The EMBO Journal Vol.18 No.20 pp.5735–5744, 1999

TRF1 binds a bipartite telomeric site with extreme spatial flexibility

Alessandro Bianchi[1,2], Rachel M.Stansel[3], Louise Fairall[4], Jack D.Griffith[3], Daniela Rhodes[4] and Titia de Lange[1,5]

[1]The Rockefeller University, 1230 York Avenue, New York, NY 10021, [3]Lineberger Comprehensive Cancer Center, UNC, Chapel Hill, NC 27599-7295, USA and [4]MRC Laboratory Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

[2]Present address: Department of Molecular Biology, University of Geneva, 30 quai Ernest-Ansermet, CH-1211 Geneva 4, Switzerland

[5]Corresponding author
e-mail: delange@rockvax.rockefeller.edu

TRF1 is a key player in telomere length regulation. Because length control was proposed to depend on the architecture of telomeres, we studied how TRF1 binds telomeric TTAGGG repeat DNA and alters its conformation. Although the single Myb-type helix–turn–helix motif of a TRF1 monomer can interact with telomeric DNA, TRF1 predominantly binds as a homodimer. Systematic Evolution of Ligands by Exponential enrichment (SELEX) with dimeric TRF1 revealed a bipartite telomeric recognition site with extreme spatial variability. Optimal sites have two copies of a 5′-YTAGGGTTR-3′ half-site positioned without constraint on distance or orientation. Analysis of binding affinities and DNase I footprinting showed that both half-sites are simultaneously contacted by the TRF1 dimer, and electron microscopy revealed looping of the intervening DNA. We propose that a flexible segment in TRF1 allows the two Myb domains of the homodimer to interact independently with variably positioned half-sites. This unusual DNA binding mode is directly relevant to the proposed architectural role of TRF1.

Keywords: dimerization/DNA-binding/electron microscopy/telomeres/TRF1

Introduction

TRF1 is a duplex telomeric DNA binding factor present at mammalian chromosome ends in interphase and mitosis (Chong *et al*., 1995), where it controls the steady-state length of the telomeric repeat tract (van Steensel and de Lange, 1997). TRF1 regulates telomere length maintenance without affecting the expression of telomerase, suggesting that TRF1, like the yeast factor Rap1p (Marcand *et al*., 1997), acts *in cis*, possibly by inhibiting the interaction of telomerase with telomere termini. It is not known how TRF1 controls events at the telomere terminus. One possibility is that TRF1 induces an architectural change that limits the access of telomerase. Indeed, the ends of telomeres can be sequestered in large duplex loops (t-loops) (Griffith *et al*., 1999) that could represent such an inaccessible state. Given these considerations, it is important to understand how TRF1 binds to telomeric DNA and affects the overall conformation of telomeric tracts.

TRF1 was first identified as a HeLa-cell-derived DNA binding activity capable of forming a specific complex with tandem arrays of duplex TTAGGG repeats (Zhong *et al.*, 1992). Neither single-stranded telomeric DNA nor heterologous telomeric repeats are effective TRF1 substrates (Zhong *et al.*, 1992; Hanish *et al.*, 1994; Chong *et al.*, 1995). TRF1 is a stable dimer, and dimerization is required for efficient DNA binding (Bianchi *et al.*, 1997). Comparison of the mammalian TRF1 proteins and their homolog TRF2 has identified two well conserved domains (Bilaud *et al.*, 1997; Broccoli *et al.*, 1997a,b). The first is a C-terminal 50 amino acid domain related to the DNA binding region of the Myb proto-oncogenes. The second is an ~200 amino acid TRF-specific domain that contains sequences responsible for homodimerization (Bianchi *et al*., 1997; Broccoli *et al*., 1997b; H.Moss and T.de Lange, unpublished results). The region between the Myb and dimerization domains is poorly conserved. TRF1 and TRF2 also differ in their N-termini, where TRF1 is acidic and TRF2 basic.

The Myb-type DNA binding domain is a helix–turn–helix (HTH) motif usually present in multiple tandem copies in the large class of Myb-related proteins that function as transcriptional regulators of cell growth and differentiation (for review see Lipsick, 1996; Martin and Paz-Ares, 1997). Interestingly, the Myb family of proteins appears to be of polyphyletic origin, implying modular evolution (Rosinski and Atchley, 1998). Besides occurring in transcription factors, Myb repeats have been identified in telomere binding factors. In fact, all duplex telomeric DNA binding proteins identified to date (Rap1p, Taz1p, TRF1 and TRF2) contain at least one Myb repeat (Larson *et al.*, 1994; Chong *et al.*, 1995; Bilaud *et al.*, 1996, 1997; Konig *et al.*, 1996; Broccoli *et al.*, 1997b; Cooper *et al.*, 1997), with the possible exception of the yeast protein Tel2p (Kota and Runge, 1998). The Myb domains of TRF1 and TRF2 are most closely related to that of the *Schizosaccharomyces pombe* telomeric protein Taz1p (Bilaud *et al*., 1997; Cooper *et al*., 1997).

The number of Myb repeats found in Myb-related proteins varies. Three repeats are found in c-Myb, the c-Myb homologs A- and B-Myb, *Drosophila* Myb and *Dictyostelium* Myb, as well as in additional factors (for review see Lipsick, 1996), whereas two repeats are often, but not exclusively, found in plant proteins (for review see Martin and Paz-Ares, 1997). Although several cases are known of proteins containing one repeat (Baranowskij *et al.*, 1994; Kirik and Baumlein, 1996; Feldbrugge *et al.*, 1997), to our knowledge the binding mechanism for these single-repeat factors has not been extensively characterized. Interestingly, the *Drosophila* transcription factor

Adf-1 contains a single Myb repeat within its DNA binding domain and has been shown to form a homodimer that binds DNA with higher affinity than the monomer (Cutler *et al.*, 1998), raising the possibility that homodimerization may be a general feature of single Myb repeat proteins. The binding mechanism of two Myb-related proteins, c-Myb and Rap1p, has been revealed in great detail by structural analysis of the protein–DNA complex (Ogata *et al*., 1994; Konig *et al*., 1996).

The DNA binding domain of c-Myb is composed of three Myb repeats, R1–3, but only R2 and R3 are required for binding (Tanikawa *et al.*, 1993). NMR spectroscopy of R2 and R3 in complex with the c-Myb recognition site 5′-AACTG-3′ revealed that both Myb repeats form an HTH motif in which the third helix is a recognition helix (Ogata *et al*., 1994). These two recognition helices pack tightly in the major groove in a head-to-tail orientation and interact cooperatively with a unique site on the DNA. Although the DNA binding domain of the yeast telomeric protein Rap1p contains two Myb-type repeats (Larson *et al.*, 1994), their tertiary arrangement, as revealed by crystallography (Konig *et al.*, 1996), is entirely different from the one in c-Myb. In Rap1p, the two HTH motifs bind separately to two tandemly repeated sequence elements (5′-GGGTGT-3′ and 5′-GGTGT-3′), which are separated by 8 bp. Like the homeodomain, the HTH motifs of Rap1p have an N-terminal arm that makes contacts in the minor groove (Gehring *et al.*, 1994; Konig *et al.*, 1996; Konig and Rhodes, 1997). This arrangement allows each of the Myb motifs to recognize a separate 5–6 bp site, whereas in c-Myb, two HTH motifs are needed to specify contacts to a single pentameric site.

The solution structure of the TRF1 Myb domain (Nishikawa *et al.*, 1998) showed that its recognition helix is longer than those of the c-Myb repeats, making it unlikely that two TRF1 recognition helices could pack together in the major groove, as is the case with c-Myb. Alternatively, the TRF1 dimer may recognize separate sites similar to Rap1p. In agreement, the amino acid sequence of the TRF1 Myb motif suggests the presence of a homeodomain-like N-terminal arm, as is found in Rap1p (Konig *et al*., 1998). Furthermore, unlike R2 and R3 of c-Myb, a single TRF1 Myb domain can bind to DNA by itself, specifically interacting with the sequence 5′-GGGTTA-3′ (Konig *et al.*, 1998). However, it is not known whether the two Myb domains in the TRF1 dimer similarly bind to independent sites. In fact, efficient DNA binding *in vitro* by the dimeric protein appears to require the presence of two Myb domains in the dimer (Bianchi *et al.*, 1997), and so does the telomeric localization of the protein *in vivo* (van Steensel and de Lange, 1997; A.Bianchi, H.Moss and T.de Lange, unpublished results).

