1999

Characterization of DNA Binding Activities at Vertebrate Telomeres

Alessandro Bianchi

Characterization of DNA binding activities at vertebrate telomeres

A thesis presented to the faculty of The Rockefeller University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

By

Alessandro Bianchi

Advisor

Titia de Lange

August 1998

The Rockefeller University

New York, New York

Dedicated to Tania, just like I am...

Acknowledgments

First and foremost I would like to thank Titia. Writing a meaningful account of what it has been to work with her in these past years is no easy task. I can say that I leave the lab with very good memories, even of the hardest days, because it has all been an intense learning experience, and because Titia is a special person and an amazing scientist. I consider myself extremely lucky in having had the chance to join the lab almost at its beginning, in a moment when success was still a few steps away and uncertain. Seeing the perseverance, the enthusiasm, the hard work, the talent and especially the optimism in always looking at the good side of an experiment, have all been very good lessons. And if Titia may be a little demanding at times (...) I can say that she has always given back more than she has asked me, and that is ... more than one can ask for. Her genuine love for science has had no small part in making these years fun and exciting. I miss it already.

I want to thank all the members of the lab, past and present, without naming the whole list because I am afraid to forget somebody and also because I have pestered everybody so much that now being too nice and grateful would sound a little weird... Thanks, everybody! First place in the list to Kiki and Susan, for seniority, if nothing else... A special mention to Jason, who had the unfortunate task of having to put up with me in the last year, to Agata, who doubled student forces in the lab, to Heidi, and to Jan, who supplied kilograms of Taq polymerase. And of course Laura, who was instrumental in the biochemical purification of TRF1 in the first place, and who was always a wonderful target for daily jokes, until she ran away to Hawaii, exhausted.

I would like to thank Kenji and Magda who have always been available to give advice on many issues, whenever asked. And Claude, for the months spent in his lab. And Tom Sakmar and Chris Min, for hospitality and advice in getting the baculovirus system going. And Gisela Schnapp for providing partially purified telomerease fractions.

Many thanks go to Jack for the hospitality in his lab and for taking such a fruitful interest in TRF1 and TRF2.

And to Carolyn for her enthusiasm and friendliness, starting before, and continuing during, her stay on sabbatical in Titia's lab. And to Art with whom we shared many journal clubs.

A really special and big thank you to Daniela, for being Daniela of course, and for her decision to spend a short sabbatical in the lab, which resulted in three very enjoyable months that were also a great source of learning for me.

I am grateful to Louise, for sharing information on TRF1 binding, and to Peter for providing purified Myb domain.

I am indebted (deeply!) to Stew for being always so efficient in solving all the little computer bugs that obstruct the progress of science. Many thanks go to Matthew Turner for help in calculating the probabilities of the SELEX sites. And to Marta Delgado, for having abolished censorship on purchases of science books from student accounts. And to the members of my commettee, Nat, Magda and Carol, for reading the thesis and for whatever else is to happen in two weeks.

This being the end of student life, a special thought to Marco for having shared many of the moments, good and bad, of it.

Muito obrigado, to my mother in law, for having given up her life in Brazil for four long months to help take care of Giulia and Flavio while I was running gels and writing this thesis.

Molte grazie, to my parents, who always gave me support and freedom in my choices, and who are going through difficult moments and are probably secretly not so fond of this dissertation which has kept them from their two grandchildren, one of whom they have never met.

A final thought to my sister, who was born when I was six and has never had a brother that was not in school, and finally will!

And then Giulia and Flavio, who are simply unbelievably cute...

INDEX

List of figures

Abstract

Telomere function is carried out by protein factors that either bind directly to telomeric DNA or that are recruited there by DNA binding activities. Thus, telomeric DNA binding activities are central to telomere function. Two types of vertebrate DNA binding activities are characterized in this thesis: proteins that can potentially bind to the 3' overhang of TTAGGG repeats and proteins that bind to duplex telomeric DNA.

We have identified, characterized and partially purified an activity from chicken erythrocytes (ATEF) that has the biochemical properties expected for a telomere terminus factor. In addition, we have analyzed the DNA binding properties of the non-specific DNA-end binding factor, human Ku, which is involved in DNA end joining. We show that long G-G folded overhangs do not prevent Ku binding to telomeric DNA *in vitro*, indicating that, similar to the situation in yeast, Ku could potentially bind to chromosome ends.

