochemistry* 1986;25:1847–1851. [PubMed: 3707915]
39. Hubbard SJ, Campbell SF, Thornton JM. Molecular recognition. Conformational analysis of limited proteolytic sites and serine proteinase protein inhibitors. *J Mol Biol* 1991;220:507–530. [PubMed: 1856871]
40. Hubbard SJ, Eisenmenger F, Thornton JM. Modeling studies of the change in conformation required for cleavage of limited proteolytic sites. *Protein Sci* 1994;3:757–768. [PubMed: 7520312]
41. Smith FW, Feigon J. Quadruplex structure of *Oxytricha* telomeric DNA oligonucleotides. *Nature* 1992;356:164–168. [PubMed: 1545871]
42. Haider S, Parkinson GN, Neidle S. Crystal structure of the potassium form of an *Oxytricha nova* G-quadruplex. *J Mol Biol* 2002;320:189–200. [PubMed: 12079378]
43. Horvath MP, Schultz SC. DNA G-quartets in a 1.86 Å resolution structure of an *Oxytricha nova* telomeric protein-DNA complex. *J Mol Biol* 2001;310:367–377. [PubMed: 11428895]
44. Bohr C, Hasselbalch KA, Krogh A. Ueber einen in biologischer Beziehung wichtigen Einfluss, den die Kohlensäurespannung des Blutes auf dessen Sauerstoffbindung ubt. *Skand Arch Physiol* 1904;16:402–412.

45. Imai K, Yonetani T. PH dependence of the Adair constants of human hemoglobin. Nonuniform contribution of successive oxygen bindings to the alkaline Bohr effect. *J Biol Chem* 1975;250:2227–2231. [PubMed: 234962]
46. Kantrowitz ER, Lipscomb WN. *Escherichia coli* aspartate transcarbamylase: the relation between structure and function. *Science* 1988;241:669–674. [PubMed: 3041592]
47. 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;12:1504–1514. [PubMed: 9585510]
48. Hicke B, Rempel R, Maller J, Swank RA, Hamaguchi JR, Bradbury EM. Phosphorylation of the *Oxytricha* telomere protein: possible cell cycle regulation. *Nucl Acids Res* 1995;23:1887–1893. [PubMed: 7596814]
49. Wang W, Skopp R, Scofield M, Price C. *Euplotes crassus* has genes encoding telomere-binding proteins and telomere-binding protein homologs. *Nucl Acids Res* 1992;20:6621–6629. [PubMed: 1480483]
50. Lei M, Baumann P, Cech TR. Cooperative binding of single-stranded telomeric DNA by the Pot1 protein of *Schizosaccharomyces pombe*. *Biochemistry* 2002;41:14560–14568. [PubMed: 12463756]
51. Loayza D, Parsons H, Donigian J, Hoke K, de Lange T. DNA binding features of human POT1: a nonamer 5'-TAGGGTTAG-3' minimal binding site, sequence specificity, and internal binding to multimeric sites. *J Biol Chem* 2004;279:13241–13248. [PubMed: 14715659]
52. Murzin AG. OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J* 1993;12:861–867. [PubMed: 8458342]
53. Lei M, Podell ER, Cech TR. Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection. *Nature Struct Mol Biol* 2004;11:1223–1229. [PubMed: 15558049]
54. 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]
55. Mitton-Fry RM, Anderson EM, Hughes TR, Lundblad V, Wuttke DS. Conserved structure for single-stranded telomeric DNA recognition. *Science* 2002;296:145–147. [PubMed: 11935027]
56. Theobald DL, Wuttke DS. Prediction of multiple tandem OB-fold domains in telomere end-binding proteins Pot1 and Cdc13. *Structure (Camb)* 2004;12:1877–1879. [PubMed: 15458635]
57. Theobald DL, Cervantes RB, Lundblad V, Wuttke DS. Homology among telomeric end-protection proteins. *Structure (Camb)* 2003;11:1049–1050. [PubMed: 12962623]
58. Loayza D, De Lange T. POT1 as a terminal transducer of TRF1 telomere length control. *Nature* 2003;423:1013–1018. [PubMed: 12768206]
59. Liu D, Safari A, O'Connor MS, Chan DW, Laegeler A, Qin J, Songyang Z. PTOP interacts with POT1 and regulates its localization to telomeres. *Nature Cell Biol* 2004;6:673–680. [PubMed: 15181449]
60. Stansel RM, de Lange T, Griffith JD. T-loop assembly *in vitro* involves binding of TRF2 near the 3' telomeric overhang. *EMBO J* 2001;20:5532–5540. [PubMed: 11574485]
61. Wang RC, Smogorzewska A, de Lange T. Homologous recombination generates T-loop-sized deletions at human telomeres. *Cell* 2004;119:355–368. [PubMed: 15507207]
62. Yang Q, Zheng YL, Harris CC. POT1 and TRF2 cooperate to maintain telomeric integrity. *Mol Cell Biol* 2005;25:1070–1080. [PubMed: 15657433]
63. Gill SC, von Hippel PH. Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem* 1989;182:319–326. [PubMed: 2610349]
64. Schagger H, von Jagow G. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* 1987;166:368–379. [PubMed: 2449095]
65. Marquardt DW. An algorithm for least-squares estimation of non-linear parameters. *J Soc Ind Appl Math* 1963;11:431–441.