We have undertaken a biochemical characterization of TRF1 and have determined the DNA binding site of the dimeric protein. We show that TRF1 binds to a bipartite ‘full’ site made of two identical ‘half’ sites, each of which is contacted independently but simultaneously by one Myb domain. Unexpectedly, the relative spacing and orientation of the two half-sites on the DNA did not affect TRF1 binding, and TRF1 could contact widely spaced half-sites by looping of the intervening sequence. The results suggest a novel binding mechanism that involves free rotation of the two Myb domains in the TRF1 dimer, possibly facilitated by a flexible connection between the dimerization domain and the Myb domain in each monomer. This conformational flexibility explains how TRF1 binding can pair and loop telomeric tracts, features proposed to be crucial for the architecture of telomeres *in vivo*.

Results

Identification of the TRF1 binding site

To determine the sequence and configuration of the TRF1 binding site, we undertook a polymerase chain reaction (PCR)-based approach, Systematic Evolution of Ligands by Exponential enrichment (SELEX). Given the tandemly repeated nature of human telomeric DNA and the binding requirement for two identical Myb repeats in a TRF1 dimer, it seemed likely that the two Myb motifs would each make independent contacts with two identical DNA sites (two ‘half’ sites, which together would constitute one ‘full’ TRF1 site). In preparation for the SELEX procedure, a series of DNA substrates was constructed (Figure 1A), each bearing two copies of the sequence $(TTAGGG)_2$, which contains the 5′-GGGTTA-3′ recognition site for the isolated TRF1 Myb domain defined by Rhodes and colleagues (Konig *et al*., 1998). The presence of two copies of the $(TTAGGG)_2$ cassette improves the affinity of TRF1 significantly (Figure 1A, and see below). The spacing between the two telomeric cassettes was then varied from 6 to 30 bp while maintaining the 6 bp phasing of vertebrate telomeric tracts. Requirement for a spacing larger than 30 bp could be ruled out based on previous band-shift and electron microscopy (EM) binding experiments (Zhong *et al.*, 1992; Bianchi *et al.*, 1997; Griffith *et al.*, 1998).

Using the DNA probes with variable distances between two $(TTAGGG)_2$ cassettes, no significant change in TRF1 binding affinity was observed in band-shift experiments (Figure 1A, lanes 7–30). This binding behavior could be explained if TRF1 were able to contact two half-sites independently of their distance, or if the full-site were contained within one single $(TTAGGG)_2$ cassette, which is unlikely given the poor binding to DNA with one cassette. In any case, if TRF1 requires a fixed spacing between these cassettes, the results suggested that this distance is represented in a sequence segment of 36 bp.

Based on these findings, we designed a SELEX library of oligonucleotides containing a central randomized region of 38 bp (see Materials and methods). Molecules bound to the TRF1 dimer were selected from the library on a native acrylamide gel, amplified by PCR and subjected again to TRF1 binding. After seven rounds of selection, the PCR products were cloned and their sequences determined.

A total of 43 clones were analyzed and each was found to contain stretches of perfect homology to the vertebrate telomeric sequence (Figure 1B). More specifically, at least one exact copy of the sequence 5′-AGGGTT-3′ was present in all instances. The selected sequences could be divided into two classes: one class containing two adjacent copies of the hexamer (Figure 1B, top panel); and one class containing two copies separated by a linker of variable length (Figure 1B, bottom panel). Only in two instances was a single copy of the hexamer present (Figure 1B, bottom panel, last two lines). Alignment of all the sites in each class resulted in the derivation of

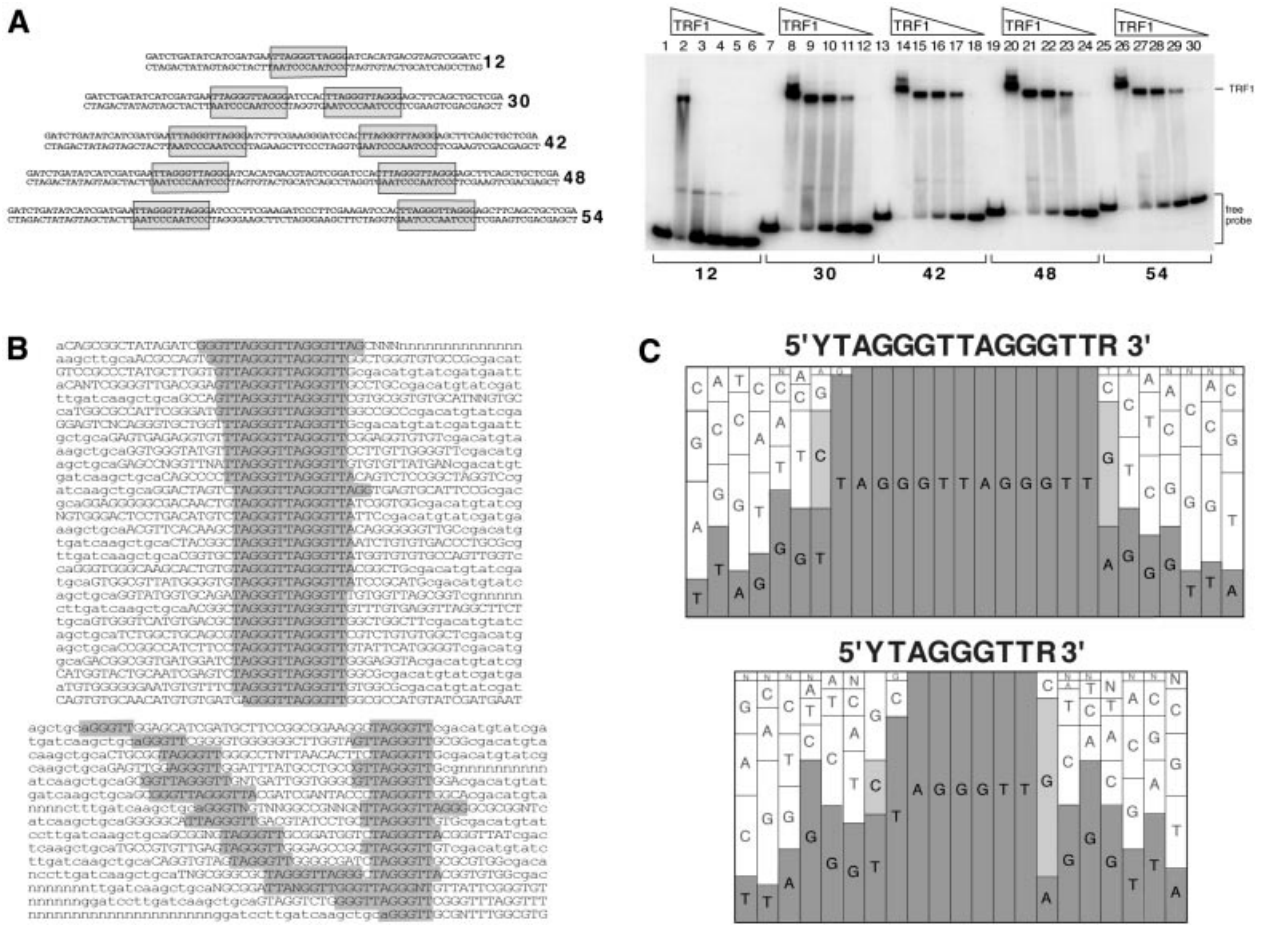

Fig. 1. Determination of the TRF1 DNA binding site by SELEX. (**A**) TRF1 band-shift assays on DNA probes (sequence shown to the left) containing one or two copies of the sequence $(TTAGGGG)_2$ (highlighted). The numbers of the probes refer to the number of bp spanning telomeric sequences and their intervening DNA. The position of the unbound DNA and the TRF1–DNA complex are indicated to the right of the gel. In all band-shift experiments with baculovirus-expressed TRF1, the higher-mobility complexes observed at the high protein concentrations are probably the result of protein–protein interactions. Protein concentrations decreased in 3-fold steps from left to right for each probe, from ~0.9×10^{-7} M (lanes 2, 8, 14, 20 and 26). (**B**) DNA sequences from the randomized oligonucleotide library recovered after seven rounds of selection with TRF1. Upper-case letters indicate sequences derived from the randomized portion of the oligonucleotide; lower-case letters refer to flanking sequences. Stretches of >5 bp identity to telomeric DNA are highlighted. The top panel lists adjacent copies of the 5′-AGGGTT-3′ motif. The bottom panel lists non-adjacent or single-copy AGGGTT sites. (**C**) Sequences from both classes were aligned and a consensus derived. In-phase identities to telomeric DNA are highlighted in dark gray, whereas lighter gray indicates bases included in the consensus site.

consensus TRF1 binding sites of very similar nature, with the first consensus (Figure 1C, top panel) simply representing two partially overlapping copies of the second (Figure 1C, bottom panel). Confirming and extending previous results (Zhong *et al.*, 1992; Hanish *et al.*, 1994), it was immediately apparent that the sequence requirements for TRF1 binding are very strict, as essentially no base changes were allowed within the core sequence, 5′-AGGGTT-3′. Almost equally strong were the requirements for a pyrimidine and a thymidine preceding the core sequence and for a purine following it.

