doi:10.1016/j.jmb.2006.02.082 J. Mol. Biol. (2006) 359, 1217–1234

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Thermodynamic Characterization of Binding *Oxytricha nova* Single Strand Telomere DNA with the Alpha Protein N-terminal Domain

Pawel Buczek and Martin P. Horvath*

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

**Corresponding author*

The *Oxytricha nova* telemere binding protein alpha subunit binds single strand DNA and participates in a nucleoprotein complex that protects the very ends of chromosomes. To understand how the N-terminal, DNA binding domain of alpha interacts with DNA we measured the stoichiometry, enthalpy (ΔH), entropy (ΔS), and dissociation constant ($K_{D\text{-}DNA}$) for binding telomere DNA fragments at different temperatures and salt concentrations using native gel electrophoresis and isothermal titration calorimetry (ITC). About 85% of the total free energy of binding corresponded with non-electrostatic interactions for all DNAs. Telomere DNA fragments d(T_2G_4), d(T_4G_4), d($G_3T_4G_4$), and d($G_4T_4G_4$) each formed monovalent protein complexes. In the case of d($T_4G_4T_4G_4$), which has two tandemly repeated d(TTTTGGGG) telomere motifs, two binding sites were observed. The high-affinity ''A site'' has a dissociation constant, $K_{D\text{-}DNA(A)} = 13(\pm 4)$ nM, while the low-affinity ''B site'' is characterized by $K_{D\text{-}DNA(B)} = 5600(\pm 600)$ nM at 25 °C. Nucleotide substitution variants verified that the A site corresponds principally with the 3′-terminal portion of d($T_4G_4T_4G_4$). The relative contributions of entropy (ΔS) and enthalpy (ΔH) for binding reactions were DNA length-dependent as was heat capacity (ΔCp). These trends with respect to DNA length likely reflect structural transitions in the DNA molecule that are coupled with DNA–protein association. Results presented here are important for understanding early intermediates and subsequent stages in the assembly of the full telomere nucleoprotein complex and how binding events can prepare the telomere DNA for extension by telomerase, a critical event in telomere biology.

Keywords: telomere-binding protein; telomere biology; protein–single strand DNA interactions; ITC; binding energetics

Introduction

Telomere DNA consists of tandem repeats of a simple T and G-rich sequence such as d(TTGGGG) in *Tetrahymena thermophila*, d(TTAGGG) in vertebrates and some fungi, and d(TTTTGGGG) in *Oxytricha nova* and *Euplotes crassus*.[1–4] The 3′-terminal strand extends past the duplex DNA creating a single strand DNA extension that can form lariat DNA structures, seen by electron microscopy,[5–9] through interaction with internal double-stranded repeats and proteins in humans, chickens, and developing macronuclei of ciliates. In mature macronuclei of ciliates, where the length of both duplex and single strand telomere DNA is precise and short, single strand DNA binds with telomere end-specific proteins to form a protective nucleoprotein complex.[10–14] Sequence and structural homology among telomere end binding proteins[15] and direct tests using gene deletion studies in yeast[16] or RNA-interference in *Stylonychia lemnae*[17] indicate that telomere end binding proteins are essential for telomere function and cell viability, sequestering telomere DNA and coordinating timely exchange with telomerase.

The simple, well-defined telomeres of ciliates provide a unique opportunity for understanding telomere complex assembly and disassembly.

Abbreviations used: ITC, isothermal titration calorimetry; nt, nucleotides; aa, amino acids; PDB, Protein Data Bank.

E-mail address of the corresponding author: horvath@biology.utah.edu

Oxytricha nova single strand telomere DNA consists of two d(TTTTGGGG) repeats and associates with alpha and beta subunits of the telomere end binding protein, OnTEBP.[12–14] Short telomere DNA fragments form binary complexes with the alpha N-terminal DNA-binding domain. Since the alpha N-terminal domain retains DNA binding properties measured for alpha in the absence of beta,[13,18] these DNA complexes likely represent early intermediates in a step-wise assembly pathway.[19,20] A co-crystal structure of telomere-derived DNA complexed with the alpha N-terminal domain provided atomic details for DNA–protein interactions showing how this DNA-binding domain establishes contacts consistent with 3′-terminal and sequence-specific recognition of DNA.[19,20]

In order to understand early steps in assembly of DNA–protein telomere end structures, we have measured the stoichiometry (n), enthalpy (ΔH), entropy (ΔS), and dissociation constant ($K_{D\text{-}DNA}$) for binding telomere DNA fragments with the alpha protein N-terminal domain at different temperatures and salt concentrations using native gel electrophoresis and isothermal titration calorimetry (ITC). These measures complement and extend what we know about DNA–protein interactions for a sequence-specific single strand DNA binding protein and for T and G-rich DNA substrates. Short $d(T_2G_4)$ and $d(T_4G_4)$ DNAs bound the alpha N-terminal domain with a modest negative heat capacity change of binding. Longer DNAs showed systematic trends with respect to enthalpy and entropy of protein binding. As G nucleotides were added to the 5′ end of the $d(T_4G_4)$ DNA, binding was associated with increasingly favorable positive entropy contributions compensated by decreases in the negative enthalpy contribution, consistent with the idea that these longer telomere-derived DNAs form hairpin structures stabilized by G•G base-pairs in their uncomplexed state.[21] Binding with the alpha N-terminal domain resolves these structures, and in the case of $d(T_4G_4T_4G_4)$ DNA, exposes a second protein-binding site. Binding at 3′-terminal and 3′-distal sites may form the basis of higher-order structures and may play an important role in ensuring that DNA remains in an extended conformation suitable for primer utilization by telomerase.

Results

Stoichiometry of DNA binding

To determine the stoichiometry of protein binding with DNA containing single or double $d(T_4G_4)$ telomere repeats we examined the electrophoretic mobility of the alpha N-terminal domain as a function of DNA concentration for $d(T_4G_4)$, $d(G_3T_4G_4)$ and $d(T_4G_4T_4G_4)$ DNAs. Figure 1 shows Coomassie-stained non-denaturing acrylamide gels with bands indicating the mobility of free and DNA-bound protein. Increasing the concentration of $d(T_4G_4)$ and $d(G_3T_4G_4)$ shifted the free protein (band *a*) to a new position corresponding to a monovalent DNA complex (band *b*). The $d(T_4G_4T_4G_4)$ DNA produced two shifted bands. One of these $d(T_4G_4T_4G_4)$-shifted bands corresponds to a negatively charged DNA complex that continually increased in intensity with respect to DNA concentration (band *b*, in Figure 1(c)). The second $d(T_4G_4T_4G_4)$-shifted band corresponds to a positively charged DNA complex that is only observed at DNA:protein ratios less than 1:1 (band *c*, in Figure 1(c)).

Assignment of the two $d(T_4G_4T_4G_4)$-shifted bands follows from consideration of size and charge differences characterizing the uncomplexed protein and DNA–protein complexes. We assign band *c* in Figure 1(c) to a protein•DNA•protein complex since band *c* is positioned midway between that of the free protein (band *a*) and the species encountered with excess DNA (band *b*). Note that band *c* is only observed when DNA is limiting. As more DNA is added, protein in the protein•DNA•protein complex repartitions to form a DNA•protein complex (band *b*). Similar results were observed for the full-length alpha subunit (Figure 1(d)–(f)), indicating that the presence or absence of the alpha C-terminal domain does not change the stoichiometry of binding DNA and that the full-length alpha subunit also forms a protein•DNA•protein complex when DNA binding sites are limiting. These results are consistent with previous characterization of DNA–alpha complexes by gel electrophoresis that showed two alpha proteins can bind one $d(T_4G_4T_4G_4)$ DNA molecule.[22]

Stoichiometry of DNA binding was further evaluated by ITC. Titration of protein into solutions of DNA resulted in well-resolved binding phases and the heat of final injections corresponded with that observed in control titrations of protein into buffer. The reverse titration, injection of DNA into a solution of protein, however, resulted in thermograms for which the heats of final injections were exothermic while the heats of control titrations of DNA into buffer were endothermic. The non-correspondence of final injections can be attributed to a combination of dissociating non-specific DNA–DNA interactions upon dilution and non-specific aggregation upon mixing concentrated DNA solutions with protein. This aggregation effect is visible as a cloudy precipitate that quickly forms and then dissolves as routinely observed in preparation of diffraction quality single crystals. By keeping the DNA in the sample cell at a more modest concentration of 33 μM, these phase transitions and non-specific residual heats were avoided.

The dependence of heat evolved upon binding the DNA-binding domain with $d(T_2G_4)$, $d(T_4G_4)$, $d(G_3T_4G_4)$, and $d(G_4T_4G_4)$ DNAs was characteristic for single-site binding reactions and consistently yielded a 1:1 stoichiometry of DNA binding (Table 1). Representative titrations and isotherms for $d(T_4G_4)$ DNA are shown in Figure 2

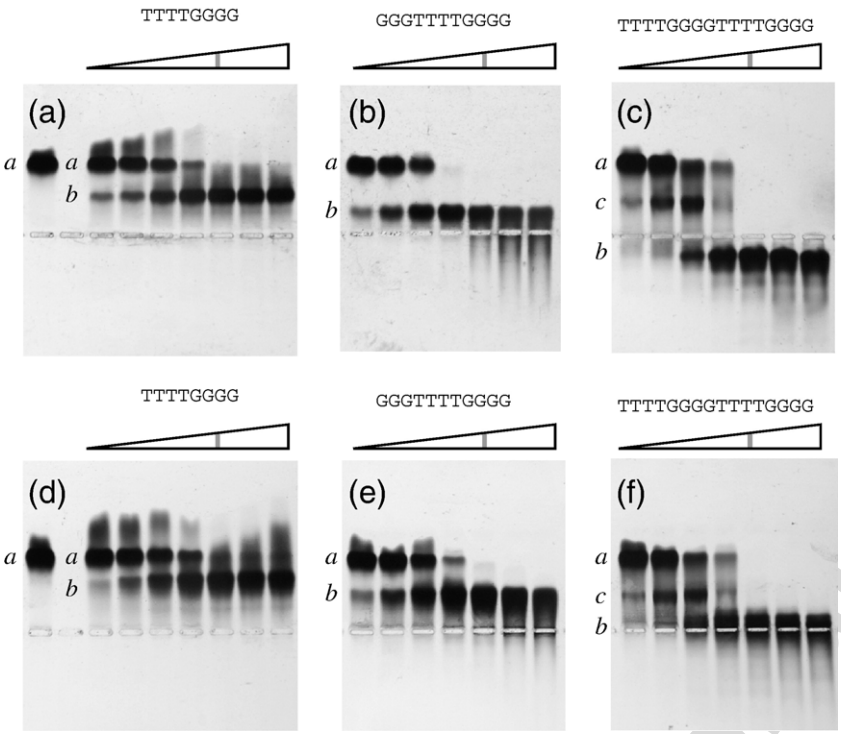

Figure 1. Non-denaturing protein gel electrophoresis mobility assay. The *Oxytricha nova* telomere end binding protein alpha subunit ((d)–(f)) and its N-terminal DNA-binding domain ((a)–(c)) were incubated with increasing amounts of telomere DNA at the following DNA:protein stoichiometric ratios: 0.125:1, 0.25:1, 0.5:1, 0.75:1, 1:1, 1.5:1, and 2:1 before separating on a 7.8% acrylamide gel. The 1:1 binding reaction is indicated with a gray line. Gels were stained with Coomassie blue to reveal protein and protein-associated species. Band *a* corresponds to unbound protein, band *b* corresponds to a DNA•protein complex. The addition of d($T_4G_4T_4G_4$) to alpha or the alpha N-terminal domain resulted in formation of an additional species assigned as a protein•DNA•protein complex (band *c*, in (c) and (f)).

(a)–(c). The d($T_4G_4T_4G_4$) DNA, with two telomere repeats, has potentially two independent and equivalent protein binding sites and was expected to produce a similar isotherm with a stoichiometry parameter $n=2$. Surprisingly, we instead observed isotherms for the binding of protein with d($T_4G_4T_4G_4$) DNA that show two sequential and non-equivalent binding events each with a stoichiometry parameter $n=1$ (Figure 2(d)–(f) and Table 1).