Most of the mammalian telomeric DNA is in double-stranded form and is bound *in vivo* by two related factors, TRF1 and TRF2, which both contain one Myb-type repeat. We have conducted a detailed analysis of the DNA binding mode of TRF1 and shown that it binds DNA as a dimer by contacting a bipartite DNA site with each Myb domain. The

relative orientation and spacing of the two sites does not affect binding. We propose a model for TRF1 binding that predicts the existence of a flexible region, located between the highly conserved dimerization and DNA binding domains, that would allow free relative rotation of the Myb domains. This is a novel binding mechanism that, although different from the one displayed by the yeast telomeric protein Rap1p, could be of relevance to other telomeric proteins, such as the human TRF2 and the fission yeast Taz1p, that share a similar domain organization. Finally, we have shown that TRF1 or TRF2 bound on telomerase substrates *in vitro* do not affect the activity of the enzyme.

Introduction

An approach to the characterization of the telomeric complex

The concept of telomere as a specialized terminal section of the chromosome, devoted to prevent fusion with other chromosomes, goes back in time many decades, to the pioneering studies of Muller and McClintock (Muller 1938; McClintock 1941). However, it is only in the last 20 years that it has been possible to move beyond the abstract functional-cytological concept of "telomere" into a molecular characterization of the structure. In parallel, the list of proposed telomeric functions has grown considerably, to include the by-pass of the DNA end replication problem, the control of cellular life-span, tumorigenesis, and the meiotic pairing of homologous chromosomes.

A full understanding of telomere function rests on the molecular characterization of the telomeric complex. The work of the last several years has established that, in most organisms, telomeric DNA is composed of short repeats arranged in tandem arrays, and that the maintenance of these repeats is essential for telomeric function. However, telomeric DNA alone is not sufficient and telomeres function as nucleoprotein structures. Telomeric repeats recruit a variety of DNA binding factors at the telomere and these in turn interact, in either a stable or transient manner, with additional proteins. Thus, for simplicity, we can imagine the telomeric complex as structurally organized at four different levels of complexity: first, the DNA; second, protein factors bound directly to the DNA; third, additional protein factors stably recruited to the telomere by the DNA binding activities; fourth, additional factors, possibly enzymatic (for instance, kinases), interacting with the telomere in a transient but regulated manner.

The cloning of telomeric DNA in vertebrates (humans) dates to 1988 (Moyzis et al. 1988). In comparison, it was only in 1995 that the first vertebrate telomeric protein (TRF1, TTAGGG Repeat biding Factor 1) was cloned (Chong et al. 1995). In order to

characterize the telomeric complex in its structural organization and protein components, we have undertaken the approach of identifying and characterizing DNA binding activities capable to interact with telomeric DNA repeats, either in single-stranded or double-stranded form. In Chapters 1-4 experiments are described that are aimed at clarifying our understanding of this level of organization of the telomeric complex in vertebrates. In this introductory section I will first discuss the DNA component of telomeres and its replication by the specialized enzyme telomerase. I will then briefly review the current understanding of how the maintenance of telomeric DNA is regulated and how it is critical for cellular function. Finally, I will describe how telomere function appears to be mediated by the higher level of telomere organization, the protein components.

The DNA component of the telomeric complex

The first insight into the molecular structure of telomeres came with the determination of the DNA sequence of telomeres in *Tetrahymena thermophila* (Blackburn and Gall 1978). In this ciliated protozoan telomeric DNA was found to be composed of tandemly repeated copies of the sequence 5'-TTGGGG-3'. It was subsequently determined that similarly short (up to 25 bp, but typically 5-8 bp) tandemly repeated sequences are found in most other organisms (see (Henderson 1995) for review). The total number of repeats at each telomere is normally heterogeneously distributed (with the exception of hypotrichous ciliates, see below) and varies considerably among species: from about 300-400 bp in yeast and *Tetrahymena* (Shampay et al. 1984; Larson et al. 1987; Shampay and Blackburn 1988) to about 10-100 kb in mouse (Kipling and Cooke 1990; Zijmans et al. 1997). Human cells normally have between 2 and 20 kb

of telomeric DNA (Allshire et al. 1988; Moyzis et al. 1988; Allshire et al. 1989; Brown 1989; Cross et al. 1989; de Lange et al. 1990; Hastie et al. 1990). The fact that telomere length is species-specific suggests that the process is under genetic control.