Unexpectedly, not only were all the selected double sites present in the same orientation, but all the telomeric sequences showed the same strand polarity with respect to the orientation in the oligonucleotide library. In other words, the G-rich strand of the TRF1 site always occurred in the same strand, representing the chemically synthesized oligonucleotides of the original library. This bias is probably due to a skewed base composition in the chemically synthesized oligonucleotide, which displayed an excess of G residues (frequency of 0.45 instead of 0.25, see Materials and methods). Identical results were obtained with a second oligonucleotide library that had a different sequence on the 5′ side of the randomized sequence (data not shown), and attempts to obtain a less biased library from other suppliers were unsuccessful.

TRF1 binds to DNA by engaging both Myb domains

The SELEX data identify the 5′-YTAGGGTTR-3′ site as the sequence recognized by a single Myb domain of TRF1, with 5′-AGGGTT-3′ representing the invariable core of the half-site (see below). The frequent occurrence of two copies of the 5′-AGGGTT-3′ sequence was in agreement with a model for TRF1 binding in which each Myb motif contacts one site independently, in a manner similar to the yeast telomeric protein Rap1p (Konig *et al.*, 1996). However, the results of the SELEX in themselves did not establish whether the two Myb domains are engaged on DNA simultaneously. In fact, SELEX often yields repeated sites because multiple binding sites increase the apparent binding affinity (see Materials and

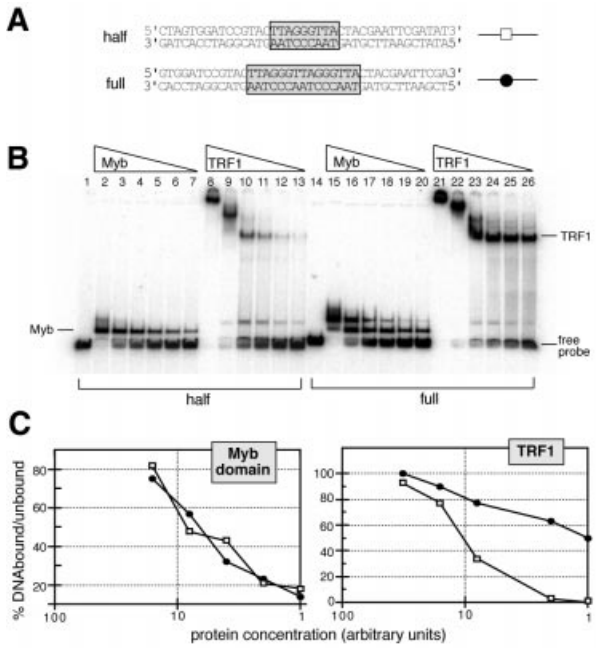

Fig. 2. Band-shift analysis of binding of TRF1 and the Myb domain to DNA fragments containing one or two copies of the half-site. (**A**) DNA probes used in binding assays. DNAs 'half' and 'full' were restriction fragments of 111 and 117 bp, respectively, containing the telomeric sequences indicated. (**B**) Band-shift assay for TRF1 and the TRF1 Myb domain binding to the 'half' and 'full' DNA probes. Assays were performed under standard TRF1 conditions with addition of 0.1% NP-40 and poly(dI·dC) (100 ng). Gel was run in 0.5× TBE at 4°C. Protein concentrations varied in 2-fold serial dilutions starting from ~4×10^{-7} M for TRF1 and ~1.0×10^{-7} M for the Myb domain. (**C**) Phosphorimager quantitation of reactions shown in (A). All complexes were scored as protein-bound DNA (both complexes in the case of the Myb domain).

methods) (Cantor and Schimmel, 1980; Wilson *et al.*, 1993). This phenomenon is well documented when selections have been performed with long oligonucleotides (Baranowskij *et al.*, 1994; Hsu *et al.*, 1994; Wotton *et al.*, 1994).

If TRF1 binds to DNA by engaging only one Myb domain, both the dimeric protein and the isolated Myb domain are expected to show a 2-fold increase in affinity with DNA containing two half-sites compared with a DNA bearing a single site. On the other hand, if both Myb domains in a TRF1 dimer are engaged on DNA simultaneously, a cooperative effect is to be expected, with an increase >2-fold for binding to DNAs with two half-sites. This cooperativity should not occur with the isolated Myb domain.

Accordingly, we compared the relative affinities of dimeric TRF1 and the isolated Myb domain for DNAs containing either one or two copies of the half-site (Figure 2). The isolated Myb domain showed nearly the same affinity for both substrates (Figure 2B, lanes 2–7 and 15–20; and Figure 2C, left panel), or at most a 2-fold higher affinity for the substrate with two half-sites (data not shown). As expected, the probe with two half-sites bound two Myb domains, confirming that each half-site represents a binding site for a single Myb domain. In contrast, binding of one TRF1 dimer to the two types of substrates occurred with a difference in affinity of ~10-fold (Figure 2B, lanes 8–13 and 21–26; and Figure 2C, right panel). This cooperative effect was not attributable to simple multimerization of the minimal binding site (see Materials and methods for theoretical consideration), supporting the idea that the TRF1 dimer can bind to DNA by engaging both Myb domains simultaneously. In agreement with cooperativity attributed to the binding of two Myb domains, the phenomenon occurred only when the number of half-sites was increased from one to two but not when the number of half-sites was increased from two to five or eight (data not shown). Furthermore, while TRF1 complexes are readily observed during gel electrophoresis at room temperature, Myb domain complexes appear unstable under these conditions and are only detected at 4°C (data not shown). This difference is consistent with both Myb domains in the TRF1 dimer conferring greater thermal stability by binding simultaneously.

No effect of half-site orientation or spacing on TRF1 binding

The sequences recovered from SELEX did not reveal a bias for a specific relative arrangement of the two half-sites (Figure 1B). When the base frequencies in the SELEX starting material were taken into account, the frequencies of different arrangements that were recovered suggested that DNAs with directly juxtaposed half-sites were not a better substrate than half-sites spaced at a distance (for calculation see Materials and methods). This was confirmed by a direct analysis of the TRF1 binding affinities for two representatives from each class of sites (Figure 3A). The SELEX products also suggested that many different half-site distances are tolerated, including those lacking the 6 bp phasing typical of natural telomeric DNA. The only notable bias in the SELEX products was the absence of combinations of half-sites in opposite orientation. Because this bias could be explained from the relative G-rich nature of the 'top' strand of the input library, we tested the effect of site orientation on TRF1 binding directly. In quantitative band-shift assays with two DNA molecules bearing two half-sites in either a direct or inverted orientation, TRF1 did not show a difference in affinity for these two molecules (Figure 3B). Thus, neither the relative orientation nor the spacing between the two TRF1 half-sites significantly affected the affinity of TRF1 for the DNA.