DNA is not folded into G-quartets in these experiments

DNA with telomere repeat sequences will form extremely stable four-stranded G-quartet-stabilized

Table 1. Thermodynamic parameters describing telomere-derived single strand DNA binding as a function of temperature at 0.15 M LiCl

DNA	T (°C)	ΔH (kcal/mol)	ΔS (cal/(mol K))	ΔG (kcal/mol)	$K_{D\text{-}DNA}$ (nM)	n [a]
TTGGGG	10	–6.03±0.03	8.1±0.1	–8.32±0.03	380±20	1.05
TTGGGG	20	–7.25±0.05	4.0±0.2	–8.42±0.04	540±40	1.05
TTGGGG	25	–8.01±0.06	1.0±0.2	–8.32±0.04	790±50	1.05
TTGGGG	30	–8.64±0.06	–0.6±0.2	–8.44±0.04	820±50	1.03
TTTTGGGG	10	–7.81±0.02	5.9±0.1	–9.48±0.04	48±3	1.04
TTTTGGGG	20	–9.63±0.04	–0.1±0.2	–9.60±0.04	69±5	1.00
TTTTGGGG	25	–10.43±0.01	–3.0±0.1	–9.53±0.03	102±6	0.98
TTTTGGGG	30	–11.17±0.03	–5.3±0.1	–9.56±0.02	127±4	1.01
GGGTTTTGGGG	10	–1.96±0.02	25.8±0.6	–9.3±0.2	70±20	1.07
GGGTTTTGGGG	20	–5.06±0.02	16.1±0.1	–9.78±0.03	51±3	1.03
GGGTTTTGGGG	25	–6.65±0.02	10.1±0.1	–9.66±0.04	80±5	1.08
GGGTTTTGGGG	30	–8.12±0.02	6.0±0.1	–9.92±0.03	71±4	1.05
GGGGTTTTGGGG	20	–2.68±0.04	23.1±0.5	–9.4±0.1	90±20	1.03
GGGGTTTTGGGG	25	–4.05±0.03	18.0±0.2	–9.42±0.06	120±20	1.01
GGGGTTTTGGGG	30	–5.61±0.03	15.1±0.3	–10.19±0.08	44±6	0.99
GGGGTTTTGGGG	35	–7.49±0.04	8.2±0.2	–10.02±0.05	79±7	1.03
(A) TTTTGGGGTTTTGGGG	10	+3.71±0.04	48.8±0.4	–10.1±0.1	16±3	0.97
(A) TTTTGGGGTTTTGGGG	25	–1.58±0.04	30.8±0.6	–10.8±0.2	13±4	0.94
(A) TTTTGGGGTTTTGGGG	27	–2.32±0.03	27.9±0.6	–10.6±0.2	16±5	1.01
(A) TTTTGGGGTTTTGGGG	30	–3.59±0.03	25±2	–11.1±0.5	10±8	0.88
(B) TTTTGGGGTTTTGGGG	10	–4.3±0.2	10.1±0.8	–7.13±0.09	3200±500	1.01
(B) TTTTGGGGTTTTGGGG	25	–5.3±0.2	6.4±0.7	–7.18±0.06	5600±600	1.03
(B) TTTTGGGGTTTTGGGG	27	–5.4±0.2	6.0±0.7	–7.19±0.05	5800±500	1.06
(B) TTTTGGGGTTTTGGGG	30	–5.6±0.2	5.3±0.7	–7.23±0.06	6100±600	1.00

[a] Uncertainty in the parameter n was, on average, 0.005 and ranged from 0.001–0.02, for all data excluding site B of d($T_4G_4T_4G_4$). For measures related to site B of d($T_4G_4T_4G_4$), uncertainty in n was 0.03 for all measures.

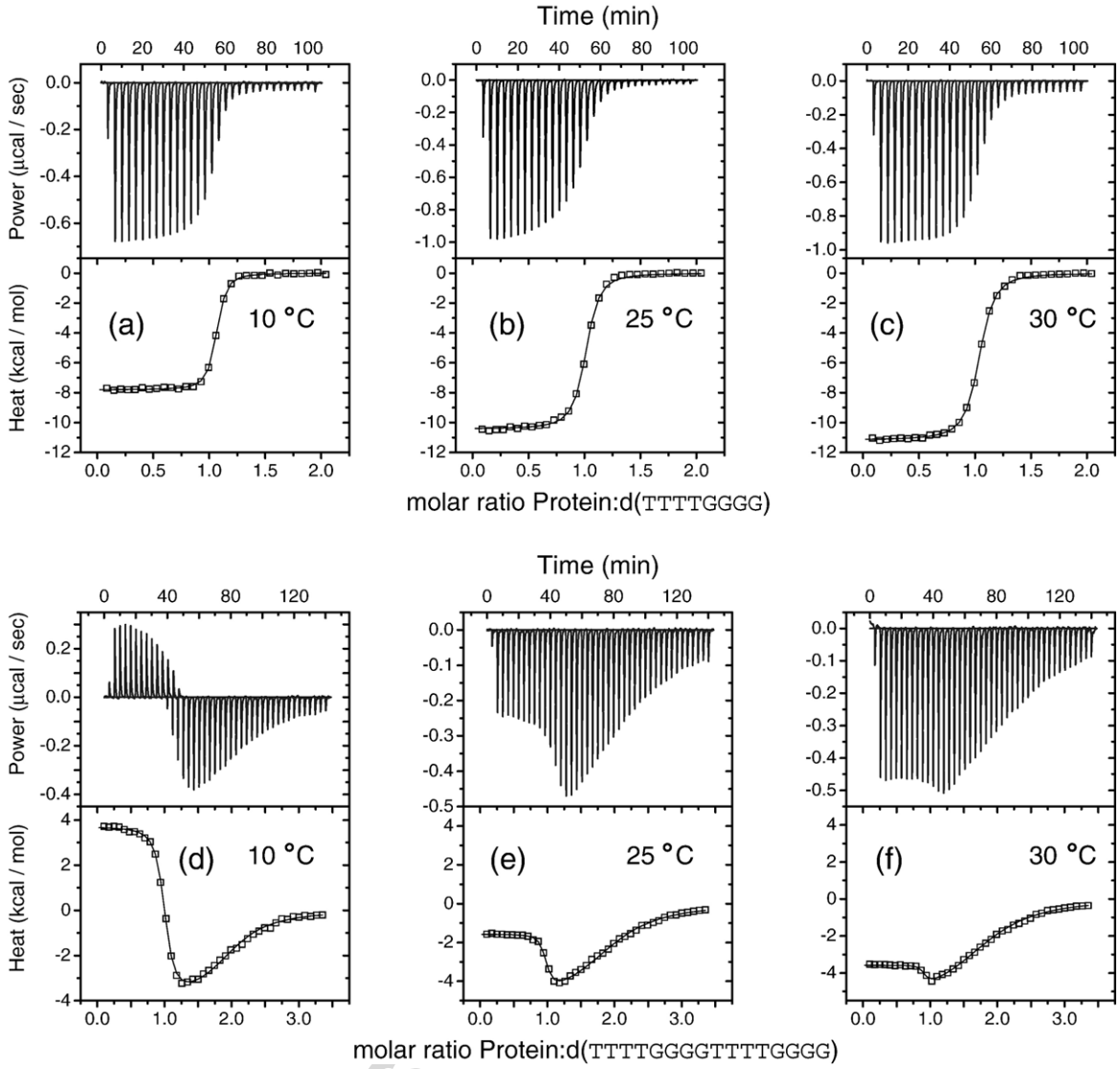

Figure 2. Representative ITC data. Thermograms (upper half of each panel) and binding isotherms (bottom half of each panel) obtained for the binding of alpha N-terminal domain with $d(T_4G_4)$ ((a)–(c)) and $d(T_4G_4T_4G_4)$ ((d)–(f)) at 10, 25, and 30 °C. A total of 31 and 42 injections were performed respectively for $d(T_4G_4)$ and $d(T_4G_4T_4G_4)$. Non-linear least-squares fit of the $d(T_4G_4)$ data with a ''one set of site'' model indicated one binding site with a stoichiometric parameter $n=1$. Non-linear least-squares fit of the $d(T_4G_4T_4G_4)$ data with a ''two set of sites'' model indicated two binding sites with a stoichiometric parameter $n=1$ for each site.

structures in the presence of appropriate cations.[23–31] These G-quartet structures are likely involved with telomere biology,[17,32–34] but in the absence of other factors these folded DNAs interfere with alpha protein binding[35–37] and hamper the otherwise processive extension of telomeres by telomerase.[38,39] Experiments were carried out in lithium chloride-containing solutions, and we are convinced that under these conditions G-quadruplexed DNA is at most a small component of the ensemble of DNA structures in these experiments since (1) adding potassium chloride eliminates all measurable enthalpy of DNA–protein interaction (data not shown), and (2) in the absence of added potassium ions, the measured stoichiometry of DNA–protein association is very close to one (Tables 1 and 3). If a substantial portion of the DNA formed stable G-quadruplexes, then this stoichiometry parameter would be substantially less than one, since a significant fraction of DNA would be unavailable for association with protein. Because $n \sim 1$ in all of our calorimetric titrations, we conclude that kinetically inert G-quadruplexed DNA was not present at significant levels in our experiments.

Affinity and free energy of binding DNA

The DNA containing a single telomere repeat, $d(T_4G_4)$, binds the alpha N-terminal DNA-binding domain at 25 °C with $K_{D\text{-}DNA}=102(\pm6)$ nM ($\Delta G=-9.5$ kcal/mol, Table 1). Shortening the DNA by removing the 5′-terminal TT nucleotides reduces the binding affinity such that $d(T_2G_4)$ binds protein with $K_{D\text{-}DNA}=790(\pm50)$ nM ($\Delta G=-8.3$ kcal/mol) at

25 °C. The affinities measured for d($G_3T_4G_4$) and d($G_4T_4G_4$) are $K_{D\text{-}DNA}$ = 80(±5) (ΔG = –9.7 kcal/mol) and $K_{D\text{-}DNA}$ = 120(±20) (ΔG = –9.4 kcal/mol), respectively, indicating that adding G nucleotides to the 5′-end of the d(T_4G_4) DNA has almost no effect on the overall free energy of binding. Comparing these measures shows that an eight-nucleotide d(T_4G_4) may represent the minimal binding site, and additional nucleotides have only a marginal effect on association.

The d($T_4G_4T_4G_4$) DNA possesses two binding sites, one strong site, which we will call the ''A site'' characterized by $K_{D\text{-}DNA(A)}$ = 13(±4) nM (ΔG = –10.8 kcal/mol) and a weaker ''B site'' with $K_{D\text{-}DNA(B)}$ = 5600(±600) nM (ΔG = –7.2 kca/mol) at 25 °C. Comparing all of the DNAs yields the following trend for DNA–protein complex stability: A site of d($T_4G_4T_4G_4$) > d($G_4T_4G_4$) ≈ d($G_3T_4G_4$) ≈ d(T_4G_4) > d(T_2G_4) > B site of d($T_4G_4T_4G_4$), a series that spans 3.6 kcal/mol in binding free energy.

A molecular interpretation of the two binding events associated with d($T_4G_4T_4G_4$) is that site A corresponds with the 3′-terminal TTTTGGGG and site B corresponds with the 3′-distal TTTTGGGG repeat. Although these sites are identical in sequence each is chemically distinct, since one ends with a 3′ OH group while the other ends with a bridging phosphodiester linkage. An alternative explanation for weaker site B binding invokes site-to-site interactions. In this model protein occupancy at the first binding site influences binding at the second site. We recognize that heterogeneous sites and equivalent but interacting sites cannot be easily differentiated;[40] however, previous binding studies have shown that the alpha N-terminal domain prefers sites with a 3′ OH group,[19] a result we confirmed through experiments with DNA sequence variants (this work), meaning that a portion of the difference in binding affinity measured for sites A and B is due to intrinsic differences between these sites.

Enthalpy, entropy and heat capacity for DNA binding

The ITC method provides measures of enthalpy and entropy of binding reactions in addition to binding free energy and dissociation constants.[41–43] Enthalpy and entropy reflect various interactions such as formation of van der Waals contacts, hydrogen bonds and release of water and hydrated ions as hydrophobic and polar groups become buried at the DNA–protein interface. Binding reactions with d(T_2G_4) and d(T_4G_4) DNAs were strongly exothermic, indicating a favorable negative enthalpy (ΔH) of binding (Figure 3(a) and (b)). In contrast, binding reactions with d($G_3T_4G_4$), d($G_4T_4G_4$), and the A site of d($T_4G_4T_4G_4$) showed more favorable positive entropy (ΔS) of binding (Figure 3(a) and (c)).

The balance between enthalpy and entropy is temperature-dependent, since both ΔH and ΔS decrease monotonically as function of temperature (Figure 3(b) and (c)). In our entropy and enthalpy measures we observed strikingly systematic trends with negative enthalpy decreasing in magnitude and positive entropy increasing in magnitude as a function of DNA length (Figure 3(a)). Comparing d(T_4G_4) and d($T_4G_4T_4G_4$), there was complete reversal of the relative importance of enthalpy and entropy. At 20 °C, the d(T_4G_4) binding reaction is enthalpy-driven ($\Delta H < 0$, $\Delta S \sim 0$), and the d($T_4G_4T_4G_4$) binding reaction at the A site is entropy-driven ($\Delta H \sim 0$, $\Delta S > 0$). As described further in Discussion, this trend with respect to DNA length likely reflects structures in the uncomplexed DNA, which become increasingly more stable as additional G and T nucleotides are added to the 5′ end of the DNA, and which are resolved upon binding DNA with protein.

We measured ΔCp from the slope of ΔH *versus* temperature plots (Figure 3(b)). The DNA binding reactions were associated with negative heat capacity changes. Overall, the magnitude of heat capacity change ($|\Delta Cp|$) increased with DNA length; however, the enthalpy changes accompanying binding of d(T_2G_4) and d(T_4G_4), are significantly less temperature-dependent than seen for d($G_3T_4G_4$), d($G_4T_4G_4$), and the A site of d($T_4G_4T_4G_4$) (Figure 3(d) and Table 2). We believe that the pattern of ΔCp with respect to DNA length is consistent with the idea that structural changes in the DNA are coupled with binding.

The magnitude of ΔCp is dramatically different for binding at sites A and B of d($T_4G_4T_4G_4$). While binding at site A is characterized by the largest magnitude in heat capacity change with ΔCp = –360(±8) cal mol^{-1} K^{-1}, binding at site B has the smallest magnitude of ΔCp for all DNAs examined in this study with ΔCp = –67(±1) cal mol^{-1} K^{-1}. These heat capacity changes suggest that binding events associated with A and B sites are different in some manner.