As in the *Tetrahymena* case, most known examples of telomeric repeats show a bias in base composition with respect to strand polarity. Usually the strand that runs in a 5' to 3' direction toward the end of the chromosome is rich in T and G residues, with short runs of Gs often present. In most cases, all the repeats found at telomeres in a particular organism are identical (Tommerup et al. 1994), with the important exception of budding and fission yeast and *Paramecium tetraurelia*, where considerable variability is observed (Shampay et al. 1984; Sugawara and Szostak 1986; McCormick-Graham et al. 1997).

Subtelomeric DNA elements of various kinds but generally much longer than the short repeats found terminally, have been described in a variety of organisms (see (Henderson 1995) for review). No function has been assigned to such elements in mammalian cells. In budding yeast two types of subtelomeric elements, X and Y', probably originating from transposition events, are found at most chromosomes (Chan and Tye 1983; Louis and Haber 1992). When the replication of telomeric repeats is impaired by the introduction of mutations in telomerase genes, Y' elements apparently take over in the telomere maintenance pathway in a small fraction of the cells and are found at all or most chromosome termini in these cells (Lundblad and Blackburn 1993). Interestingly, a non-telomerase based mechanism for the maintenance of terminal sequences also appear to be at work in wild type *Drosophila melanogaster*, where transposable elements instead of telomeric repeats carry out telomere function (Levis et al. 1993). However, both in yeast and in mammals, under normal circumstances the

short telomeric repeats appear to represent the only DNA element required for telomeric function, as demonstrated by various kinds of telomere regeneration experiments in yeast (Szostak and Blackburn 1982; Brown 1989; Cross et al. 1989; Riethman et al. 1989; Sandell and Zakian 1993), and mammalian cells (Wilkie et al. 1990; Farr et al. 1991; Barnett et al. 1993; Hanish et al. 1994).

Hypotrichous ciliates, with their peculiar chromosome biology (see Introduction of Chapter 1), have provided the unique opportunity to determine the precise configuration of telomeric DNA at the terminus. The fact that in the vegetative macronucleus chromosomes undergo fragmentation and duplications and that all the telomeres of these minichromosomes have exactly the same length has allowed direct sequencing of both strands, which in turn has revealed that in these organisms short overhangs of the G-rich strand are present (Klobutcher et al. 1981). Overhangs of the same polarity have subsequently been observed during S-phase in budding yeast (Wellinger et al. 1993; Wellinger et al. 1996) and both in cycling and quiescent mammalian cells (Makarov et al. 1997; McElligott and Wellinger 1997; Wright et al. 1997). In these organisms, though, the length of the single-stranded tails appears to be heterogeneous and considerably longer than in ciliates. The possibility that unusual DNA structures may be formed at these overhangs exists and is discussed in Chapter 2. The fact that single-stranded overhangs are a common feature throughout such diverse organisms at least at some point during the life cycle, raises the obvious possibility that G-overhangs may carry out some conserved and essential function at telomeres. In any case, the presence of single-stranded DNA with its recognized ability to act as a signal in the DNA-damage recognition pathway (Garvik et al. 1995; Lydall and Weinert 1995),

strongly implies the existence of binding activities capable of preventing activation of DNA-damage processing.

Replication and maintenance of telomeric DNA

The replication of linear DNA molecules poses a problem due to the requirement of a primer for DNA synthesis by DNA polymerases (Watson 1972; Olovnikov 1973). After the removal of the primer from the 5' terminus of the newly synthesized strand, a gap is left that cannot be adequately filled-in. Over generations and subsequent rounds of DNA replication, the recurring loss of terminal DNA would ultimately prove fatal in the absence of a mechanism to replace the lost DNA. Telomeres provide such a mechanism with the DNA-polymerase activity of the enzyme telomerase ((Greider and Blackburn 1985), reviewed in (Greider 1995)). The biochemical characteristics of telomerase are described in more detail in Chapter 4. Here, I will just point out that the enzyme is a ribonucleoprotein, with an RNA component that provides a template for the synthesis of the repeats (Greider and Blackburn 1987). A polypeptide with homology to reverse transcriptases is essential for telomerase activity (Counter et al. 1997; Harrington et al. 1997b; Lingner et al. 1997b; Meyerson et al. 1997; Nakamura et al. 1997; Nakamura and Cech 1998) and various other proteins associated with the enzyme have been reported (Collins et al. 1995; Lingner and Cech 1996; Bednenko et al. 1997; Harrington et al. 1997a; Nakayama et al. 1997).