Figure 1.

(a) Overview of the telomere DNA-protein complex from *O. nova* (adapted using Protein Data Bank entries 1OTC and 1JB7). The DNA complex is shown with OnTEBP alpha and beta subunits as molecular surface representations. Alpha N-terminal domain (magenta) and beta (cyan) form a deep cleft which binds the single strand 3'-terminal telomere DNA. Double strand B-form DNA has been added to the model. The biological role of the OnTEBP-DNA complex is to cap and protect the very ends of chromosomes. The enlargement shows protein-protein interactions involving the C-terminal domain of alpha (green) and an extended peptide loop of beta (dark blue). (b) The 16-nucleotide single strand DNA binding site of OnTEBP. Contacts with the N-terminal domain of alpha and beta are colored magenta and cyan, respectively. Text labels for nucleotides 9, 10, 11, and 14 are placed at magenta/cyan boundaries, since nucleotides 9, 10, 11, and 14 make contact with both alpha and beta. The DNA seen in OnTEBP co-crystals consists of nucleotides 6–16. A Cy5-labeled form of this d(GGGTTTGGGG) DNA was used for DNA-protein binding experiments.

Figure 2.

Fusion proteins, subunits, and domains used in this study. The OnTEBP alpha (iv) and beta (v) subunits are color coded with alpha N-terminal domain magenta, alpha C-terminal domain green, and beta 28 kDa core domain blue. Residues 314–327 α of alpha constitute a natural linker connecting alpha N and C-terminal domains. The extended protein-interaction peptide loop of beta (aa 155–201 β) is shown in a darker shade of blue. Short names for each fusion protein, subunit, and domain were derived as a combination of the molecular mass (kDa) followed by characters indicating the subunits and domains included (a for alpha, b for beta, aN for alpha N-terminal domain, aC for alpha C-terminal domain, and b Δ for beta with its protein-interaction peptide loop deleted). In the p85ab fusion protein (i) alpha and beta subunits are joined through an eight amino acid residue linker (aa sequence SGANSGSG). In the p66aNb fusion protein (ii), the same eight amino acid residue linker joins alpha N-terminal domain and beta. In the p61aNb Δ fusion protein (iii) a four amino acid residue linker (GSAG) similarly joins alpha N-terminal domain with beta. The main difference between p66aNb and p61aNb Δ is that the extended peptide protein–protein interaction loop of beta is deleted in the later. The alpha N-terminal domain (aa 2–326 α) was expressed as a 37 kDa protein, p37aN (vi), and the alpha C-terminal domain (aa 327–495 α) was expressed as a 19 kDa protein, p19aC (vii). Single letter amino acid abbreviations: A, alanine; G, glycine; N, asparagine; and S, serine.

Figure 3.