Polyelectrolyte effect

In order to determine the number of ions paired upon binding of DNA with protein, the association constant was measured at different concentrations of lithium chloride (Figure 4 and Table 3). Lithium (Li^+) was chosen to avoid complications arising from kinetically inert G-quartet stabilized structures that form in the presence of other cations such as Na^+, K^+, and NH_4^+.[23,24,28] Although Li^+ and these other ions interact very differently with the O^6 coordination sphere of G-quartets, their hydrated forms likely interact very similarly with the polyanion phosphodiester groups of DNA.[44,45] For instance, studies with the bacterial single strand DNA binding protein in a variety of salts showed that DNA–protein interactions were the same whether measured in solutions containing NaCl or LiCl.[45]

Binding reactions for each of the single strand DNA fragments showed relatively weak dependence on salt concentration. The number of ion pairs

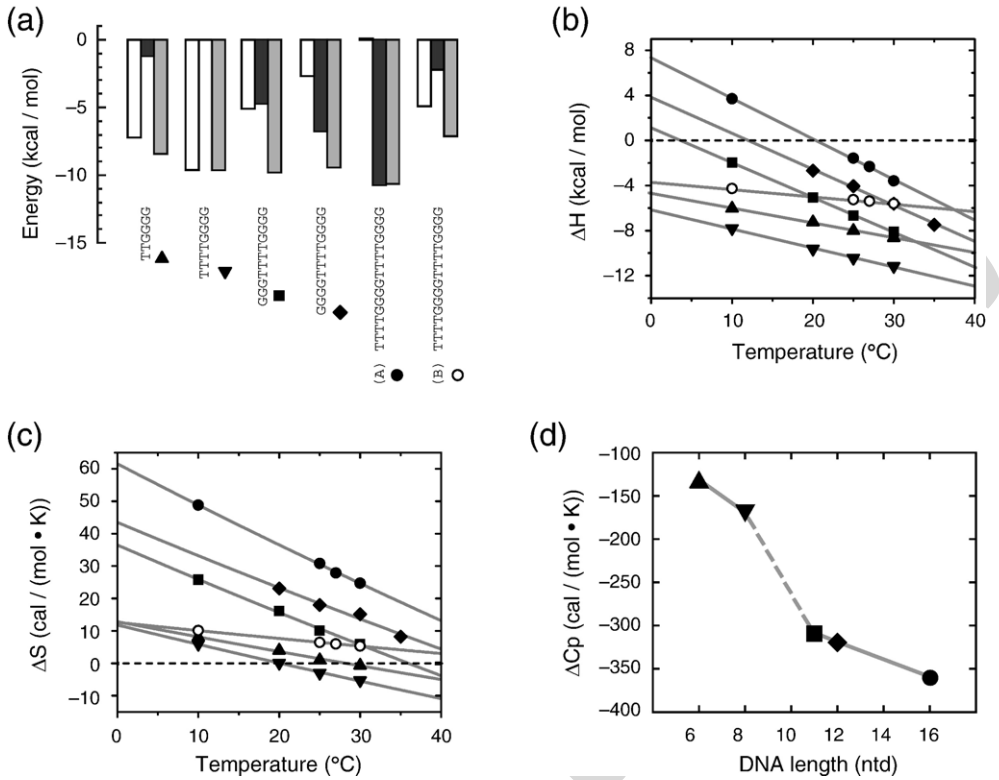

Figure 3. Calorimetric characterization of DNA binding. (a) Thermodynamic parameters for binding $d(T_2G_4)$, $d(T_4G_4)$, $d(G_3T_4G_4)$, $d(G_4T_4G_4)$ and sites A and B of $d(T_4G_4T_4G_4)$ at 20 °C are shown with free energy change (ΔG) indicated with a gray bar, enthalpy change (ΔH) with a white bar, and entropy change ($-T\Delta S$) with a black bar. The enthalpic and entropic contributions to free energy show systematic trends with $-T\Delta S$ increasingly favorable for $d(G_3T_4G_4)$, $d(G_4T_4G_4)$ and site A of $d(T_4G_4T_4G_4)$. (b) Binding enthalpy change for $d(T_2G_4)$ (▲), $d(T_4G_4)$ (▼), $d(G_3T_4G_4)$ (■), $d(G_4T_4G_4)$ (♦), site A of $d(T_4G_4T_4G_4)$ (●), and site B of $d(T_4G_4T_4G_4)$ (○) is plotted as a function of temperature. Model lines represent linear least-squares fitting of the data to ΔCp (T–$T_{\Delta H=0}$), where the slope defines ΔCp, the change in heat capacity, and the x-intercept defines $T_{\Delta H=0}$, the temperature at which enthalpy change is zero. (c) Entropy change is plotted as a function of temperature. Model curves represent non-linear least-squares fitting of the data to $\Delta Cp\ \ln(T/T_{\Delta S=0})$ where $T_{\Delta S=0}$ is the temperature at which entropy change is zero. Values of ΔCp, $T_{\Delta H=0}$ and $T_{\Delta S=0}$ are reported in Table 2. (d) Heat capacity change is plotted as a function of DNA length for $d(T_2G_4)$, $d(T_4G_4)$, $d(G_3T_4G_4)$, $d(G_4T_4G_4)$ and site A of $d(T_4G_4T_4G_4)$. The line connects data points. The broken line highlights a sharp increase in the magnitude of ΔCp experienced upon addition of a second G tract. As discussed in the text, the second G tract allows for the formation of intramolecular hairpin structure stabilized by G•G base-pairs that, if formed, must dissociate upon binding with protein.

(Z) was calculated from the slope of $\ln(K_{A\text{-}DNA})$ *versus* ln([LiCl]) plots (Figure 4) normalized by –0.71, the accepted value for single strand DNA[46] (Table 4). For $d(T_2G_4)$ $Z=1.3$, whereas for $d(T_4G_4)$ $Z=2.1$. $d(G_3T_4G_4)$, $d(G_4T_4G_4)$, and site A of $d(T_4G_4T_4G_4)$ each showed a salt dependence of protein association nearly identical with that of $d(T_4G_4)$ with $Z=2.0$–2.5. The B site of $d(T_4G_4T_4G_4)$ was characterized by $Z=0.9$, further indication that binding events associated with site B are different from those associated with site A. It should be noted that the salt-dependence of DNA association was determined at 20 °C for the $d(T_2G_4)$, $d(T_4G_4)$, $d(G_3T_4G_4)$, and $d(G_4T_4G_4)$ DNAs while data were obtained at 25 °C for the $d(T_4G_4T_4G_4)$ DNA. Comparison with $d(T_4G_4T_4G_4)$ is nevertheless reasonable, since the slope of $\ln(K_{A\text{-}DNA})$ *versus* ln([LiCl]) plots is generally only weakly temperature-dependent,[47] except in the case of strongly polar interactions.[48]

Table 2. Temperature dependence of enthalpy and entropy changes in 0.15 M LiCl

DNA	$\Delta C_P^{observed}$ (cal/(mol K))	$T_{\Delta H=0}$ (K)	$T_{\Delta S=0}$ (K)
TTGGGG	–131±4	237±2	301±1
TTTTGGGG	–169±6	236±3	293±1
GGGTTTTGGGG	–309±3	277±1	309±1
GGGGTTTTGGGG	–320±20	285±1	318±2
(A) TTTTGGGGTTTTGGGG	–360±8	293±1	325±1
(B) TTTTGGGGTTTTGGGG	–67±1	219±1	327±1

Extrapolation to 1 M LiCl provides a free energy reference point ($\Delta G_{1M\ LiCl}$) and allows calculation of the relative contributions of electrostatic and non-electrostatic interactions under physiological conditions of 0.15 M LiCl (Table 4). Approximately 85% of the total free energy of binding corresponds with non-electrostatic interactions for all DNAs tested (Table 4). A similarly high non-electrostatic contribution for binding telomere-derived DNA has also been observed for a telomere single strand DNA binding system from fission yeast.[49] Single strand

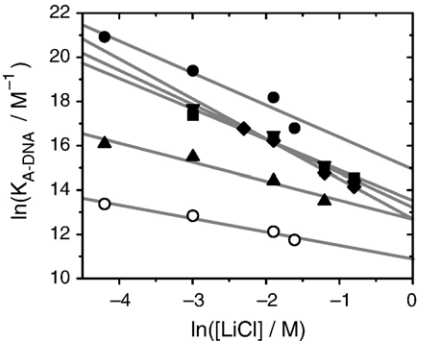

Figure 4. Salt dependence of DNA binding. Natural log-log plots of association binding constant ($K_{A\text{-}DNA}$) *versus* lithium chloride concentration were fit by linear regression. Symbols for each DNA are the same as in Figure 3. The slopes ($\partial \ln K_a / \partial \ln[LiCl]$) define $-Z\psi$, where Z is the apparent number of ion pairs formed between DNA and protein and ψ is the theoretical number of cations released into solution upon formation of one ionic pair. These parameters describing the poly-electrolyte effect are reported in Tables 3 and 4. Note that data for $d(T_2G_4)$, $d(T_4G_4)$, $d(G_3T_4G_4)$, $d(G_4T_4G_4)$ DNAs were obtained at 20 °C, while the data for sites A and B of $d(T_4G_4T_4G_4)$ were obtained at 25 °C.

telomere DNA binding proteins apparently adapted to recognize DNA in a sequence-specific manner by relying less on electrostatic interaction in favor of other non-electrostatic interactions such as stacking of aromatic groups and hydrogen-bonding.[49–51]

Binding of DNA variants

The penultimate G nucleotide is critical for binding to the alpha DNA-binding domain as determined for d(TTTTGGGG)-derived DNAs.[20] We replaced this nucleotide with cytosine (C) at G7 and G15 in the d(GGGTTTTGGGG) DNA† and at G15 in d(TTTTGGGGTTTTGGGG) in order to address the question of whether A and B sites correspond to particular locations on the telomere DNA fragment. Titration of the alpha N-terminal domain into d(GGGTTTTGG**C**G) showed no detectable binding at 20 °C and 30 °C, whereas titration into d(G**C**GTTTTGGGG) at 25 °C showed a similar thermogram and binding isotherm as seen for the cognate $d(G_3T_4G_4)$ DNA (Figure 5(a)). The measured enthalpy, entropy, free energy of binding, and dissociation constants measured for these DNA variants are reported in Table 5. The position-dependence of the G-to-C substitution confirmed that the 3′-terminal G tract is critical for binding while the 3′-distal G tract is less important. The small differences in the magnitude of negative enthalpy and free energy of binding measured for the d(G**C**GTTTTGGGG) likely reflects differences in the enthalpy and stability of G•C Watson-Crick *versus* G•G Hoogsteen base-pairs present in the ensemble of unbound DNA structures.

Nucleotide substitution variants of the 11-mer $d(G_3T_4G_4)$ DNA indicated that the alpha N-terminal domain binds preferentially to sites with a 3′ OH group rather than to internal sites, consistent with interactions seen in the DNA–alpha N-terminal domain co-crystal structure[19] and conclusions from previous binding studies.[20] Since replacement of the penultimate G in $d(G_3T_4G_4)$ completely abolished protein association, we reasoned that a similar replacement in $d(T_4G_4T_4G_4)$ should provide a route to measure the binding properties of the 3′-distal site of d(TTTTGGGGTTTTGG**C**G) before the 3′-terminal site becomes occupied. For comparison, $d(T_4G_4T_4)$ DNA was evaluated, since this DNA also places the TTTTGGGG site at a 3′-distal location. Figure 5(b) compares binding isotherms obtained for $d(T_4G_4T_4G_4)$, d(TTTTGGGGTTTTGG**C**G), and $d(T_4G_4T_4)$. The similarity of the first part of the binding isotherm for d(TTTTGGGGTTTTGG**C**G) and second part of the binding isotherm for $d(T_4G_4T_4G_4)$ and also the complete lack of binding with d(GGGTTTGG**C**G) suggest that there is a reversal in binding order with site B occupied first after which the mutated site A becomes occupied. Additional proof of this suggestion comes from the binding profile obtained with $d(T_4G_4T_4)$ DNA, which shows one binding step with a $K_{D\text{-}DNA}$ = 1000(±50) nM, very similar to the $K_{D\text{-}DNA}$ = 1400(±700) nM measured for the first B site binding transition of d(TTTTGGGGTTTTGG**C**G).

On the basis of these comparisons we assign the first binding event for d(TTTTGGGGTTTTGG**C**G) to a 3′-distal site. The $K_{D\text{-}DNA(B)}$ measured for site B in the context of cognate $d(T_4G_4T_4G_4)$ is slightly higher than measured for the 3′-distal site in d(TTTTGGGG-TTTTGG**C**G) and $d(T_4G_4T_4)$; however, these sites are likely related. Differences in free energy of binding at site B for these different DNAs ($\Delta\Delta G \sim 1$ kcal/mol) are not unexpected given that the A site is complexed with protein in $d(T_4G_4T_4G_4)$ but unoccupied in d(TTTTGGGGTTTTGG**C**G) during the binding step at site B. The picture emerging indicates that interaction between the two TTTTGGGG sites makes each responsive to the binding state of the other site. Such site-to-site interactions could be mediated either through DNA structural changes or through direct protein–protein contact.

Further evidence for interaction between binding sites comes from comparison of the binding properties of the mutated sequences in d(GGGTTTTGG**C**G) and d(TTTTGGGGTTTTGG**C**G). As noted above, no binding was detected with the shorter DNA at either of two temperatures. Binding of d(TTTTGGGGTT-TTGG**C**G) at 25 °C is characterized by two transitions, the weaker of which we assign to an impaired A site with $K_{D\text{-}DNA(A\text{-}mutated)}$ = 20,000(±7000) nM. This impaired A site was not accessible in the context of the 11-mer d(GGGTTTTGG**C**G) DNA but is available in d(TTTTGGGGTTTTGG**C**G) so long as

† We will refer to nucleotide positions by their number in the d(TTTTGGGGTTTTGGGG) sequence.