Footprinting analysis of the isolated Myb domain and dimeric TRF1

To test further whether the TRF1 dimer engages both Myb domains on DNA, we performed DNase I footprinting experiments comparing the binding of the isolated Myb domain with that of TRF1. The Myb domain and TRF1 produced essentially the same footprints, both on DNAs containing one or two copies of the half-site (Figure 4A–H). Since the binding of two Myb domains on two half-sites gave the same footprint as one TRF1 dimer (Figure 4E–H; band-shift data not shown), it is likely that both Myb domains in the dimer were bound. The footprint is centered on the sequence TAGGGTTAGGGTT and hence corresponds very well to the sequence of a full binding site identified by SELEX (Figure 1B). In addition, the footprint of one TRF1 molecule bound to two half-sites separated

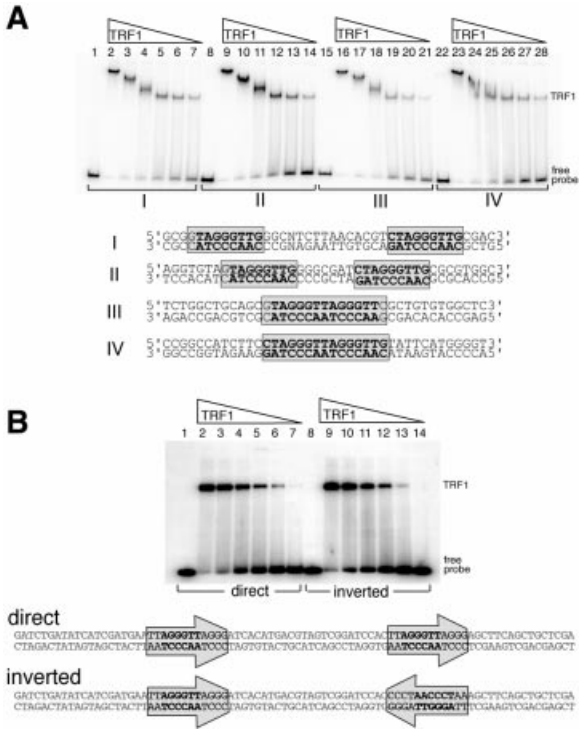

Fig. 3. Effects of spacing and orientation of half-sites on TRF1 binding. (**A**) TRF1 band-shift assays on four representative clones recovered from SELEX, two from each class of sites. DNA probes were 170 bp, with the TRF1 half-sites (boxed) at a central position. Within each box, matches to the consensus are indicated in bold. The protein concentrations were a dilution range with 1.5-fold steps starting at ~1.7×10^{-7} M (lanes 2, 9, 16 and 23). The position of the unbound DNA and the TRF1–DNA complex are indicated at the right of the gel. (**B**) TRF1 band-shift assays on two DNA fragments bearing two half-sites in either direct or inverted orientation. Protein concentrations varied in 2-fold steps beginning at about 1.4×10^{-7} M (lanes 2 and 9). The size of the DNA fragments was 80 bp. Telomeric cassettes are highlighted on gray background, and core sites are in bold.

by 18 bp resulted in a split footprint, in which each half corresponds to the footprint of a Myb domain centered on the half-site TAGGGTTA (Figure 4I–J). This is consistent with the simultaneous interaction of both Myb domains on DNA. Thus, these results further validated the conclusion that the full TRF1 binding site is composed of two equal half-sites, each capable of binding one Myb domain, and that the TRF1 dimer has the capability to bind to a large collection of spatial arrangements of the half-sites.

TRF1-induced DNA looping between two distant half-sites

We reasoned that the unexpected ability of TRF1 to engage two distant half-sites simultaneously could lead to looping of the DNA if the intervening sequence was sufficiently long. To address this possibility, we used electron microscopy to examine TRF1 bound to a 2196 bp *Nae*I–*Sca*I fragment of pBluescript II KS(+), which carries two AGGGTT sites. These sites are spaced 205 bp apart at 249 and 454 bp from the *Nae*I site (Figure 5A). In all the TRF1-bound molecules DNA loops were observed, and TRF1 was always present at the base of the loop (Figure 5A). TRF1 binding was never observed elsewhere in the molecule (data not shown). Consistent with looping being due to TRF1 binding at the two half-sites, most loops were in the 150–250 bp size range (Figure 5B) and their position clustered around the AGGGTT sites (Figure 5C).

Since TRF1 is known to form higher order complexes through homotypic interactions (Griffith *et al.*, 1998), we performed control experiments to determine whether the observed looping was simply due to interactions between two TRF1 dimers, each one bound at one of the half-sites. If this were the case, TRF1 should also form loops between distantly placed full-sites. To test this prediction, three plasmids were constructed based on pBluescript KS(–), which itself bears five AGGGTT sites. To each

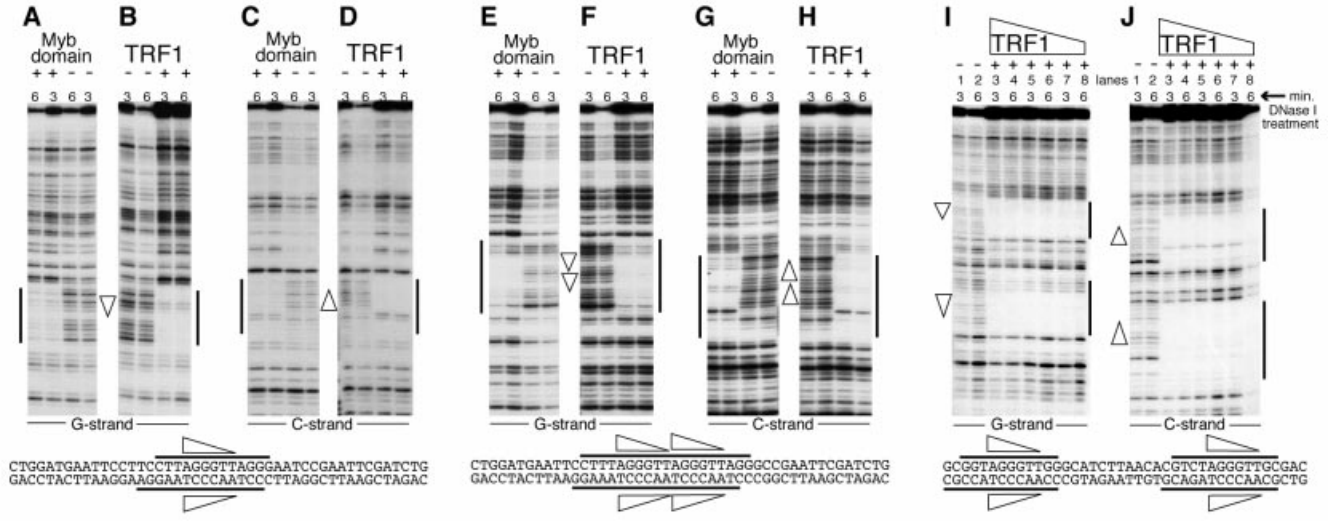

Fig. 4. Footprinting of TRF1 and the isolated Myb domain on DNA fragments containing one or two copies of the half-site. (**A**–**D**) DNase I footprint of the isolated TRF1 Myb domain and dimeric TRF1 bound to a sjngle half-site. (**E**–**H**) DNase I footprint of the Myb domain and TRF1 bound to two adjacent half-sites. (**I** and **J**) DNase I footprint of dimeric TRF1 bound to two non-adjacent half-sites. Protected positions were assessed from G+A Maxam–Gilbert sequencing reactions run in parallel (not shown). Open triangles indicate the 5′-AGGGTT-3′ core sequence. Bars highlight protected regions. The TRF1 concentration was 2×10^{-8} M for (B), (D), (F) and (H), and reactions were performed with 10 μg/ml sonicated calf thymus competitor DNA. For (I) and (J), TRF1 concentrations were 2×10^{-8} M (lanes 3 and 4), 6×10^{-9} M (lanes 5 and 6) and 2×10^{-9} M (lanes 7 and 8), and sonicated *E.coli* DNA was used as a competitor.