Chymotrypsin digestion time-course for p56a, p28b, p85ab, and p61aNb Δ . Chymotrypsin digestion products were analyzed by SDS-PAGE.⁶⁴ Under similar conditions, misfolded proteins are completely processed into small peptides in less than two minutes (data not shown). The mobilities and molecular masses (kDa) of standard proteins are shown on the left. For each of the alpha N-terminal domain-containing proteins (p56a: the native alpha subunit shown in (a), p85ab shown in (c), and p61aNb Δ shown in (d)), chymotrypsin cleaves the peptide bond following amino acid Leu324 α (see Table 1) to give two fragments that resist further processing for >30 minutes. One of these fragments (mass = 36,544 Da, indicated by*) corresponds to the N-terminal amino acid residues 2–324 α . The second protease-resistant fragment corresponds to the C-terminal portion that varies in size according to each protein. These bands are labeled with lower case letters and correspond to the alpha C-terminal domain (*a*), the alpha C-terminal domain fused with beta (*c*), and the beta subunit with peptide loop removed (*d*). Even though there are three leucine and phenylalanine residues in the extended peptide loop comprising residues 156–200 β , the beta 28 kDa core domain is, nonetheless, resistant to chymotrypsin cleavage (band *b* in (b)), suggesting that residues 156–200 β of beta refold into some new structure when not complexed with alpha.

Figure 4.

Size exclusion chromatography and equilibrium sedimentation of DNA–protein complexes are consistent with a 1:1 DNA–protein complex. (a) Fusion proteins p85ab (upper traces) and p61aNbΔ (lower traces) and the DNA complexes formed by each were analyzed by size exclusion chromatography. The elution volumes and molecular masses (in units of kDa) for molecular mass standards are shown above the chromatograms. Protein and DNA were detected by absorbance at 280 nm. The identity of the small molecular mass peak labeled DNA was confirmed by A_{260}/A_{280} ratio and by its blue color, characteristic of the Cy5 fluorophore. The elution volume expected for each 1:1 DNA–protein species is marked as monomer and the corresponding volume expected for a 2:2 DNA–protein species is marked as dimer. Peaks seen in each chromatogram are consistent with monomeric fusion proteins and 1:1 DNA–protein complexes. A slight leading shoulder seen for the p61aNbΔ–DNA complex is less than 3% of the total integrated area. (b) Representative radial absorbance scans for equilibrium sedimentation of the p61aNbΔ–DNA complex at three angular velocities: 23,000 rpm (top), 21,000 rpm (middle), and 19,000 rpm (bottom). Measured absorbance data points (circles) are compared with two model curves calculated assuming either 100% 1:1 DNA–protein species (continuous curve) or a mixture of 90% 1:1 DNA–protein and 10% 2:2 DNA–protein species (broken curve).

Figure 5.

DNA-binding performance of fusion proteins. (a) Agarose gel electrophoresis of Cy5-d (GGGTTTTGGGG) DNA at increasing concentrations of p61aNbΔ (left-hand panel) and p85ab (right-hand panel). The free DNA travels towards the anode (bottom) while the DNA–protein complex travels towards the cathode (top). Both fusion proteins are fully active for binding single strand DNA. (b) Anisotropy of the Cy5-DNA fluorophore measured as a function of p85ab concentration (square symbols), p66aNb (filled circles), and p61aNbΔ (open circles). Three separately obtained titrations are shown for p85ab. Curves are the results of non-linear least-squares fitting of the data to a bimolecular binding reaction. The p66aNb and p61aNbΔ fusion proteins require higher protein concentration to achieve a DNA–protein complex meaning these proteins, that lack an alpha C-terminal domain, bind single strand DNA with diminished free energy of binding (Table 2).

Figure 6.

Binding linkage measured for a “*trans* complementation” p66aNb–p19aC–DNA complex and the native alpha-beta–DNA complex. (a) Binding isotherms for p66aNb binding with Cy5-DNA at various fixed concentrations of p19aC. The concentrations of p19aC are (from bottom to top): 0 μ M (open circles), 0.1 μ M, 0.5 μ M, 2.0 μ M, 8 μ M, and 16 μ M (filled circles). For each binding isotherm, the point obtained when concentration of p66aNb is equal to the apparent dissociation constant is marked with a vertical line. Model curves represent the result of non-linear least-squares fitting of all the data (including three titrations not shown) with equation (1). (b) Binding isotherms measured for the native alpha subunit (p56a) binding with Cy5-DNA at different concentrations of beta (p28b). The concentrations of beta are (from bottom to top): 0 μ M (open circles), 0.2 μ M, 1 μ M, 2 μ M, 10 μ M, and 19 μ M (filled circles). (c) Binding isotherms measured for p61aNb Δ in the absence of (bottom) and presence of 20 μ M p19aC (top). In contrast to p66aNb, the DNA affinity of p61aNb Δ is insensitive to added p19aC. (d) The apparent association constant for A binding with DNA plotted as a function of $[B]$. B is either the alpha C-terminal domain (p19aC, filled circles) interacting with p66aNb (A of equation (1)), or B is the native beta subunit (p28b, open circles) that interacts with alpha (p56a). The alpha C-terminal domain and beta subunit each enhance DNA binding by occupying a site on the DNA–protein complex that saturates with increasing concentration. Modeling of equation (1) with thermodynamic parameters reported in Table 3 generated the fitted curves. Shaded regions highlight limiting values of K_a measured for p85ab (upper limit) and for p61aNb Δ (lower limit).