Table 3. Thermodynamic parameters describing telomere-derived single strand DNA binding as a function of salt concentration

DNA	T (°C)	[LiCl] (M)	ΔH (kcal/mol)	ΔS (cal/(mol K))	ΔG (kcal/mol)	$K_{D\text{-}DNA}$ (nM)	n [a]
TTGGGG	20	0.015	–7.94±0.04	4.9±0.2	–9.38±0.03	102±5	0.85
TTGGGG	20	0.050	–7.92±0.06	3.8±0.4	–9.03±0.09	180±30	0.98
TTGGGG	20	0.150	–7.25±0.05	4.0±0.2	–8.42±0.04	540±40	1.05
TTGGGG	20	0.300	–7.07±0.08	2.8±0.4	–7.89±0.08	1400±200	0.94
TTTTGGGG	20	0.050	–8.98±0.03	4.5±0.2	–10.30±0.05	21±2	1.02
TTTTGGGG	20	0.150	–9.63±0.04	–0.1±0.2	–9.60±0.04	69±5	1.00
TTTTGGGG	20	0.300	–9.36±0.06	–1.9±0.3	–8.80±0.04	270±20	1.02
TTTTGGGG	20	0.450	–10.06±0.01	–6.0±0.9	–8.3±0.3	600±300	0.99
GGGTTTTGGGG	20	0.050	–4.93±0.06	17.8±0.6	–10.1±0.2	28±8	1.00
GGGTTTTGGGG	20	0.150	–5.06±0.02	16.1±0.1	–9.78±0.03	51±3	1.03
GGGTTTTGGGG	20	0.300	–5.11±0.03	12.2±0.1	–8.68±0.02	330±10	0.94
GGGTTTTGGGG	20	0.450	–4.71±0.03	12.8±0.1	–8.46±0.02	480±20	1.07
GGGGTTTTGGGG	20	0.100	–2.32±0.03	25.5±0.4	–9.8±0.1	50±10	0.93
GGGGTTTTGGGG	20	0.150	–2.68±0.04	23.1±0.5	–9.4±0.1	90±20	1.03
GGGGTTTTGGGG	20	0.300	–2.23±0.04	21.8±0.4	–8.6±0.1	380±70	0.98
GGGGTTTTGGGG	20	0.450	–1.73±0.05	22.2±0.6	–8.2±0.2	700±200	1.11
(A) TTTTGGGGTTTTGGGG	25	0.015	–2.21±0.05	34.2±0.8	–12.4±0.2	1.0±0.4	0.98
(A) TTTTGGGGTTTTGGGG	25	0.050	–1.68±0.05	32.9±0.5	–11.5±0.1	4±1	1.00
(A) TTTTGGGGTTTTGGGG	25	0.150	–1.58±0.04	30.8±0.6	–10.8±0.2	13±4	0.94
(A) TTTTGGGGTTTTGGGG	25	0.200	–1.29±0.06	29.0±0.8	–9.9±0.2	50±20	0.93
(B) TTTTGGGGTTTTGGGG	25	0.015	–7.9±0.1	0.2±0.4	–7.92±0.03	1600±100	1.03
(B) TTTTGGGGTTTTGGGG	25	0.050	–6.3±0.2	4.5±0.8	–7.61±0.06	2600±300	1.03
(B) TTTTGGGGTTTTGGGG	25	0.150	–5.3±0.2	6.4±0.8	–7.18±0.06	5600±600	1.03
(B) TTTTGGGGTTTTGGGG	25	0.200	–5.4±0.6	5±2	–6.96±0.08	7700±1200	0.91

[a] Uncertainty in the parameter n was, on average, 0.007 and ranged from 0.001–0.02, for all data excluding site B of $d(T_4G_4T_4G_4)$. For measures related to site B of $d(T_4G_4T_4G_4)$, uncertainty in n was, on average, 0.04 and ranged from 0.01–0.09.

the B site is bound to protein. These results are consistent with the idea that site-to-site interactions influence binding behavior at the two adjacent TTTTGGGG repeats of the $d(T_4G_4T_4G_4)$ telomere fragment.

Comparison of $d(T_4G_4T_4)$ and any of the other DNAs with one TTTTGGGG sequence gives an estimate for the free energy difference distinguishing binding at 3′-distal and 3′-terminal sites, $\Delta\Delta G$ = +1.5 kcal/mol, corresponding to an approximately tenfold higher $K_{D\text{-}DNA}$ at 25 °C for the 3′-distal site. Extending this estimate to the $d(T_4G_4T_4G_4)$ telomere-derived DNA, the first binding event at site A involves binding at either the 3′-terminal telomere repeat or the 3′-distal repeat to form a DNA●protein complex with approximately 90% of the DNA occupied at the 3′-terminal site and 10% occupied at the 3′-distal site. That the second binding event at site B is associated with a less favorable free energy of binding could be a consequence of at least three effects: (1) most of the second binding event involves association with the intrinsically weaker 3′-distal site; (2) the statistical likelihood of association at the second site is less, since there is only one way to add the second protein but two ways to dissociate the protein●DNA●protein complex; and (3) the available binding site is adjacent to an occupied site and therefore involves placing two proteins close to each other.

Effects (1) and (2) work against binding the second protein, while effect (3) could either help or discourage binding depending on whether close approach of the two proteins leads to overall favorable cooperative interactions or overall unfavorable anti-cooperative interactions. The +3.6 kcal/mol free energy of DNA binding distinguishing site B from site A predicts either that a certain degree of steric strain or other repulsive terms are introduced through juxtaposition of the two alpha DNA-binding domains, or that the second protein makes a new set of interactions with DNA, which are collectively weaker than those made by the first protein, in order to avoid steric collisions. We are currently

Table 4. Polyelectrolyte effect for DNA binding

DNA	T (°C)	$-\partial \ln K_A / \partial \ln[LiCl]$	Z	$\Delta G_{1M\ LiCl}$ (kcal/mol)	$\Delta G_{electrostatic}$ (kcal/mol)	$\Delta G_{non\text{-}electrostatic}$ (kcal/mol)	$F_{non\text{-}electrostatic}$ (%)
TTGGGG	20	0.9±0.1	1.3±0.2	–7.4±0.2	–0.8±0.2	–7.6±0.2	90±2
TTTTGGGG	20	1.5±0.2	2.1±0.3	–7.7±0.2	–1.5±0.2	–8.1±0.2	84±2
GGGTTTTGGGG	20	1.4±0.3	2.0±0.3	–7.8±0.2	–1.6±0.2	–8.2±0.2	84±2
GGGGTTTTGGGG	20	1.8±0.1	2.5±0.1	–7.4±0.1	–1.5±0.2	–7.9±0.1	84±2
(A) TTTTGGGGTTTTGGGG	25	1.4±0.2	2.0±0.3	–8.8±0.4	–1.6±0.5	–9.2±0.4	85±4
(B) TTTTGGGGTTTTGGGG	25	0.6±0.1	0.9±0.1	–6.5±0.1	–0.5±0.1	–6.7±0.1	93±2

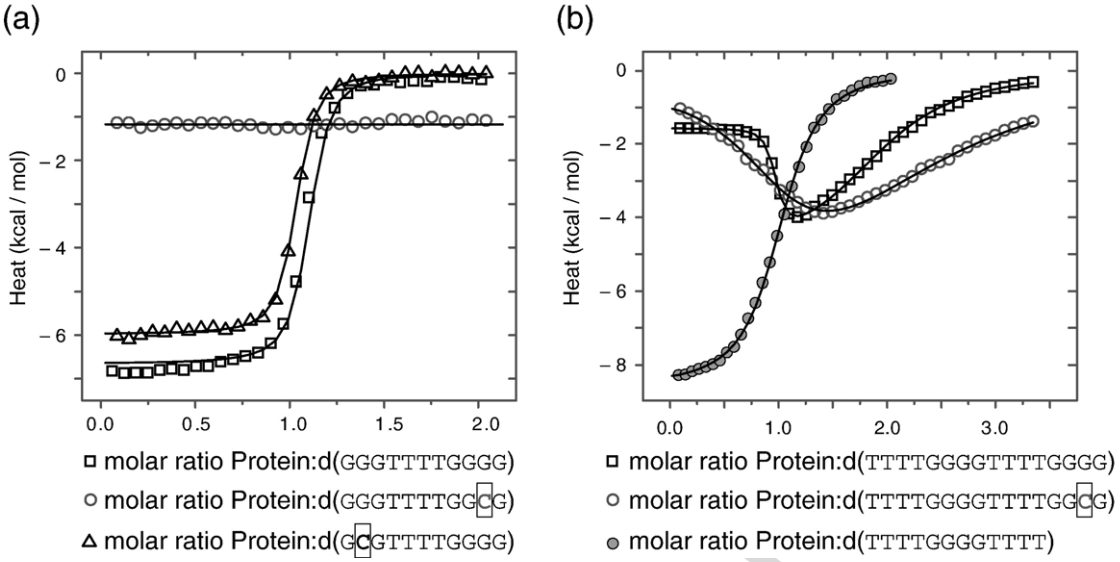

Figure 5. Binding isotherm profiles for nucleotide substitution variants. (a) Titration of GGGTTTTGGGG (□) and G**C**GTTTTGGGG (△) at 25 °C showed high similarity in binding, whereas titration of GGGTTTTG**C**G (○) at 20 or 30 °C showed no detectable binding. (b) A comparison of isotherms performed at 25 °C is shown for the DNA-binding domain binding with TTTTGGGGTTTTGGGG (□), TTTTGGGGTTTTGG**C**G (○) and TTTTGGGGTTTT (●). The single binding site observed for TTTTGGGGTTTT has a similar $K_{D\text{-}DNA}$ compared with site B′ of TTTTGGGGTTTTGG**C**G and site B of TTTTGGGGTTTTGGGG (see Table 5).

undertaking structural studies in order to distinguish between these two possibilities.

Discussion

Thermodynamic character of binding d(T_4G_4)

Experiments presented here showed that binding of the alpha N-terminal domain with a single strand telomere d(T_4G_4) repeat at 20 °C is associated with a favorable change in enthalpy (ΔH=–9.6 kcal/mol) and a near zero change in entropy (ΔS=–0.1 cal mol^{-1} K^{-1}). The temperature dependence of ΔH indicated a modest and negative change in heat capacity (ΔCp=–169 cal mol^{-1} K^{-1}), while the salt dependence of DNA association showed that most of the free energy is derived from non-electrostatic interactions at physiological salt concentrations. The thermodynamic character of d(T_4G_4)–alpha N-terminal domain association can now be compared with interactions seen in co-crystal structures containing the alpha N-terminal domain and DNA.[19,52] These structures revealed four classes of interactions: (1) arginine-phosphate ion pairs; (2) π–π tyrosine-base stacking interactions; (3) hydrogen bonds between base edges and polar carboxylate and amide bearing amino acid residues; and (4) hydrophobic type interactions between moieties of the deoxyribose and apolar side-chains. Here we discuss how these interactions may contribute to the thermodynamics of DNA–protein binding measured in our experiments.

The co-crystal structures show two phosphate groups of d(T_4G_4) making salt-bridges with two arginine residues[19,52] that likely contribute to salt-sensitive DNA–protein contacts totaling ~15% of

Table 5. Thermodynamic parameters for nucleotide substitution variants in 0.15 M LiCl

DNA	T (°C)	ΔH (kcal/mol)	ΔS [cal/(mol K)]	ΔG (kcal/mol)	K_{D-DNA} (nM)	n [a]
GGGTTTTGGGG	20	–5.06±0.02	16.1±0.1	–9.78±0.03	51±3	1.03
GGGTTTTGGGG	25	–6.65±0.02	10.2±0.1	–9.66±0.04	80±5	1.08
G**C**GTTTTGGGG	20	–4.01±0.04	17.8±0.3	–9.23±0.09	130±20	1.02
G**C**GTTTTGGGG	25	–5.97±0.03	12.6±0.2	–9.73±0.05	72±6	1.01
GGGTTTTGG**C**G	20	none detected	———	———	———	—
GGGTTTTGG**C**G	30	none detected	———	———	———	—
(A) TTTTGGGGTTTTGGGG	25	–1.58±0.04	30.8±0.6	–10.8±0.2	13±4	0.94
(B) TTTTGGGGTTTTGGGG	25	–5.3±0.2	6.4±0.7	–7.18±0.06	5600±600	1.03
(B′) TTTTGGGGTTTTGG**C**G	25	–0.3±0.2	26±1	–8.0±0.3	1400±700	1.01
(A′) TTTTGGGGTTTTGG**C**G	25	–10±4	–12±10	–6.4±0.2	20,000±7000	1.2
TTTTGGGGTTTT	25	–8.61±0.04	–1.5±0.2	–8.15±0.3	1000±50	1.03

[a] Uncertainty in the parameter n was, on average, 0.008 and ranged from 0.002–0.03, for all data excluding ″site B″ and ″site A″ of d(TTTTGGGGTTTTGGCG). For ″site B″ and ″site A″ of d(TTTTGGGGTTTTGGCG) the uncertainty in n was 0.09 and 0.4 respectively.

the overall free energy of DNA binding at 0.15 M lithium chloride. Stacking interactions between bases of the DNA and four tyrosine residues of the protein complement salt-bridges seen in the co-crystal structures. An additional base makes a similar stacking-type interaction with the guanidinium group of an arginine residue for a total of five such π−π contacts at the DNA−protein interface. Model compound studies[53] and protein−DNA examples[54] showed that stacking of aromatic residues is associated with favorable changes in enthalpy. Given the number of these types of contacts, the stacking of aromatic and heteroaromatic groups may be a dominant force leading to a stable single strand DNA−protein complex.