Table I. Analysis of DNA loop formation in the presence of full- and/or half-sites in the DNA molecule

Plasmid	Full-sites in plasmid	Half-sites in plasmid	No. of molecules	Molecules with one TRF1 dimer bound	Molecules with two TRF1 dimers bound	Looped molecules
pH96H	0	7	300	18 (6.0%)	0	15 (5.0%)
pH96F	1	6	400	91 (22.7%)	0	4 (1.0%)
pF96F	2	5	250	65 (26.0%)	19 (7.6%)	0

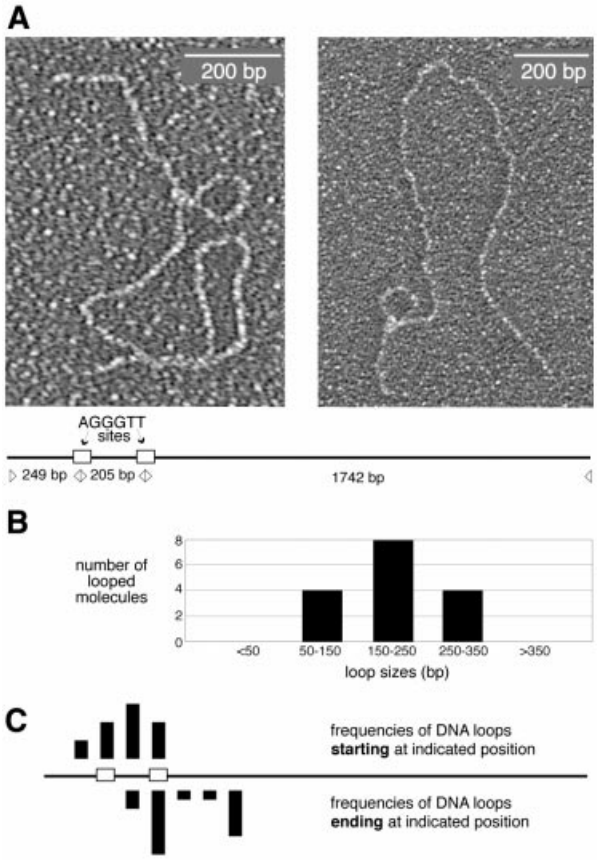

Fig. 5. EM analysis of TRF1-inducing loops between two half-sites separated by 205 bp. (**A**) Electron micrographs of linear DNA molecules containing two AGGGTT sites bearing small DNA loops anchored at their base by a TRF1 protein ball. Schematic of the DNA fragment [*Nae*I–*Sca*I fragment from pBluescript II KS(+)] is shown below the micrographs. (**B**) Histogram of the TRF1-induced loop sizes on the *Nae*I–*Sca*I fragment. (**C**) Histogram of the positions of bound TRF1 molecules in the *Nae*I–*Sca*I fragment.

plasmid, two additional TRF1 binding sites were added, separated by 96 bp of non-telomeric DNA. The added sites were either half-sites (TTAGGGTTA) or full-sites (TTAGGGTTAGGGTTA) and resulted in plasmids containing two added full-sites (pF96F), two half-sites (pH96H) or one of each (pH96F). These DNA molecules were then incubated with TRF1 and analyzed by EM for frequency of TRF1 binding and loop formation (Table I).

Consistent with the data obtained using band-shift analysis, binding improved significantly (4- to 6-fold) when at least one full-site was present in the plasmid in addition to the multiple half-sites. Remarkably, for DNAs not bearing full-sites, 100% (for pBluescript; data not shown) and 83% (for pH96H) of the bound molecules had loops, further implying that high-affinity binding by TRF1 requires simultaneous binding of two half-sites. Due to the presence of multiple half-sites in pH96H, the distribution of loop sizes was broader and flatter with this plasmid (data not shown), with a profile that followed the curve for DNA circularization determined by Shore and Baldwin (Shore *et al.*, 1981), suggesting that a limiting factor for TRF1 binding to this DNA lies in bringing the two half-sites into close proximity in three-dimensional space. Most importantly, no loops were observed with the plasmid bearing two full-sites, even in those cases where two bound TRF1 molecules were clearly visible. This result argues against the possibility that DNA loops are formed only through interactions between TRF1 dimers. Furthermore, the apparent size of the protein ball at the base of the loops was consistent with a single TRF1 dimer rather than a larger protein mass.

In previous EM studies on TRF1–DNA complexes, we have failed to observe looping or shortening of the DNA (Griffith *et al.*, 1998). These experiments were conducted on continuous arrays of TTAGGG repeats, and it is possible that under these conditions the two Myb domains tend to occupy adjacent (or close) sites due to the higher probability of encountering a second site that is closely linked. Similarly, the loop frequencies observed here dropped from 5 to 0–1% when a full-site was present in the DNA, suggesting that binding of adjacent half-sites was favored.

Discussion

The maintenance and protection of mammalian chromosome ends depends on the action of two related TTAGGG-repeat binding factors, TRF1 and TRF2 (see Figure 6A). Merging evidence suggests that TRF1 and TRF2 primarily act as architectural proteins that modify the overall conformation of the telomeric tract (Bianchi *et al.*, 1997; Griffith *et al.*, 1998, 1999). For instance, *in vitro* TRF1 pairs two telomeric tracts in both parallel and anti-parallel orientations, and TRF2 induces the formation of large duplex loops (t-loops) through strand-invasion of the telomere terminus. T-loops may represent one of the main mechanisms by which telomeric DNA is protected and maintained. In this study, we have used SELEX, analysis of binding affinities, DNase I footprinting and EM analysis to establish the DNA binding mode of TRF1. The results indicate that TRF1 binds a bipartite telomeric site composed of two 5′-YTAGGGTTR-3′ half-sites. Remarkably, there is no detectable constraint on the spatial arrangement of the two half-sites (illustrated in Figure 6B). TRF1 binds half-sites in direct and inverted orientations equally well. Furthermore, TRF1 tolerates a large variation in half-site distances and can induce DNA loops in the intervening segment. This extensive flexibility predicts that TRF1 can fold back telomeres *in vivo* and can hold the telomeric tracts in an anti-parallel orientation. Such changes in

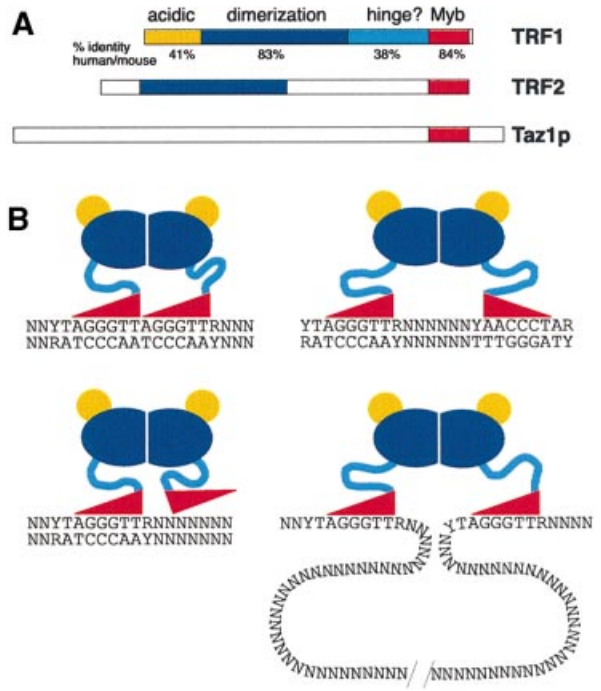

Fig. 6. Model for the flexible binding of TRF1 to telomeric sites. (**A**) Domain structure of TRF1 and comparison with TRF2 and Taz1p. (**B**) Illustration of different ways in which one or two Myb domains in the TRF1 dimer can contact specific half-sites. See text for discussion.

telomere conformation are highly relevant for the formation of t-loops and other aspects of telomere function.

Although the isolated Myb domain of TRF1 can bind telomeric sites (Bilaud *et al*., 1996; Konig *et al*., 1998), full-length TRF1 forms a homodimer and requires both Myb domains for the formation of a stable complex *in vitro* and *in vivo* (Bianchi *et al*., 1997; van Steensel and de Lange, 1997). In agreement with a dimeric DNA binding mode, the current results reveal that the two Myb domains of full-length TRF1 each contact a 5′-YTAGGGTTR-3′ half-site independently and that both Myb domains are bound to DNA simultaneously (schematized in Figure 6B). A cooperative effect is observed when TRF1 binds to two half-sites as opposed to one. The effect is specific for TRF1 and does not occur with the isolated Myb domain. Although the cooperative effect is relatively small (10-fold), similar values have been reported for other proteins (Wilson *et al.*, 1993; Payre *et al.*, 1997; Yie *et al.*, 1997). Stronger cooperativity, examples of which are reported for several HTH proteins and homeodomains (Wilson *et al.*, 1993; Smith and Sauer, 1995), usually involves specific protein–protein interactions. In the case of TRF1, we propose that the cooperative effect is simply created by increased proximity of the second Myb after binding of the first, and not by specific interactions between the two DNA binding domains.