Figure 7.

Stabilities of DNA–protein complexes suggest a route to facilitate DNA exchange during telomere replication. Each DNA–protein complex is shown with a black curved line representing single strand DNA, a magenta square for the alpha N-terminal domain, a blue shape for the beta subunit, and a green circle for the alpha C-terminal domain. In the absence of beta the DNA binds alpha N-terminal domain at the alpha site (i, ii). When the DNA-binding portion of beta is present (iii, iv, vi, vii), the DNA shifts in register to occupy both the alpha site and the alpha-beta site. As measured in this study, all of the DNA–protein complexes that lack productive protein–protein associating domains (i–iv, vii) bind nearly isoenergetically with single strand DNA ($\Delta G_{\text{DNA}} \sim 9$ kcal/mol). With protein–protein interactions included, the DNA complex is significantly more stable ($\Delta G_{\text{DNA}} \sim 12$ kcal/mol). The high-affinity complex (vi) can be formed by multiple routes, either by combining native alpha and beta subunits with DNA, by binding the alpha-beta (p85ab) fusion protein with DNA, or by combining alpha C-terminal domain and p66aNb fusion protein with DNA. Any mechanism that disrupts the protein–protein interactions (proteolysis is depicted in vii) would result in a destabilized DNA protein complex enabling more rapid exchange of telomere DNA with telomerase.

Table 1

Molecular mass of chymotrypsin digestion products

Protein	Fragment	Measured ^a (Da)	Calculated (Da)	Δ^b (Da)
p85ab	2–324L ^c	36,542.8	36,544.4	+1.6
	K325–762	48,560.0	48,562.9	+2.9
p61aNbΔ	2–324L	36,542.0	36,544.4	+2.4
	K325–546	24,783.3	24,784.1	+0.8
p56a	2–324L	36,542.5	36,544.4	+1.9
	K325–495	19,424.0	19,424.9	+0.9
p28b	2–260	28,538.9	28,538.5	–0.4

^a Mass determined by electrospray ionization mass spectrometry.

^b Δ is the difference between mass predicted based on amino acid sequence and the experimentally measured mass.

^c Proteins expressed in bacteria are commonly processed to remove the initiator methionine residue.

Table 2

Telomere DNA–protein complex stability at 20 °C

Protein	K_{D-DNA} (nM)	ΔG_{DNA} (kcal mol ⁻¹)
p85ab	1.4±0.2	-11.9±0.1
p66aNb	290±40	-8.8±0.1
p61aNbΔ	220±30	-8.9±0.1
p56a	220±20	-8.9±0.1
p37aN	200±10	-9.0±0.1

Table 3

Binding linkage in a telomere DNA–protein complex

A	ΔG_1	B	ΔG_2	$\Delta \Delta G$	ΔG_3
p66aNb	-8.9 ± 0.2	p19aC	-7.1 ± 0.2	-2.7 ± 0.2	-18.6 ± 0.1
p56a	-9.0 ± 0.1	p28b	-7.1 ± 0.2	-2.8 ± 0.2	-18.9 ± 0.1

All energy terms are expressed in units of kcal mol⁻¹. ΔG_1 is the intrinsic free energy of binding protein A with DNA (see equation (1)). ΔG_2 is the free energy of forming an A·B complex without added DNA. $\Delta \Delta G$ is the added stability realized upon completing both protein–protein and protein–DNA interactions to form the A·B·DNA complex. ΔG_3 is the overall stability of A·B·DNA, equal to the sum of ΔG_1 , ΔG_2 , and $\Delta \Delta G$ (equation (1)).