Telomere single strand DNA binding proteins have evolved to recognize a particular sequence of DNA, and some of the interactions seen in alpha N-terminal domain containing co-crystal structures involve apparently sequence-specific hydrogen bonds between the edge groups of DNA bases and polar asparate, asparagine, arginine, and glutamine side-chains of the protein.[14,19,52] H-bond formation and removal of polar groups from water are enthalpy-driven processes,[55–57] so these DNA sequence-specific interactions should augment the favorable enthalpy term of DNA−protein complex formation.

Hydrophobic-type contacts are seen at the DNA−protein interface where aliphatic moieties of the DNA, especially the deoxyribose group, nestle with valine, isoleucine, and phenylalanine residues of the protein. These types of contacts with the 2′-methylene group likely contribute to DNA *versus* RNA discrimination.[14,19,52] Removing hydrophobic groups from solvent is entropy-driven.[58–60] However, there are fewer of these hydrophobic interactions in comparison with the enthalpy-driven aromatic stacking and hydrogen-bonding interactions, consistent with the thermodynamic signature determined here showing that enthalpy favors DNA−protein complex formation while the net entropy change coupled with $d(T_4G_4)$ binding is close to zero.

The interactions seen in the *O. nova* telomere end binding protein−single strand DNA complexes are different in detail, but similar in general terms to interactions seen in other telomere single strand DNA−protein complexes from human and fungi. The human and fission yeast protection of telomere protein (POT1) DNA binding domain each share one folding unit with the *O. nova* alpha N-terminal domain,[61,62] and each protein makes hydrogen bonding contacts with sequence-specific groups of the DNA, stacking interactions involving nucleotide bases and aromatic residues, and a few salt-bridges with DNA phosphate groups.[61,62] The single strand telomere DNA binding domain from the CDC13 protein of budding yeast, is constructed from a structurally similar fold.[63] A scanning alanine mutagenesis study for the CDC13 single strand telomere DNA binding domain confirmed and highlighted the importance of aromatic side-chains for DNA affinity.[50] The repeated use of similar types of interactions in these telomere single strand DNA binding proteins likely reflects a common evolutionary pressure to recognize and bind with the telomere repeat in a sequence-specific manner.[15,50]

Structural changes coupled with binding DNA

The 8 nt $d(T_4G_4)$ DNA represents only a portion of the 16 nt single strand $d(T_4G_4T_4G_4)$ DNA found at the ends of mature telomeres in *O. nova*. During *de novo* telomere addition and telomere replication the single strand DNA may be even longer.[64,65] In order to understand how longer DNAs interact with the alpha DNA-binding domain, we compared the thermodynamic parameters characterizing binding the short DNAs, $d(T_2G_4)$ and $d(T_4G_4)$, with those for binding $d(G_3T_4G_4)$, $d(G_4T_4G_4)$ and $d(T_4G_4T_4G_4)$. The results were consistent with the idea that nucleotides added to the 5′ end of a TTTTGGGG repeat likely interact only weakly, if at all, with protein. Enthalpy and entropy changes accompanying protein association correlate in a systematic manner to DNA length, however. To explain this pattern required careful consideration of the unbound DNA structure and changes in DNA structure coupled with DNA−protein association.

Several lines of experimental evidence suggest that single strand DNA in solution retains a large degree of organized structure built around base-base stacks. Calorimetric studies of structural transitions in oligo-A indicated that purine bases in a single strand RNA stack with a substantial enthalpy of –3 kcal/(mol base stack).[66,67] Other calorimetric analyses of 13 and 14 –base-pair DNAs containing dG•dC and dA•dT base-pairs, concluded that a substantial portion of enthalpy in duplex DNA is preserved in the dissociated single strands in the form of ordered helical base stacked structure.[55,68] Accordingly, in Figure 6(a) the uncomplexed form of $d(T_4G_4)$ is sketched in one hypothetical form closely resembling a strand of *B*-form DNA with bases making base-stacking interactions. We realize that free in solution, the single strand DNA may adopt a range of different conformations, with differing degrees of base stacking. However, the model highlights the point that single strand DNA is something other than a ''random coil'' ensemble and further illustrates that DNA−protein association is associated with bases unstacking[69] (Figure 6(c)).

The potential for forming organized structure is even greater in the longer DNAs, $d(G_3T_4G_4)$ $d(G_4T_4G_4)$ and $d(T_4G_4T_4G_4)$, since the additional GGG or GGGG tracts can form intramolecular G•G base-pairs with the 3′-terminal GGGG tract. A hypothetical model for $d(G_4T_4G_4)$ that is consistent with spectroscopic evidence is shown in Figure 6(b). The $d(G_4T_4G_4)$ molecule has been studied in lithium chloride by NMR,[70] and under low concentrations of sodium chloride by hydrogen-deuterium exchange, Raman spectroscopy, and gel electrophoresis.[21] These studies indicated the DNA forms a

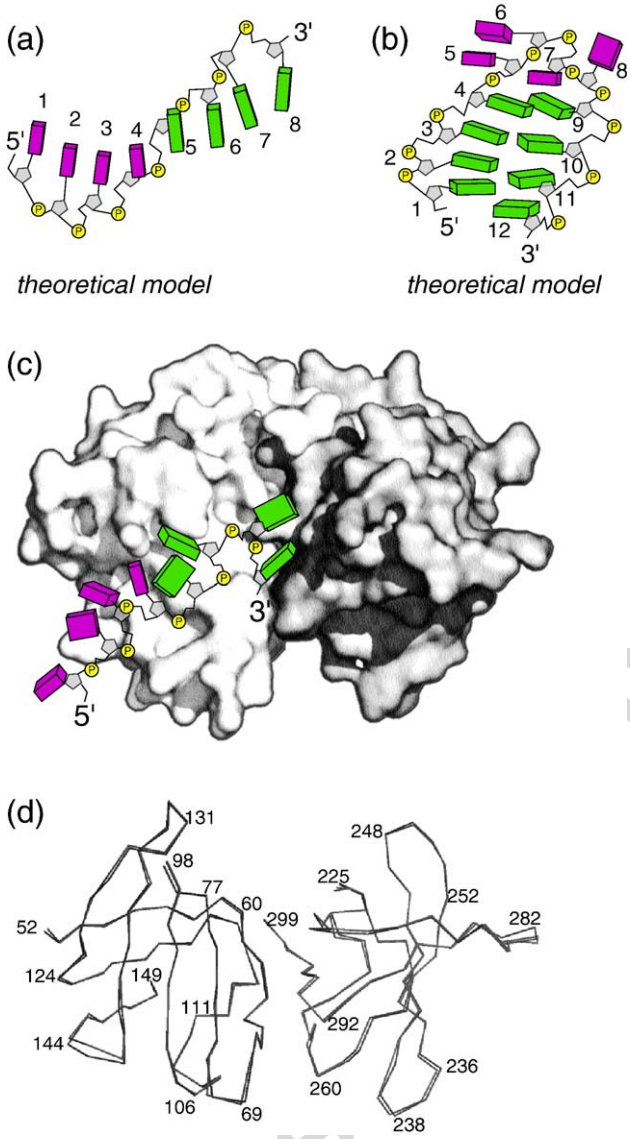

Figure 6. Structural models for telomere DNA and the alpha N-terminal domain. (a) $d(T_4G_4)$ in a conformation characterized by base-base stacking interactions. (b) Model of $d(G_4T_4G_4)$ in a hairpin conformation stabilized by G•G base-pairs. (c) $d(T_4G_4)$ as modelled in complex with protein. The position of the 5′-most nucleotide is not well determined by contact with protein, however the positions of the remaining nucleotides are clearly seen in electron density maps calculated for alpha N-terminal domain containing crystals[19,52] and crystals of $d(T_4G_4T_4G_4)$ complexed with the DNA-binding portions of alpha and beta (M.P.H., unpublished results). Note how bases are largely unstacked in the DNA–protein complex. (d) Superposition of OB folds contained within the alpha subunit N-terminal domain for the uncomplexed protein and for the DNA-bound protein. Coordinates for the two conformations were obtained from chains C and B of PDB ID 1K8G.[19] OB fold 1 comprises residues 52–149. OB fold 2 comprises residues 225–299. DNA appears to induce a subtle repositioning of OB folds. These subtle movements are best discerned in a movie available online (Movie S1).

hairpin structure with Hoogsteen G•G base-pairs.[21,70] In moderate concentrations of Na^+, K^+ or NH_4^+, $d(G_4T_4G_4)$ forms related but different structures with two molecules associating in G-quartet stabilized dimers.[30,71,72] Recent experiments in a closely related ciliate, *Stylonychia lemnae*, provide compelling evidence that telomere end binding proteins interact with such structures *in vivo*.[17]

Binding the $d(G_3T_4G_4)$, $d(G_4T_4G_4)$ and $d(T_4G_4T_4G_4)$ DNAs could be coupled to dissociation of G•G Hoogsteen base-pairs in addition to base unstacking. Formation of G•G Hoogsteen base-pairs is an enthalpy-driven process compensated by an unfavorable decrease in entropy. Differential scanning calorimetry with triplexed[73–77] and quadruplexed[78] DNAs measured fairly consistent values of $\Delta H = 3$–5 kcal mol^{-1} and $\Delta S = 9$ cal mol^{-1} K^{-1} for Hoogsteen base-pair dissociation. Resolving G•G base-pairs should therefore contribute an unfavorable positive enthalpy and favorable positive entropy to the binding reaction, exactly as is observed comparing $d(G_4T_4G_4)$, $d(G_4T_4G_4)$, and the A site of $d(T_4G_4T_4G_4)$ with $d(T_4G_4)$ (Figure 3(a)). The more favorable positive entropy change characterizing binding reactions of these longer DNAs likely derives from increased flexibility in the 5′-terminal nucleotides along with interactions of these nucleotides with solvent.

In addition to G•G base-pairs, $d(T_4)$ nucleotides shape the structure of $d(G_3T_4G_4)$, $d(G_4T_4G_4)$ and $d(T_4G_4T_4G_4)$. Evidence that $d(T_4)$ loops form ordered structures comes from studies of Watson–Crick base-paired hairpin DNAs and G-quartet stabilized DNA dimers. In the context of a Watson–Crick base-paired hairpin DNA, $d(T)_4$ loops augmented thermal stability of the duplex to a greater degree than $d(C)_4$, $d(G)_4$, or $d(A)_4$.[79,80] NMR and X-ray studies confirmed that the $d(T_4)$ nucleotides within such loops are highly organized.[81–84] Within G-quartet stabilized DNA dimers, $d(T_4)$ nucleotides connect diagonally opposed GGGG tracts. The same $d(T_4)$ structure, with T–T and T–G base stacking, was observed in several crystal forms[31,85] and in solution

by NMR,[28,86] further indication that d(T_4) nucleotides have an intrinsic propensity for forming structure. Since d(T_4) loops adopt organized structures in several different contexts, we believe that the d(T_4) loop of hairpins is structured in d($G_3T_4G_4$), d($G_4T_4G_4$) and d($T_4G_4T_4G_4$) DNAs and that protein binding reactions are coupled with breaking this d(T_4) structure.

These proposed structural changes in the DNA may be accompanied by subtle adjustments in the structure of the alpha N-terminal domain coupled with binding DNA. The N-terminal domain is constructed from two oligonucleotide/oligosaccharide/oligopeptide binding folds[14,87] (OB folds). Figure 6(d) shows the superposed alpha carbon traces of these two OB folds from uncomplexed and DNA-bound forms of the N-terminal domain. Overall, the structures are very similar with a RMS deviation of less than 0.4 Å for all alpha carbons. As noted previously,[19] atom displacements within each OB fold are concerted, however, so that OB fold-1 and OB fold-2 rotate slightly with respect to each other. This subtle relative movement of OB folds is more easily detected in a two-frame movie (Movie S1), or as intramolecular distance differences that show certain loops have shifted by 1.8 Å comparing the uncomplexed and DNA-bound proteins. Although these DNA-binding induced shifts in atom positions are small, they may trigger more dramatic changes in structure of the alpha protein that subsequently lead to protein–protein association between alpha and beta subunits in the full *O. nova* telomere end binding protein. Such an allosteric mechanism could potentially explain cooperativity between DNA binding and protein–protein association uncovered in earlier studies of the *O. nova* telomere end binding protein[13] and recently described quantitatively in the form of a thermodynamic binding cycle.[18]

Heat capacity change of DNA–protein binding

The structural changes coupled to DNA–protein binding proposed above could contribute to the observed change in heat capacity (ΔCp) associated with binding reactions. It is widely recognized that heat capacity is an important thermodynamic property intrinsic to folding and binding of macromolecules, yet the molecular basis for changes in heat capacity is still not well understood.[88–90] A negative ΔCp is considered a hallmark for sequence-specific *versus* non-sequence-specific recognition of double-stranded DNA.[59,91–93] Single strand DNA is less restrained and likely samples more conformations compared with double-stranded DNA. Since heat capacity depends critically on the degeneracy of states accessible within a molecular ensemble,[90] trends observed for binding double-stranded DNA may or may not apply generally to binding single strand DNA. The present work provides examples of ΔCp measured for the sequence-specific recognition of telomere-derived single strand DNAs.