The spatial arrangement of the two TRF1 monomers in the dimer remains to be determined. If TRF1, like many other homodimeric proteins, has rotational symmetry (as depicted in Figure 6B), one of the Myb domains would have to swivel 180° to contact the second half-site in tandemly arranged telomeric repeats. The region of TRF1 that spans amino acids 265–376 seems a good candidate for a structural hinge between the dimerization and DNA binding domains. This part of TRF1 is poorly conserved (only 38% identity in human and mouse TRF1; Broccoli *et al*., 1997a; see Figure 6A) and has three protease-sensitive sites (L.Fairall and D.Rhodes, unpublished data). Assuming a length of ~3 Å for an unstructured amino acid, this putative linker domain of ~110 amino acids could extend over 33 nm, sufficient to span the most distantly spaced sites tested by band-shift (56 bp). Several other examples have been reported of proteins that can change the relative orientation of DNA binding domains with a flexible linker [e.g. p53 (Arrowsmith and Morin, 1996), Oct-1 and other POU domain proteins (van Leeuwen *et al.*, 1997) and HAP1 (Zhang and Guarente, 1996; King *et al*., 1999)].

The structural flexibility of TRF1 and its ability to mediate long-range DNA interactions is likely to have far-reaching consequences for telomere structure *in vivo*. The binding of TRF1 to telomeres *in vivo* could occur in several alternative modes. First, TRF1 has the ability to form a filament of protein dimers bound along duplex telomeric tracts. Although there is no detectable cooperative interaction between TRF1 dimers binding to a telomeric repeat array [Hill coefficient of 0.75 (Bianchi, 1999)], they can become tightly packed on telomeric repeats, as visualized by EM analysis (Griffith *et al*., 1998). In a second binding mode, TRF1 induced the formation of DNA loops that are held together by a TRF1 dimer at their base. DNA looping probably requires a minimal distance between the TRF1 half-sites, but such a requirement is easily met at human telomeres. Presumably, binding of the second Myb domain at a nearby half-site is favored over a spatially distant site, and the formation of loops *in vivo* may therefore depend on additional factors that can affect higher-order DNA structure (e.g. nucleosomes). The ability of TRF1 to induce a shallow bend in telomeric DNA (Bianchi *et al*., 1997) could further facilitate telomere folding.

In its third binding mode, TRF1 mediates side-by-side association of unlinked telomeric tracts. The possibility of TRF1-dependent telomere pairing was previously inferred from EM analysis in which parallel and anti-parallel associations were observed between TRF1-covered telomeric tracts (Griffith *et al*., 1998). The current results indicate that such synaptic structures might be facilitated by the extensive flexibility of TRF1, which would allow two Myb domains to engage half-sites in separate telomeric tracts.

These three DNA binding modes of TRF1 (formation of a tightly packed protein–DNA filament, DNA looping and synapse formation) are relevant to the generation of t-loops. Acquisition of TRF1 molecules along a telomeric tract may result in both the formation of a looped structure and the maintenance of such a folded conformation by multiple dimers engaging in a pairing reaction. The relatively high off-rate of TRF1 in *vitro* could facilitate dynamic rearrangements within the complex, while the high site density at telomeres might ‘trap’ TRF1 locally, as we have observed in Surface Plasmon Resonance experiments (A.Bianchi, J.D.Griffith and T.de Lange, unpublished data). Strand invasion by the 3′ telomeric overhang, as facilitated by TRF2, could eventually stabilize the complex in its final t-loop form. According to this view, the flexible hinge region of TRF1 is a crucial

(although non-conserved) region of the protein, in that it allows a multiplicity of DNA binding modes and thus facilitates architectural changes in the TTAGGG repeat arrays. This is a testable prediction of this model.

In addition to revealing architectural features of TRF1, the SELEX data underscore the very strong sequence preference of this protein. The requirement for the core 5′-AGGGTT-3′ in the half-sites appears nearly absolute and there are additional strong base preferences in three adjacent positions. Since TRF1 needs to contact two half-sites, the full TRF1 binding site is at least 12 and possibly 15–18 bp long, a fact that should ensure great specificity of binding *in vivo*. These findings are in agreement with previous results on the sequence requirement for TRF1 binding (Zhong *et al*., 1992; Hanish *et al*., 1994) and are relevant to the specificity of TRF1 for telomeric DNA *in vivo*. We note that the core of the TRF1 half-site, 5′-AGGGTT-3′, occurs in the genes encoding the mammalian telomerase RNAs (Blasco *et al*., 1995; Feng *et al*., 1995), where it represents part of the template and alignment region. Although TRF1 can specifically footprint the human and mouse telomerase RNA template areas (Bianchi, 1999), this interaction is unlikely to be relevant in the absence of additional stabilizing factors, and regulation of the expression of human telomerase RNA by TRF1 has not been observed (A.Bianchi, J.Karlseder and T.de Lange, unpublished data).

The available evidence and similar domain organization (Figure 6A) suggest that TRF2 may bind DNA in a manner similar to TRF1 (Broccoli *et al*., 1997b; Bianchi, 1999). The binding mode of TRF1 may also be a model for telomeric proteins in other organisms. As in the TRF proteins, a C-terminally located single Myb repeat is found in the telomeric protein Taz1p from *S.pombe* (Cooper *et al*., 1997). The mechanism of binding of Taz1p to DNA remains unknown. However, the short DNA sequences used in the one-hybrid screen employed to identify the protein all contained at least two tandem copies of the sequence 5′-GGTTAC-3′, which is similar to the most frequent repeat found at fission yeast telomeres (5′-GGTTACA-3′) (Hiraoka *et al*., 1998). It is possible, given the similar domain organization between the TRFs and Taz1p, that these proteins bind telomeric DNA in a similar manner and all have the ability to facilitate changes in the architecture of the telomeric complex *in vivo*.

Materials and methods

Band-shift assays

TRF1 band-shift assays were as described (Zhong *et al*., 1992) with purified baculovirus-expressed His_6-tagged protein (Bianchi *et al*., 1997). Binding reactions [in 5% glycerol, 4% Ficoll, 20 mM HEPES–KOH pH 7.9, 150 mM KCl, 1 mM $MgCl_2$, 0.1 mM EDTA, 5 ng/ml β-casein, 0.5 mM dithiothreitol (DTT) and 0.1% NP-40] were at room temperature (rt) for 30 min and run on a 5% polyacrylamide gel (29:1) in 1× TBE at rt, or 0.5× TBE at 4°C, at 130–140 V for 2–3 h. Dried gels were analyzed by autoradiography or on a phosphorimager. Myb domain binding reactions were as above, with *Escherichia coli*-derived protein ($TRF1_{371–439}$; Konig *et al*., 1998), and gels were run at 4°C. DNA probes were gel-purified restriction fragments labeled by fill-in with Klenow fragment. About 0.5–1 pmol of DNA were labeled in a 20 µl reaction and the final probe concentration in binding assays was <100 pM. TRF1 concentrations (approximations based on Bradford assays) are given in the figure legends.