A modest negative change in heat capacity ($\Delta Cp < 0$) characterizes association of the alpha N-terminal domain with each of the DNAs, and the magnitude of this negative heat capacity change ($|\Delta Cp|$) increases with DNA length (Figure 3(d)). The DNA length dependence of $|\Delta Cp|$ is not uniform, however, with a significantly larger $|\Delta Cp|$ associated with the binding reactions of longer DNAs, d($G_3T_4G_4$), d($G_4T_4G_4$) and site A of d($T_4G_4T_4G_4$), compared with those of the shorter DNAs, d(T_2G_4) and d(T_4G_4). The simplest interpretation is that, along with base unstacking,[69] dissociation of intramolecular G•G base-pairs contributes to $|\Delta Cp|$. The heat capacity change measured for DNA strand dissociation[55,94–98] is positive, and thus apparently inconsistent with the idea that G•G dissociation should lead to a more negative heat capacity change. However, since base unstacking and G•G disscociation processes are coupled with protein binding, an ''apparent'' negative ΔCp term is anticipated,[69,90] exactly as measured for the binding reactions of d($G_3T_4G_4$), d($G_4T_4G_4$) and the A site of d($T_4G_4T_4G_4$). The same kind of structural transitions may explain the very large and negative heat capacity change observed for binding of G-rich DNA by F factor relaxase.[99] Larger than expected negative heat capacity changes for DNA-binding events have often been discussed in terms of changes in protein structure coupled with binding.[99–102] Structural changes in single strand DNA coupled with binding can also contribute to the final value of ΔCp.[69,103,104] The heat capacity changes measured in our system are consistent with the idea that binding DNA is accompanied by base unstacking and base-pair dissociation in the telomere single strand DNA.

Biological function and evolution of the telomere end binding protein

Although the molecular picture is far from complete it is important to consider how the protein binding events measured here relate to subsequent interactions essential for the biology of telomeres. The thermodynamic binding properties of alpha and its N-terminal domain described here and in previous *in vitro* studies[18] show how alpha might participate in loading the 3′-terminal single strand DNA into a productive complex with telomerase by resolving G•G base-paired structures (Figure 7). Because telomerase from *O. nova* is unable to initiate extension for folded forms of the single strand DNA,[38] and formation of the alpha–DNA complex is incompatible with these folded structures, protein binding can maintain the DNA in an extended conformation required for productive interaction with telomerase. Froelich-Ammon *et al.* observed that formation of an alpha•DNA•alpha complex diminished telomerase activity, whereas an alpha•DNA complex served as a suitable primer for telomerase-mediated DNA extension.[105] In the closely related ciliate *S. lemnae*,[17] alpha is important for attachment of DNA to the nuclear envelope and for regulation of DNA structures that are immunoreactive with G-quartet-recognizing antibodies.[17]

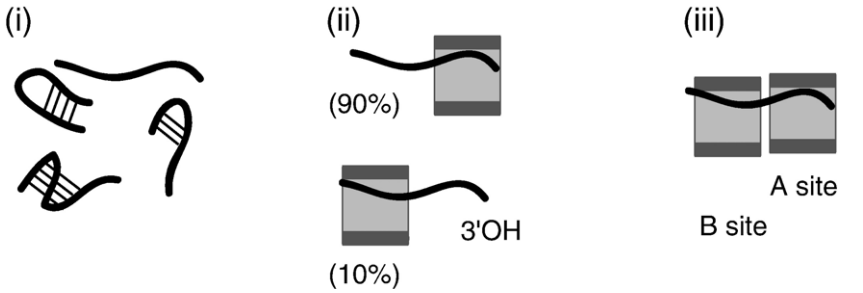

Figure 7. Model for interaction of the *O. nova* alpha N-terminal domain with telomere DNA. In the uncomplexed state (i), telomere DNA is represented by an ensemble with some members containing structures stabilized by G•G base-pairing. The alpha N-terminal domain (square shape) binds extended DNA molecules to form an initial DNA–protein complex (ii). This binding event resolves intramolecular DNA interactions, including base-pairing and base stacking interactions. Association with protein at the second site results in a DNA–protein complex with two DNA-binding domains associated with one DNA (iii).

These results demonstrate an active role for alpha in forming and resolving DNA structures in a manner timed with DNA replication.[17] According to our measures, given a limited amount of alpha the 3′-terminal TTTTGGGG binding site will be occupied by protein leaving the 3′-distal nucleotides exposed and thus free to interact with telomerase.

If alpha binds preferentially to the 3′-terminal site, how does telomerase gain access to the 3′ OH group required for nucleotidyl transfer and DNA extension? The single strand DNA fragments that we have studied only approximate complete telomeres, and the presence of a complementary DNA strand present in authentic telomere DNA may guide the alpha protein away from the 3′-end. Alternatively, as suggested for human telomere binding protein POT1,[106] there may be other cellular factors that move the end-specific protein from a 3′-terminal site to an internal binding site. In the full *O. nova* telomere complex, beta, the second telomere protein, could be a key player in this interconversion event. Upon association with this second protein subunit, the 3′-terminal TTTTGGG portion of the DNA makes new interactions with a binding pocket created by juxtaposition of alpha and beta subunits. Dissociation of beta, potentially triggered by phosphorylation,[17,107] would leave the 3′-terminal region exposed and ready to interact with telomerase.

During *de novo* telomere synthesis and telomere replication, the 3′-terminal single strand DNA is extended well past the 16 nt length found at mature telomeres.[64,65,108,109] Binding to 3′-distal repeats as observed during the second binding transition associated with site B may resemble the function of an ancestral, non-sequence-specific single strand DNA binding protein from which the telomere end binding protein was derived. All organisms have such a non-sequence-specific single strand DNA binding protein that acts at the replication fork to maintain separated strands in a form suitable to act as template for DNA polymerase during DNA synthesis.[110,111] The crystal structure of the human replication protein A single strand DNA binding domain[112] revealed two OB folds[87] tandemly arrayed to interact with eight nucleotides of single strand DNA. In bacteria the ssDNA binding protein is a homo-tetramer, also constructed from multiple OB folds.[113] The N-terminal domain of alpha comprises two such OB folds.[14] Structural homology among these proteins lends support to the idea that high affinity, sequence-specific recognition of 3′-terminal single strand DNA evolved through gene duplication and divergence.[14,15] The site B interactions described here further suggest that interaction with non-terminal telomere single strand DNA was retained even as the protein became more specialized at capping the very ends of telomeres.

Experimental Methods

Protein and DNA purification

Alpha (aa 1–495) and alpha N-terminal domain (aa 1–326) from the *O. nova* telomere end binding protein were expressed in *Escherichia coli* BL21 (DE3) pLysS and purified as described,[14,18] except that the final size-exclusion chromatography was developed in binding buffer consisting of 0.15 M lithium chloride and 0.05 M Tris (pH 7.5). Protein concentration was determined with the use of absorbance at 280 nm and extinction coefficients calculated according to the number of tyrosine and tryptophan residues in each protein.[114]

Trityl-group protected DNA oligonucleotides were synthesized by the University of Utah DNA/peptide research core facility as multiple 1 μmol batches and resuspended in HPLC buffer A (10 mM triethylamine ammonium acetate (pH 6.0)). Protected oligonucleotides were purified by reverse phase HPLC using 40 min linear gradients from 14%–22% HPLC buffer B (10 mM triethylamine ammonium acetate (pH 6.0), 75% acetonitrile). The trityl protecting group was removed by resuspending the lyophilized material with 10 ml of 80% acetic acid. Reactions were incubated at room temperature for 20 min and then quenched with 50 ml of ice cold water followed by lyophilization. Deprotected DNAs were resuspended in HPLC buffer A, insoluble material was removed by filtration through a 0.2 μm PVDF low-protein binding filter (Millipore), and a second reverse phase HPLC purification was developed using a 30 min linear gradient from 5%–10% HPLC buffer B. Solvents were removed by repeatedly resuspending lyophilized material with 10 ml of water. Pure DNA oligonucleotides were resuspended with a minimal volume of binding buffer and dialyzed overnight against ×1000 volumes of binding buffer. DNA concentrations were determined with use of absorbance at 260 nm and molar extinction coefficients calculated using the nearest neighbor model.[115]

Since sodium ions promote the formation of kinetically inert G-quartet stabilized DNA quadruplexes while lithium ions do not stabilize G-quartets, calorimetry experiments described here were conducted in lithium chloride-containing solutions and reasonable precautions were taken to exclude potassium or sodium ions. For instance, EDTA was purchased as the free acid and titrated with Tris base, not sodium hydroxide, and the pHs of buffered solutions were measured in separate aliquots to avoid contact with a potassium chloride equilibrated electrode.

Native gel electrophoresis

7.6% (w/v) acrylamide:0.2% bis-acrylamide gels were cast in a 12 cm×14 cm horizontal tray with *n*-butanol layered on top to promote uniform polymerization. The 89 ml gel mixture contained 10% 0.25×TBE running buffer. Polymerization proceeded for 2 h. The gel was submerged in 0.25×TBE (22.5 mM Tris, 22.5 mM boric acid, 0.5 mM EDTA) and pre-run for 30 min at 110 mV and +4 °C. 10 μl samples were loaded in each lane and electrophoresis proceeded at 110 mV for an additional 4 h or 5 h, respectively, for alpha N-terminal domain and full-length alpha subunit. The concentration of protein in each sample was 110 μM for alpha N-terminal domain or 80 μM for the full-length alpha subunit.

Isothermal titration calorimetry

Isothermal titration calorimetry was performed with use of a MicroCal VP-ITC instrument. To prepare binding components, DNA and protein solutions were loaded into separate 2.5 ml, 1000 molecular weight cutoff dialysis units (Amersham Biosciences) and dialyzed at +4 °C overnight against one common 500 ml volume of binding buffer. The DNA was diluted to 30 μM, degassed, and placed in the calorimeter sample cell. For titration of all DNAs except d($T_4G_4T_4G_4$), the protein was diluted to 330 μM, degassed, and placed in the titration syringe. For titration of d($T_4G_4T_4G_4$), the protein concentration was adjusted to 540 μM. All dilutions were prepared by adding the equilibrating solution obtained after dialysis was complete. During the experiment, the sample cell was stirred at 300 rpm. Titration of all DNAs except d($T_4G_4T_4G_4$) started with one initial 3-μl injection followed by 30 8-μl injections spaced by 200-s intervals. Titration of d($T_4G_4T_4G_4$) started with one initial 2-μl injection followed by 41 6-μl injections.

Heat evolved or absorbed upon each injection was quantified by integration of the power *versus* time plot corrected for the heat of injecting buffer alone into the sample solution. Enthalpy (ΔH), dissociation constant ($K_{D\text{-}DNA}$), entropy (ΔS), and stoichiometry of binding (n) were obtained by non-linear least-squares fitting of the heat *versus* protein:DNA data to one-site or two-site models provided with the data analysis software (Origin). Free energy of binding (ΔG) was calculated from Gibb's equation $\Delta G = \Delta H - T\Delta S$.

Calculation of the relative proportion of non-electrostatic free energy at 0.15 M lithium chloride

Plots of $\ln(K_{A\text{-}DNA})$ *versus* $\ln([LiCl])$ defined by four measurements were extrapolated to obtain the free energy of DNA association at 1 M LiCl, $\Delta G_{1M\ LiCl}$. At this reference point, free energy of DNA association can be assigned to non-electrostatic interactions arising from hydrogen bonds, van der Waals contacts, and hydrophobic interactions plus the energy of each ion pair at standard conditions. To obtain the relative contributions of electrostatic and non-electrostatic interactions at physiological salt concentration of 0.15 M LiCl, we applied equations (1) and (2):

$$\Delta G_{0.15\ M\ LiCl} = \Delta G_{non\text{-}electrostatic} + \Delta G_{electrostatic} \quad (1)$$

$$\Delta G_{non\text{-}electrostatic} = \Delta G_{1\ M\ LiCl} - Z\Delta G_{ion\text{-}pair} \quad (2)$$

where $\Delta G_{ion\text{-}pair}$ describes the energetic contribution of a lysine/arginine-phosphate ion pair under standard state conditions estimated at +0.18 kcal/mol ion pair,[46,116–118] and Z is the number of such ion pairs determined from the slope of $\ln(K_{A\text{-}DNA})$ *versus* $\ln([LiCl])$ data.

Building model structures of d(T_4G_4) and d($G_4T_4G_4$)

Models of single strand DNA structures free in solution were constructed using coordinates for *B*-form DNA (PDB ID 9BNA)[119,120] and antiparallel/parallel stranded G-quadruplex DNA (PDB ID 1JB785). Extra strands were deleted and the sequence of each DNA was altered to be consistent with the d(T_4G_4) repeat of *O. nova* telomeres. In the case of hairpin structures, the connecting d(T_4) loop was repositioned in *O*[121] to splice base-paired GGGG tracts into one DNA strand. Structures were energy minimized with *CNS*[122] using updated DNA-RNA parameters,[123] and additional restraints describing hydrogen bonds between G•G base-pairs.

The structure of d(T_4G_4) DNA as complexed with alpha N-terminal domain was obtained from examination of three crystal structures: d(T_2G_4) in complex with alpha N-terminal domain[19]; d(T_4G_4) in complex with alpha,[52] and d($T_4G_4T_4G_4$) complexed with the DNA-binding portions of alpha and beta subunits (M.P.H., unpublished results). The DNA in all three examples are super-imposable for internal G nucleotides of d($T_4G_4T_4G_4$) and 3′-terminal G nucleotides of d(T_2G_4) and d(T_4G_4). Details concerning the 5′-terminal nucleotides vary slightly among the structures. Our current model, for the 5′-TTTT portion, is consistent with photo-crosslinking experiments, conducted in solution, which revealed a close contact between Tyr142 of alpha and the nucleotide at position 2 in d($T_4G_4T_4G_4$).[124]

Acknowledgements

We are grateful for insightful critical comments provided by anonymous reviewers and S. Classen. We thank D. Goldenberg for use of an isothermal titration calorimeter and for helpful suggestions during preparation of the manuscript. 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).

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2006.02.082

References

1. Greider, C. W. & Blackburn, E. H. (1987). The telomere terminal transferase of Tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell,* **51**, 887–898.
2. Greider, C. W. & Blackburn, E. H. (1985). Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. *Cell,* **43**, 405–413.
3. Klobutcher, L. A., Swanton, M. T., Donini, P. & Prescott, D. M. (1981). 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,* **78**, 3015–3019.
4. McElligott, R. & Wellinger, R. J. (1997). The terminal DNA structure of mammalian chromosomes. *EMBO J.* **16**, 3705–3714.
5. Griffith, J. D., Comeau, L., Rosenfield, S., Stansel, R. M., Bianchi, A., Moss, H. & de Lange, T. (1999). Mammalian telomeres end in a large duplex loop. *Cell,* **97**, 503–514.
6. Murti, K. G. & Prescott, D. M. (1999). Telomeres of polytene chromosomes in a ciliated protozoan terminate in duplex DNA loops. *Proc. Natl Acad. Sci. USA,* **96**, 14436–14439.
7. Stansel, R. M., de Lange, T. & Griffith, J. D. (2001). T-loop assembly in vitro involves binding of TRF2 near the 3′ telomeric overhang. *EMBO J.* **20**, 5532–5540.
8. de Lange, T. (2004). T-loops and the origin of telomeres. *Nature Rev. Mol. Cell Biol.* **5**, 323–329.
9. Nikitina, T. & Woodcock, C. L. (2004). Closed chromatin loops at the ends of chromosomes. *J. Cell Biol.* **166**, 161–165.
10. Price, C. M. & Cech, T. R. (1989). Properties of the telomeric DNA-binding protein from *Oxytricha nova*. *Biochemistry,* **28**, 769–774.
11. Gottschling, D. E. & Zakian, V. A. (1986). Telomere proteins: specific recognition and protection of the natural termini of Oxytricha macronuclear DNA. *Cell,* **47**, 195–205.
12. Fang, G. & Cech, T. R. (1993). Oxytricha telomere-binding protein: DNA-dependent dimerization of the alpha and beta subunits. *Proc. Natl Acad. Sci. USA,* **90**, 6056–6060.
13. Fang, G., Gray, J. T. & Cech, T. R. (1993). Oxytricha telomere-binding protein: separable DNA-binding and dimerization domains of the alpha-subunit. *Genes Dev.* **7**, 870–882.
14. Horvath, M. P., Schweiker, V. L., Bevilacqua, J. M., Ruggles, J. A. & Schultz, S. C. (1998). Crystal structure of the Oxytricha nova telomere end binding protein complexed with single strand DNA. *Cell*, **95**, 963–974.
15. Theobald, D. L., Cervantes, R. B., Lundblad, V. & Wuttke, D. S. (2003). Homology among telomeric end-protection proteins. *Structure (Camb.),* **11**, 1049–1050.
16. Baumann, P. & Cech, T. R. (2001). Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science,* **292**, 1171–1175.
17. Paeschke, K., Simonsson, T., Postberg, J., Rhodes, D. & Lipps, H. J. (2005). Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo. *Nature Struct. Mol. Biol.* **12**, 847–854.
18. Buczek, P., Orr, R. S., Pyper, S. R., Shum, M., Kimmel, E., Ota, I. & Horvath, M. P. (2005). Binding linkage in a telomere DNA-protein complex at the ends of *Oxytricha nova* chromosomes. *J. Mol. Biol.* **350**, 938–952.
19. Classen, S., Ruggles, J. A. & Schultz, S. C. (2001). 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.* **314**, 1113–1125.
20. Classen, S., Lyons, D., Cech, T. R. & Schultz, S. C. (2003). Sequence-specific and 3′-end selective single-strand DNA binding by the *Oxytricha nova* telomere end binding protein alpha subunit. *Biochemistry,* **42**, 9269–9277.
21. Laporte, L. & Thomas, G. J. Jr (1998). A hairpin conformation for the 3′ overhang of Oxytricha nova telomeric DNA. *J. Mol. Biol.* **281**, 261–270.
22. Gray, J. T., Celander, D. W., Price, C. M. & Cech, T. R. (1991). Cloning and expression of genes for the Oxytricha telomere-binding protein: specific subunit interactions in the telomeric complex. *Cell,* **67**, 807–814.
23. Sen, D. & Gilbert, W. (1988). Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature,* **334**, 364–366.
24. Sundquist, W. I. & Klug, A. (1989). Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. *Nature,* **342**, 825–829.
25. Williamson, J. R., Raghuraman, M. K. & Cech, T. R. (1989). Monovalent cation-induced structure of telomeric DNA: the G-quartet model. *Cell,* **59**, 871–880.
26. Sen, D. & Gilbert, W. (1990). A sodium-potassium switch in the formation of four-stranded G4-DNA. *Nature,* **344**, 410–414.
27. Sen, D. & Gilbert, W. (1992). Guanine quartet structures. *Methods Enzymol.* **211**, 191–199.
28. Smith, F. W. & Feigon, J. (1992). Quadruplex structure of Oxytricha telomeric DNA oligonucleotides. *Nature,* **356**, 164–168.
29. Smith, F. W. & Feigon, J. (1993). Strand orientation in the DNA quadruplex formed from the Oxytricha telomere repeat oligonucleotide d(G4T4G4) in solution. *Biochemistry,* **32**, 8682–8692.
30. Schultze, P., Hud, N. V., Smith, F. W. & Feigon, J. (1999). The effect of sodium, potassium and ammonium ions on the conformation of the dimeric quadruplex formed by the *Oxytricha nova* telomere repeat oligonucleotide d(G(4)T(4)G(4)). *Nucl. Acids Res.* **27**, 3018–3028.
31. Haider, S., Parkinson, G. N. & Neidle, S. (2002). Crystal structure of the potassium form of an *Oxytricha nova* G-quadruplex. *J. Mol. Biol.* **320**, 189–200.
32. Riou, J. F., Guittat, L., Mailliet, P., Laoui, A., Renou, E. Petitgenet, O. *et al.* (2002). Cell senescence and telomere shortening induced by a new series of specific G-quadruplex DNA ligands. *Proc. Natl Acad. Sci. USA,* **99**, 2672–2677.
33. Gomez, D., Aouali, N., Londono-Vallejo, A., Lacroix, L., Megnin-Chanet, F. Lemarteleur, T. *et al.* (2003). Resistance to the short term antiproliferative activity of the G-quadruplex ligand 12459 is associated with telomerase overexpression and telomere capping alteration. *J. Biol. Chem.* **278**, 50554–50562.
34. Gomez, D., Aouali, N., Renaud, A., Douarre, C., Shin-Ya, K. Tazi, J. *et al.* (2003). Resistance to senescence induction and telomere shortening by a G-quadruplex ligand inhibitor of telomerase. *Cancer Res.* **63**, 6149–6153.
35. Raghuraman, M. K. & Cech, T. R. (1990). Effect of monovalent cation-induced telomeric DNA structure on the binding of Oxytricha telomeric protein. *Nucl. Acids Res.* **18**, 4543–4552.

36. Fang, G. & Cech, T. R. (1993). Characterization of a G-quartet formation reaction promoted by the beta-subunit of the Oxytricha telomere-binding protein. *Biochemistry,* **32**, 11646–11657.
37. Fang, G. & Cech, T. R. (1993). The beta subunit of Oxytricha telomere-binding protein promotes G-quartet formation by telomeric DNA. *Cell,* **74**, 875–885.
38. Zahler, A. M., Williamson, J. R., Cech, T. R. & Prescott, D. M. (1991). Inhibition of telomerase by G-quartet DNA structures. *Nature,* **350**, 718–720.
39. Zaug, A. J., Podell, E. R. & Cech, T. R. (2005). Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro. *Proc. Natl Acad. Sci. USA,* **102**, 10864–10869.
40. Wyman, J. & Gill, S. J. (1990). *Binding and Linkage. Functional Chemistry of Biological Macromolecules.* University Science Books, Mill Valley, CA.
41. Jelesarov, I. & Bosshard, H. R. (1999). Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition. *J. Mol. Recognit.* **12**, 3–18.
42. Pierce, M. M., Raman, C. S. & Nall, B. T. (1999). Isothermal titration calorimetry of protein-protein interactions. *Methods,* **19**, 213–221.
43. Velazquez-Campoy, A. & Freire, E. (2005). ITC in the post-genomic era...? Priceless. *Biophys. Chem.* **115**, 115–124.
44. Acevedo, O. L., Dickinson, L. A., Macke, T. J. & Thomas, C. A. Jr (1991). The coherence of synthetic telomeres. *Nucl. Acids Res.* **19**, 3409–3419.
45. Kozlov, A. G. & Lohman, T. M. (1998). Calorimetric studies of E. coli SSB protein-single-stranded DNA interactions. Effects of monovalent salts on binding enthalpy. *J. Mol. Biol.* **278**, 999–1014.
46. Record, M. T. Jr, Lohman, M. L. & De Haseth, P. (1976). Ion effects on ligand-nucleic acid interactions. *J. Mol. Biol.* **107**, 145–158.
47. Dragan, A. I., Liggins, J. R., Crane-Robinson, C. & Privalov, P. L. (2003). The energetics of specific binding of AT-hooks from HMGA1 to target DNA. *J. Mol. Biol.* **327**, 393–411.
48. Lohman, T. M., Overman, L. B., Ferrari, M. E. & Kozlov, A. G. (1996). A highly salt-dependent enthalpy change for *Escherichia coli* SSB protein-nucleic acid binding due to ion-protein interactions. *Biochemistry,* **35**, 5272–5279.
49. Lei, M., Baumann, P. & Cech, T. R. (2002). Cooperative binding of single-stranded telomeric DNA by the Pot1 protein of *Schizosaccharomyces pombe*. *Biochemistry,* **41**, 14560–14568.
50. Anderson, E. M., Halsey, W. A. & Wuttke, D. S. (2003). Site-directed mutagenesis reveals the thermodynamic requirements for single-stranded DNA recognition by the telomere-binding protein Cdc13. *Biochemistry,* **42**, 3751–3758.
51. Mitton-Fry, R. M., Anderson, E. M., Theobald, D. L., Glustrom, L. W. & Wuttke, D. S. (2004). Structural basis for telomeric single-stranded DNA recognition by yeast Cdc13. *J. Mol. Biol.* **338**, 241–255.
52. Peersen, O. B., Ruggles, J. A. & Schultz, S. C. (2002). Dimeric structure of the *Oxytricha nova* telomere end-binding protein alpha-subunit bound to ssDNA. *Nature Struct. Biol.* **9**, 182–187.
53. Makhatadze, G. I. & Privalov, P. L. (1990). Heat capacity of proteins: I. Partial molar heat capacity of individual amino acid residues in aqueous solution: hydration effect. *J. Mol. Biol.* **213**, 375–384.
54. Lopez, M. M., Yutani, K. & Makhatadze, G. I. (1999). Interactions of the major cold shock protein of *Bacillus subtilis* CspB with single-stranded DNA templates of different base composition. *J. Biol. Chem.* **274**, 33601–33608.
55. Holbrook, J. A., Capp, M. W., Saecker, R. M. & Record, M. T. Jr (1999). Enthalpy and heat capacity changes for formation of an oligomeric DNA duplex: interpretation in terms of coupled processes of formation and association of single-stranded helices. *Biochemistry,* **38**, 8409–8422.
56. Taylor, J. W., Greenfield, N. J., Wu, B. & Privalov, P. L. (1999). A calorimetric study of the folding-unfolding of an alpha-helix with covalently closed N and C-terminal loops. *J. Mol. Biol.* **291**, 965–976.
57. Dragan, A. I. & Privalov, P. L. (2002). Unfolding of a leucine zipper is not a simple two-state transition. *J. Mol. Biol.* **321**, 891–908.
58. Privalov, P. L. & Gill, S. J. (1988). Stability of protein structure and hydrophobic interaction. *Advan. Protein Chem.* **39**, 191–234.
59. Ha, J. H., Spolar, R. S. & Record, M. T. Jr (1989). Role of the hydrophobic effect in stability of site-specific protein-DNA complexes. *J. Mol. Biol.* **209**, 801–816.
60. Murphy, K. P., Privalov, P. L. & Gill, S. J. (1990). Common features of protein unfolding and dissolution of hydrophobic compounds. *Science,* **247**, 559–561.
61. Lei, M., Podell, E. R., Baumann, P. & Cech, T. R. (2003). DNA self-recognition in the structure of Pot1 bound to telomeric single-stranded DNA. *Nature,* **426**, 198–203.
62. Lei, M., Podell, E. R. & Cech, T. R. (2004). Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection. *Nature Struct. Mol. Biol.* **11**, 1223–1229.
63. Mitton-Fry, R. M., Anderson, E. M., Hughes, T. R., Lundblad, V. & Wuttke, D. S. (2002). Conserved structure for single-stranded telomeric DNA recognition. *Science,* **296**, 145–147.
64. Fan, X. & Price, C. M. (1997). Coordinated regulation of G- and C strand length during new telomere synthesis. *Mol. Biol. Cell,* **8**, 2145–2155.
65. Vermeesch, J. R. & Price, C. M. (1994). Telomeric DNA sequence and structure following de novo telomere synthesis in *Euplotes crassus*. *Mol. Cell. Biol.* **14**, 554–566.
66. Davis, R. C. & Tinoco, I. Jr (1968). Temperature-dependent properties of dinucleoside phosphates. *Biopolymers,* **6**, 223–242.
67. Breslauer, K. J. & Sturtevant, J. M. (1977). A calorimetric investigation of single stranded base stacking in the ribo-oligonucleotide A7. *Biophys. Chem.* **7**, 205–209.
68. Vesnaver, G. & Breslauer, K. J. (1991). The contribution of DNA single-stranded order to the thermodynamics of duplex formation. *Proc. Natl Acad. Sci. USA,* **88**, 3569–3573.
69. Kozlov, A. G. & Lohman, T. M. (1999). Adenine base unstacking dominates the observed enthalpy and heat capacity changes for the *Escherichia coli* SSB tetramer binding to single-stranded oligoadenylates. *Biochemistry,* **38**, 7388–7397.
70. Choi, K. H. & Choi, B. S. (1994). Formation of a hairpin structure by telomere 3′ overhang. *Biochim. Biophys. Acta,* **1217**, 341–344.
71. Hud, N. V., Schultze, P., Sklenar, V. & Feigon, J. (1999). Binding sites and dynamics of ammonium ions in a