SELEX of TRF1 binding sites

SELEX was performed on the oligonucleotide library 5′-ATCGGATC-CTTGATCAAGCTGCA$(N)_{38}$CGACATGTATCGATGAATTCGAC-3′ (1Sel38) (Genelink), rendered double-stranded by annealing to the 20mer 5′-GTCGAATTCATC GATACATG-3′ (*Eco*I) and extension with Klenow fragment. Three-hundred-and-eighty pmol of 1Sel38 and 650 pmol of *Eco*I (annealed in 20 µl of TE/100 mM NaCl at 94°C for 5 min and at 44°C for 30 min) were extended in 80 µl of 50 mM Tris pH 7.5, 10 mM $MgCl_2$, 1 mM DTT, 50 µg bovine serum albumin (BSA) per ml, 2.5 mM each dNTP and 0.25 U Klenow per ml at 30°C for 20 h. In the first TRF1 binding reaction, 1.2 µg of phenol-extracted and ethanol-precipitated double-stranded product were used (at TRF1 concentrations of ~1000, 200 and 40 nM). Binding was as above except that ΦX174 *Hae*III-digested DNA was substituted as DNA competitor. TRF1-bound DNA was isolated from a gel slice (~8 × 5 × 1 mm), crushed in 400 µl of 500 mM NH_4Cl–0.1% SDS, incubated at 90°C for 5 min and, rotating, at rt overnight. DNA was phenol-extracted, ethanol-precipitated and resuspended in 30 µl of ddH_2O. PCR reactions using primers *Eco*I (above) and *Bam*I (5′-ATCGGA TCCTTGATCAAGCT-3′) were performed with 10 µl of the isolated DNA in 50 µl of 10 mM Tris pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% (w/v) gelatin, 0.2 mM each dNTP, 2.4 pmol of each primer per ml and 0.1 U Taq polymerase per ml. Twenty cycles were performed at 94°C (30 s), 44°C (30 s) and 72°C (30 s). One volume of fresh reaction mixture (excluding DNA) was added and one more cycle performed at 94°C (2 min), 44°C (1 min) and 72°C (10 min). Samples were phenol-extracted, ethanol-precipitated and resuspended in 5 µl of TE/100 mM NaCl. Two microliters were used for the subsequent round of selection. After seven rounds, PCR products were *Eco*RI- and *Bam*HI-digested, cloned in pBluescript KS and sequenced. The base composition in the non-selected library was determined by performing four PCR cycles as above (omitting TRF1 binding). The base frequencies, from a total of five clones for 190 nucleotides, were as follows: G = 0.450, A = 0.138, T = 0.222 and C = 0.185. Based on these numbers, the expected frequency of the 5′-YTAGGGTTAGGGTTR-3′ site in the library is ~5.2×10^{-7} and of two copies of the 5′-YTAGGGTTR-3′ site is ~2.6×10^{-7}. Therefore, the two classes of sites (adjacent and non-adjacent half-sites), if bound with equal affinity by TRF1, are expected to be recovered in a 2:1 ratio, which matches the experimental result (Figure 1B).

Relationship between microscopic and macroscopic dissociation constants

For a protein with one binding site (i.e. the TRF1 Myb domain) binding to a DNA containing one copy of the binding site, the macroscopic (K) and microscopic (k) dissociation constants are equal (Cantor and Schimmel, 1980). For binding of Myb to a DNA containing two copies of the binding site, the macroscopic equilibrium is $P+D \rightleftharpoons PD$, whereas microscopically $P+D \rightleftharpoons PD_1$ and $P+D \rightleftharpoons PD_2$ (where PD is the protein–DNA complex at the macroscopic level and PD_1 and PD_2 are the microscopic complexes in which the first or second site in the DNA is occupied, respectively). Since $K = (P)(D)/PD$; $k = (P)(D)/PD_1 = (P)(D)/PD_2$ and $PD = PD_1 + PD_2$, it follows that $K = k/2$. Similarly, the binding of a protein with two independent binding sites (e.g. TRF1) to a DNA containing one copy of the binding site also gives a value of K that is half the value of k. The binding of such a protein to a DNA containing two binding sites instead gives (for the first binding event) a value of K that is one-quarter of the value of k, since there are four different protein–DNA complexes at the microscopic level: P_1D_1, P_1D_2, P_2D_1 and P_2D_2. Thus, for both Myb and TRF1, in the absence of cooperativity, a 2-fold increase in affinity is to be expected when the number of DNA sites is increased from one to two.

DNase I footprinting

Oligonucleotides shown in Figure 4 were cloned into the *Sma*I site of pBend2 (Konig *et al*., 1998) and the *Nhe*I–*Hin*dIII fragment was gel-purified. The 3′ end of the G-rich and C-rich strands were labeled by filling in the *Nhe*I and *Hin*d III sites, respectively, using reverse transcriptase. Binding reactions included DNA at ~1×10^{-9} M and the isolated Myb domain ($TRF1_{371–439}$; Konig *et al*., 1998) at 4×10^{-8} M in 20 mM HEPES pH 8.0, 100 mM KCl, 10 µg/ml BSA, 0.1% Triton X-100, 5% glycerol, 2 mM $MgCl_2$, 10 µg/ml sonicated calf thymus DNA or sonicated *E.coli* DNA. TRF1 concentrations (approximations based on activity assays) are indicated in the figure legends. DNase I (0.05 U/ml) digestions were at rt in 22 µl. Aliquots (10 µl) were removed at 3 and 6 min and the reactions terminated with 10 µl of 6 mM EDTA and 20 µl of phenol. Samples were then extracted with phenol/chloroform, 1 µg of sonicated calf thymus DNA was added and DNA was collected

by ethanol precipitation. Pellets were washed with 70% ethanol (–20°C) and dried. Products were fractionated on 8% polyacrylamide (38:2)/8 M urea gels and detected by autoradiography.

Preparation of DNA–protein complexes for electron microscopy

Plasmids pF96F, pH96F and pH96H were created by cloning into the *Bam*HI and Asp718 sites of pBluescript KS(–), a pair of TRF1 binding sites (either 5′-TTAGGGTTA-3′ or 5′-TTAGGGTTAGGGTTA-3′) in all three combinations and separated by 96 bp of random DNA sequence. Plasmids were linearized with *Sca*I before TRF1 binding and EM analysis. All binding reactions were performed for 30 min on ice in a buffer containing 20 mM HEPES–KOH pH 7.9, 100 mM KCl, 0.5 mM DTT and 0.1 mM EDTA. Reactions included DNA at 10 µg/ml and TRF1 at 25–50 dimers per full TRF1 binding site. Reactions were terminated by addition of glutaraldehyde to 0.6% for 5 min at rt. The samples were run through Bio-Gel A-5 m (Bio-Rad) columns equilibrated with 10 mM Tris–HCl pH 7.6, 0.1 mM EDTA. Filtered samples were mixed with a spermidine-containing buffer, adsorbed to glow-charged thin carbon foils, dehydrated through a series of water–ethanol washes and rotary shadowcast with tungsten as described previously (Griffith and Christiansen, 1978). Samples were visualized in a Philips EM400 instrument. Micrographs were scanned from negatives using a Nikon multiformat film scanner. The contrast was optimized and panels were arranged using Adobe Photoshop. Morphometry measurements were done using a Summagraphics digitizer with software developed by J.D.Griffith.

Acknowledgements

We thank Heidi Moss (Rockefeller University) for essential and generous help during the last stages of the work. Matthew Turner (Rockefeller University) is thanked for assistance in calculating the expected site frequencies in the oligonucleotide library, Peter Konig (University California San Francisco) for preparations of purified Myb domain and Jason Lue (Rockefeller University) for preparations of purified TRF1. Diego Loayza is thanked for comments on the manuscript. Supported by NIH grants to J.D.G. (GM31819 and CA70343) and T.de L. (GM49046) and by a Human Frontier Science Program grant to T.de L. and D.R. (RG0323). A.B. was supported by a Comitato Promotore Telethon graduate fellowship.

References

Arrowsmith,C.H. and Morin,P. (1996) New insights into p53 function from structural studies. *Oncogene*, **12**, 1379–1385.

Baranowskij,N., Frohberg,C., Prat,S. and Willmitzer,L. (1994) A novel DNA binding protein with homology to Myb oncoproteins containing only one repeat can function as a transcriptional activator. *EMBO J.*, **13**, 5383–5392.

Bianchi,A. (1999) *Characterization of DNA Binding Activities at Vertebrate Telomeres*. Doctoral dissertation. The Rockefeller University, New York, NY.

Bianchi,A., Smith,S., Chong,L., Elias,P. and de Lange,T. (1997) TRF1 is a dimer and bends telomeric DNA. *EMBO J.*, **16**, 1785–1794.

Bilaud,T., Koering,C.E., Binet-Brasselet,E., Ancelin,K., Pollice,A., Gasser,S.M. and Gilson,E. (1996) The telobox, a Myb-related telomeric DNA binding motif found in proteins from yeast, plants and human. *Nucleic Acids Res.*, **24**, 1294–1303.

Bilaud,T., Brun,C., Ancelin,K., Koering,C.E., Laroche,T. and Gilson,E. (1997) Telomeric localization of TRF2, a novel human telobox protein. *Nature Genet.*, **17**, 236–239.

Blasco,M.A., Funk,W., Villeponteau,B. and Greider,C.W. (1995) Functional characterization and developmental regulation of mouse telomerase RNA. *Science*, **269**, 1267–1270.

Broccoli,D., Chong,L., Oelmann,S., Fernald,A.A., Marziliano,N., van Steensel,B., Kipling,D., Le Beau,M.M. and de Lange,T. (1997a) Comparison of the human and mouse genes encoding the telomeric protein, TRF1: chromosomal localization, expression and conserved protein domains. *Hum. Mol. Genet.*, **6**, 69–76.

Broccoli,D., Smogorzewska,A., Chong,L. and de Lange,T. (1997b) Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2. *Nature Genet.*, **17**, 231–235.

Cantor,C.C. and Schimmel,P.R. (1980) *Biophysical Chemistry, Part III: The Behaviour of Biological Macromolecules*. Freeman, New York, NY.

Chong,L., van Steensel,B., Broccoli,D., Erdjument-Bromage,H., Hanish,J., Tempst,P. and de Lange,T. (1995) A human telomeric protein. *Science*, **270**, 1663–1667.

Cooper,J.P., Nimmo,E.R., Allshire,R.C. and Cech,T.R. (1997) Regulation of telomere length and function by a Myb-domain protein in fission yeast. *Nature*, **385**, 744–747.

Cutler,G., Perry,K.M. and Tjian,R. (1998) Adf-1 is a nonmodular transcription factor that contains a TAF-binding Myb-like motif. *Mol. Cell. Biol.*, **18**, 2252–2261.

Feldbrugge,M., Sprenger,M., Hahlbrock,K. and Weisshaar,B. (1997) PcMYB1, a novel plant protein containing a DNA-binding domain with one MYB repeat, interacts *in vivo* with a light-regulatory promoter unit. *Plant J.*, **11**, 1079–1093.

Feng,J. *et al.* (1995) The RNA component of human telomerase. *Science*, **269**, 1236–1241.

Gehring,W.J. *et al.* (1994) Homeodomain-DNA recognition. *Cell*, **78**, 211–223.

Griffith,J.D. and Christiansen,G. (1978) Electron microscope visualization of chromatin and other DNA–protein complexes. *Annu. Rev. Biophys. Bioeng.*, **7**, 19–35.

Griffith,J., Bianchi,A. and de Lange,T. (1998) TRF1 promotes parallel pairing of telomeric tracts *in vitro*. *J. Mol. Biol.*, **278**, 79–88.

Griffith,J.D., Comeau,L., Rosenfield,S., Stansel,R.M., Bianchi,A., Moss,H. and de Lange,T. (1999) Mammalian telomeres end in a large duplex loop. *Cell*, **97**, 503–514.

Hanish,J.P., Yanowitz,J. and de Lange,T. (1994) Stringent sequence requirements for telomere formation in human cells. *Proc. Natl Acad. Sci. USA*, **91**, 8861–8865.

Hiraoka,Y., Henderson,E. and Blackburn,E.H. (1998) Not so peculiar: fission yeast telomere repeats. *Trends Biochem. Sci.*, **23**, 126.

Hsu,H.L., Huang,L., Tsan,J.T., Funk,W., Wright,W.E., Hu,J.S., Kingston,R.E. and Baer,R. (1994) Preferred sequences for DNA recognition by the TAL1 helix–loop–helix proteins. *Mol. Cell. Biol.*, **14**, 1256–1265.

King,D.A., Zhang,L., Guarente,L. and Marmorstein,R. (1999) Structure of a HAP1–DNA complex reveals dramatically asymmetric DNA binding by a homodimeric protein. *Nature Struct. Biol.*, **6**, 64–71.

Kirik,V. and Baumlein,H. (1996) A novel leaf-specific myb-related protein with a single binding repeat. *Gene*, **183**, 109–113.

Konig,P. and Rhodes,D. (1997) Recognition of telomeric DNA. *Trends Biochem. Sci.*, **22**, 43–47.

Konig,P., Giraldo,R., Chapman,L. and Rhodes,D. (1996) The crystal structure of the DNA-binding domain of yeast RAP1 in complex with telomeric DNA. *Cell*, **85**, 125–136.

Konig,P., Fairall,L. and Rhodes,D. (1998) Sequence-specific DNA recognition by the myb-like domain of the human telomeric binding protein TRF1: a model for the protein–DNA complex. *Nucleic Acids Res.*, **26**, 1731–1740.

Kota,R.S. and Runge,K.W. (1998) The yeast telomere length regulator TEL2 encodes a protein that binds to telomeric DNA. *Nucleic Acids Res.*, **26**, 1528–1535.

Larson,G.P., Castanotto,D., Rossi,J.J. and Malafa,M.P. (1994) Isolation and functional analysis of a *Kluyveromyces lactis* RAP1 homologue. *Gene*, **150**, 35–41.

Lipsick,J.S. (1996) One billion years of Myb. *Oncogene*, **13**, 223–235.

Marcand,S., Gilson,E. and Shore,D. (1997) A protein-counting mechanism for telomere length regulation in yeast. *Science*, **275**, 986–990.

Martin,C. and Paz-Ares,J. (1997) MYB transcription factors in plants. *Trends Genet.*, **13**, 67–73.

Nishikawa,T., Nagadoi,A., Yoshimura,S., Aimoto,S. and Nishimura,Y. (1998) Solution structure of the DNA-binding domain of human telomeric protein, hTRF1. *Structure*, **6**, 1057–1065.

Ogata,K., Morikawa,S., Nakamura,H., Sekikawa,A., Inoue,T., Kanai,H., Sarai,A., Ishii,S. and Nishimura,Y. (1994) Solution structure of a specific DNA complex of the Myb DNA-binding domain with cooperative recognition helices. *Cell*, **79**, 639–648.

Payre,F., Buono,P., Vanzo,N. and Vincent,A. (1997) Two types of zinc fingers are required for dimerization of the serendipity δ transcriptional activator. *Mol. Cell. Biol.*, **17**, 3137–3145.

Rosinski,J.A. and Atchley,W.R. (1998) Molecular evolution of the Myb family of transcription factors: evidence for polyphyletic origin. *J. Mol. Evol.*, **46**, 74–83.

Shore,D., Langowski,J. and Baldwin,R.L. (1981) DNA flexibility studied by covalent closure of short fragments into circles. *Proc. Natl Acad. Sci. USA*, **78**, 4833–4837.

Smith,T.L. and Sauer,R.T. (1995) P22 Arc repressor: role of cooperativity in repression and binding to operators with altered half-site spacing. *J. Mol. Biol.*, **249**, 729–742.

Tanikawa,J., Yasukawa,T., Enari,M., Ogata,K., Nishimura,Y., Ishii,S. and Sarai,A. (1993) Recognition of specific DNA sequences by the c-myb protooncogene product: role of three repeat units in the DNA-binding domain. *Proc. Natl Acad. Sci. USA*, **90**, 9320–9324.

van Leeuwen,H.C., Strating,M.J., Rensen,M., de Laat,W. and van der Vliet,P.C. (1997) Linker length and composition influence the flexibility of Oct-1 DNA binding. *EMBO J.*, **16**, 2043–2053.

van Steensel,B. and de Lange,T. (1997) Control of telomere length by the human telomeric protein TRF1. *Nature*, **385**, 740–743.

van Steensel,B., Smogorzewska,A. and de Lange,T. (1998) TRF2 protects human telomeres from end-to-end fusions. *Cell*, **92**, 401–413.

Wilson,D., Sheng,G., Lecuit,T., Dostatni,N. and Desplan,C. (1993) Cooperative dimerization of paired class homeo domains on DNA. *Genes Dev.*, **7**, 2120–2134.

Wotton,D., Ghysdael,J., Wang,S., Speck,N.A. and Owen,M.J. (1994) Cooperative binding of Ets-1 and core binding factor to DNA. *Mol. Cell. Biol.*, **14**, 840–850.

Yie,J., Liang,S., Merika,M. and Thanos,D. (1997) Intra- and intermolecular cooperative binding of high-mobility-group protein I (Y) to the β-interferon promoter. *Mol. Cell. Biol.*, **17**, 3649–3662.

Zhang,L. and Guarente,L. (1996) The C6 zinc cluster dictates asymmetric binding by HAP1. *EMBO J.*, **15**, 4676–4681.

Zhong,Z., Shiue,L., Kaplan,S. and de Lange,T. (1992) A mammalian factor that binds telomeric TTAGGG repeats *in vitro*. *Mol. Cell. Biol.*, **13**, 4834–4843.

Received July 12, 1999; revised August 23, 1999; accepted August 24, 1999