telomere repeat DNA quadruplex. *J. Mol. Biol.* **285**, 233–243.
72. Wu, G. & Wong, A. (2004). Solid-state 23Na NMR determination of the number and coordination of sodium cations bound to *Oxytricha nova* telomere repeat d(G4T4G4). *Biochem. Biophys. Res. Commun.* **323**, 1139–1144.
73. Pilch, D. S., Brousseau, R. & Shafer, R. H. (1990). Thermodynamics of triple helix formation: spectrophotometric studies on the d(A)10.2d(T)10 and d(C+3T4C+3).d(G3A4G3).d(C3T4C3) triple helices. *Nucl. Acids Res.* **18**, 5743–5750.
74. Ohms, J. & Ackermann, T. (1990). Thermodynamics of double- and triple-helical aggregates formed by self-complementary oligoribonucleotides of the type rAxUy. *Biochemistry,* **29**, 5237–5244.
75. Xodo, L. E., Manzini, G. & Quadrifoglio, F. (1990). Spectroscopic and calorimetric investigation on the DNA triplex formed by d(CTCTTCTTTCTTTTC-TTTCTTCTC) and d(GAGAAGAAAGA) at acidic pH. *Nucl. Acids Res.* **18**, 3557–3564.
76. Plum, G. E. & Breslauer, K. J. (1995). Thermodynamics of an intramolecular DNA triple helix: a calorimetric and spectroscopic study of the pH and salt dependence of thermally induced structural transitions. *J. Mol. Biol.* **248**, 679–695.
77. Volker, J., Osborne, S. E., Glick, G. D. & Breslauer, K. J. (1997). Thermodynamic properties of a conformationally constrained intramolecular DNA triple helix. *Biochemistry,* **36**, 756–767.
78. Jin, R., Gaffney, B. L., Wang, C., Jones, R. A. & Breslauer, K. J. (1992). Thermodynamics and structure of a DNA tetraplex: a spectroscopic and calorimetric study of the tetramolecular complexes of d(TG3T) and d(TG3T2G3T). *Proc. Natl Acad. Sci. USA,* **89**, 8832–8836.
79. Senior, M. M., Jones, R. A. & Breslauer, K. J. (1988). Influence of loop residues on the relative stabilities of DNA hairpin structures. *Proc. Natl Acad. Sci. USA,* **85**, 6242–6246.
80. Paner, T. M., Amaratunga, M., Doktycz, M. J. & Benight, A. S. (1990). Analysis of melting transitions of the DNA hairpins formed from the oligomer sequences d[GGATAC(X)4GTATCC] (X=A, T, G, C). *Biopolymers,* **29**, 1715–1734.
81. Hare, D. R. & Reid, B. R. (1986). Three-dimensional structure of a DNA hairpin in solution: two-dimensional NMR studies and distance geometry calculations on d(CGCGTTTTCGCG). *Biochemistry,* **25**, 5341–5350.
82. Ikuta, S., Chattopadhyaya, R., Ito, H., Dickerson, R. E. & Kearns, D. R. (1986). NMR study of a synthetic DNA hairpin. *Biochemistry,* **25**, 4840–4849.
83. Wolk, S. K., Hardin, C. C., Germann, M. W., van de Sande, J. H. & Tinoco, I. Jr (1988). Comparison of the B- and Z-form hairpin loop structures formed by d(CG)5T4(CG)5. *Biochemistry,* **27**, 6960–6967.
84. Chattopadhyaya, R., Ikuta, S., Grzeskowiak, K. & Dickerson, R. E. (1988). X-ray structure of a DNA hairpin molecule. *Nature,* **334**, 175–179.
85. Horvath, M. P. & Schultz, S. C. (2001). DNA G-quartets in a 1.86 A resolution structure of an *Oxytricha nova* telomeric protein-DNA complex. *J. Mol. Biol.* **310**, 367–377.
86. Schultze, P., Smith, F. W. & Feigon, J. (1994). Refined solution structure of the dimeric quadruplex formed from the Oxytricha telomeric oligonucleotide d (GGGGTTTTGGGG). *Structure,* **2**, 221–233.
87. Murzin, A. G. (1993). OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.* **12**, 861–867.
88. Gomez, J., Hilser, V. J., Xie, D. & Freire, E. (1995). The heat capacity of proteins. *Proteins: Struct. Funct. Genet.* **22**, 404–412.
89. Cooper, A., Johnson, C. M., Lakey, J. H. & Nollmann, M. (2001). Heat does not come in different colours: entropy-enthalpy compensation, free energy windows, quantum confinement, pressure perturbation calorimetry, solvation and the multiple causes of heat capacity effects in biomolecular interactions. *Biophys. Chem.* **93**, 215–230.
90. Prabhu, N. V. & Sharp, K. A. (2005). Heat capacity in proteins. *Annu. Rev. Phys. Chem.* **56**, 521–548.
91. Lundback, T., Cairns, C., Gustafsson, J. A., Carlstedt-Duke, J. & Hard, T. (1993). Thermodynamics of the glucocorticoid receptor-DNA interaction: binding of wild-type GR DBD to different response elements. *Biochemistry,* **32**, 5074–5082.
92. Lundback, T., Zilliacus, J., Gustafsson, J. A., Carlstedt-Duke, J. & Hard, T. (1994). Thermodynamics of sequence-specific glucocorticoid receptor-DNA interactions. *Biochemistry,* **33**, 5955–5965.
93. Lundback, T. & Hard, T. (1996). Sequence-specific DNA-binding dominated by dehydration. *Proc. Natl Acad. Sci. USA,* **93**, 4754–4759.
94. Suurkuusk, J., Alvarez, J., Freire, E. & Biltonen, R. (1977). Calorimetric determination of the heat capacity changes associated with the conformational transitions of polyriboadenylic acid and polyribouridylic acid. *Biopolymers,* **16**, 2641–2652.
95. Filimonov, V. V. & Privalov, P. L. (1978). Thermodynamics of base interaction in (A)*n* and (A.U)*n*. *J. Mol. Biol.* **122**, 465–470.
96. Chalikian, T. V., Volker, J., Plum, G. E. & Breslauer, K. J. (1999). A more unified picture for the thermodynamics of nucleic acid duplex melting: a characterization by calorimetric and volumetric techniques. *Proc. Natl Acad. Sci. USA,* **96**, 7853–7858.
97. Rouzina, I. & Bloomfield, V. A. (1999). Heat capacity effects on the melting of DNA: I. General aspects. *Biophys. J.* **77**, 3242–3251.
98. Jelesarov, I., Crane-Robinson, C. & Privalov, P. L. (1999). The energetics of HMG box interactions with DNA: thermodynamic description of the target DNA duplexes. *J. Mol. Biol.* **294**, 981–995.
99. Stern, J. C., Anderson, B. J., Owens, T. J. & Schildbach, J. F. (2004). Energetics of the sequence-specific binding of single-stranded DNA by the F factor relaxase domain. *J. Biol. Chem.* **279**, 29155–29159.
100. Spolar, R. S. & Record, M. T. Jr (1994). Coupling of local folding to site-specific binding of proteins to DNA. *Science,* **263**, 777–784.
101. Berger, C., Jelesarov, I. & Bosshard, H. R. (1996). Coupled folding and site-specific binding of the GCN4-bZIP transcription factor to the AP-1 and ATF/CREB DNA sites studied by microcalorimetry. *Biochemistry,* **35**, 14984–14991.
102. Lundback, T., Chang, J. F., Phillips, K., Luisi, B. & Ladbury, J. E. (2000). Characterization of sequence-specific DNA binding by the transcription factor Oct-1. *Biochemistry,* **39**, 7570–7579.
103. Ferrari, M. E. & Lohman, T. M. (1994). Apparent heat capacity change accompanying a nonspecific protein-DNA interaction. *Escherichia coli* SSB tetramer binding to oligodeoxyadenylates. *Biochemistry,* **33**, 12896–12910.

104. Minetti, C. A., Remeta, D. P., Zharkov, D. O., Plum, G. E., Johnson, F., Grollman, A. P. & Breslauer, K. J. (2003). Energetics of lesion recognition by a DNA repair protein: thermodynamic characterization of formamidopyrimidine-glycosylase (Fpg) interactions with damaged DNA duplexes. *J. Mol. Biol.* **328**, 1047–1060.
105. Froelich-Ammon, S. J., Dickinson, B. A., Bevilacqua, J. M., Schultz, S. C. & Cech, T. R. (1998). Modulation of telomerase activity by telomere DNA-binding proteins in Oxytricha. *Genes Dev.* **12**, 1504–1514.
106. Lei, M., Zaug, A. J., Podell, E. R. & Cech, T. R. (2005). Switching human telomerase on and off with hPOT1 protein in vitro. *J. Biol. Chem.* **280**, 20449–20456.
107. Hicke, B., Rempel, R., Maller, J., Swank, R. A., Hamaguchi, J. R., Bradbury, E. M., Prescott, D. M. & Cech, T. R. (1995). Phosphorylation of the Oxytricha telomere protein: possible cell cycle regulation. *Nucl. Acids Res.* **23**, 1887–1893.
108. Zahler, A. M. & Prescott, D. M. (1988). Telomere terminal transferase activity in the hypotrichous ciliate Oxytricha nova and a model for replication of the ends of linear DNA molecules. *Nucl. Acids Res.* **16**, 6953–6972.
109. Jacob, N. K., Skopp, R. & Price, C. M. (2001). G-overhang dynamics at Tetrahymena telomeres. *EMBO J.* **20**, 4299–4308.
110. Kornberg, A. & Baker, T. (1992). *DNA Replication*, 2nd edit. Freeman, New York.
111. Karpel, R. L. (1990). T4 Bacteriophage gene 32 protein. In *The Biology of Nonspecific DNA-Protein Interactions* (Revzin, A, ed), pp. 103–126, CRC, Boca Raton, FL.
112. Bochkarev, A., Pfuetzner, R. A., Edwards, A. M. & Frappier, L. (1997). Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA. *Nature,* **385**, 176–181.
113. Raghunathan, S., Kozlov, A. G., Lohman, T. M. & Waksman, G. (2000). Structure of the DNA binding domain of *E. coli* SSB bound to ssDNA. *Nature Struct. Biol.* **7**, 648–652.
114. Gill, S. C. & von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319–326.
115. Cantor, C. R., Warshaw, M. M. & Shapiro, H. (1970). Oligonucleotide interactions: III. Circular dichroism studies of the conformation of deoxyoligonucleotides. *Biopolymers,* **9**, 1059–1077.
116. Record, M. T. Jr, Woodbury, C. P. & Lohman, T. M. (1976). Na^+ effects on transition of DNA and polynucleotides of variable linear charge density. *Biopolymers,* **15**, 893–915.
117. deHaseth, P. L., Lohman, T. M. & Record, M. T. Jr (1977). Nonspecific interaction of lac repressor with DNA: an association reaction driven by counterion release. *Biochemistry,* **16**, 4783–4790.
118. Lohman, T. M., deHaseth, P. L. & Record, M. T. Jr (1980). Pentalysine-deoxyribonucleic acid interactions: a model for the general effects of ion concentrations on the interactions of proteins with nucleic acids. *Biochemistry,* **19**, 3522–3530.
119. Westhof, E. (1987). Re-refinement of the B-dodecamer d(CGCGAATTCGCG) with a comparative analysis of the solvent in it and in the Z-hexamer d (5BrCG5BrCG5BrCG). *J. Biomol. Struct. Dynam.* **5**, 581–600.
120. Fratini, A. V., Kopka, M. L., Drew, H. R. & Dickerson, R. E. (1982). Reversible bending and helix geometry in a B-DNA dodecamer: CGCGAATTBrCGCG. *J. Biol. Chem.* **257**, 14686–14707.
121. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A,* **47**, 110–119.
122. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P. Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography, NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D,* **54**, 905–921.
123. Parkinson, G., Vojtechovsky, J., Clowney, L., Brunger, A. T. & Berman, H. M. (1996). New parameters for the refinement of nucleic acid-containing structures. *Acta Crystallog. sect. D,* **52**, 57–64.
124. Hicke, B. J., Willis, M. C., Koch, T. H. & Cech, T. R. (1994). Telomeric protein-DNA point contacts identified by photo-cross-linking using 5-bromodeoxyuridine. *Biochemistry,* **33**, 3364–3373.

Edited by J. E. Ladbury

(Received 9 November 2005; received in revised form 6 February 2006; accepted 17 February 2006)
Available online 25 April 2006