Before telomerase was discovered, various recombination mechanisms had been proposed to explain how telomeric repeats could allow the complete replication of the chromosome terminus (reviewed in (Blackburn and Szostak 1984)). It now appears that recombination is used only as a back-up mechanism in case of telomerase inactivation

(Pluta and Zakian 1989; Lundblad and Blackburn 1993; Bryan et al. 1995; Li and Lustig 1996; McEachern and Blackburn 1996), and that the primary pathway for telomere maintenance depends on telomerase (Yu et al. 1990; Singer and Gottschling 1994; McEachern and Blackburn 1995). *In vitro* expression studies suggest that the human catalytic subunit (hTERT, telomere reverse transcriptase) and the human RNA (hTER, telomerase RNA) may constitute the only essential components for catalysis, but the possibility exists that additional essential activities are contributed by the extracts used for expression (Weinrich et al. 1997; Beattie et al. 1998). At the level of expression, however, it appears that the limiting factor for telomerase regulation in primary human cells is hTERT and not hTER (Broccoli et al. 1995; Blasco et al. 1996; Meyerson et al. 1997; Bodnar et al. 1998; Counter et al. 1998; Kolquist et al. 1998).

In yeast, screens for a senescence phenotype induced by telomere attrition have identified 4 genes implicated in the telomerase pathway: EST1-4 (Lundblad and Szostak 1989; Lendvay et al. 1996). EST1 and EST4 encode nucleic acid binding activities (see below and introduction of Chapter 1), while EST2 is the catalytic subunit of telomerase and EST3 has unknown function (Nugent et al. 1996; Virta-Pearlman et al. 1996; Counter et al. 1997; Lingner et al. 1997b; Morris and Lundblad 1997). Even though the EST genes generate similar phenotypes and belong to the same complementation group, only EST2 is required for telomerase activity *in vitro* (Lingner et al. 1997a). This result suggests that, at least in yeast, other components (Est1p, Est3p and Est4p) in addition to the RNA and the catalytic subunit may be required for the *in vivo* function of telomerase, possibly by acting as loading factors or by rendering the substrate accessible.

Telomeres appear to be in a dynamic equilibrium, clustering around an average length determined by the two opposing processes of shortening and elongation (reviewed in (Greider 1996)). Shortening occurs in *Euplotes* in a developmentally regulated manner (Vermeesch and Price 1994), in *Tetrahymena* when growing conditions are altered (Larson et al. 1987), and in yeast when longer than normal telomeres are introduced in the cell (Li and Lustig 1996). A dynamic equilibrium of telomere length has been revealed in yeast by the analysis of individual telomeres in clonal cell populations (Shampay and Blackburn 1988). In addition, the replacement of the wild type telomerase RNA with a copy mutated in the template region has shown that the mutated telomeric repeats appear at telomeres first terminally and only gradually, over subsequent cell divisions, more internally, indicating that a stochastic telomere shortening mechanism is at work that is counterbalanced by telomerase elongation (McEachern and Blackburn 1995).

Telomere length regulation: a mitotic clock for cellular senescence

Yeast or ciliate cells carrying a mutated copy of the telomerase RNA that leads to the synthesis of altered telomere repeats, show various kinds of growth defects (Yu et al. 1990; Singer and Gottschling 1994; McEachern and Blackburn 1995; Krauskopf and Blackburn 1996), demonstrating that the presence of a functional telomeric complex is critical. In yeast the loss of a single telomere generates cell cycle arrest and chromosome loss (Sandell and Zakian 1993). In human cells, impairment of the function of telomeric protein TRF2 (TTAGGG Repeat binding Factor 2) gives rise to the formation of chromosome fusions and to cellular senescence (van Steensel et al. 1998). Thus, telomeres are required for normal cell division and cell cycle progression.

Consistent with the end-replication problem of linear chromosomes, in telomerase-negative human primary fibroblasts grown *in vitro*, progressive telomere shortening is observed (Harley et al. 1990), and the number of cell divisions that these cells can undergo correlates with telomere length (Allsopp et al. 1992). A correlation also exists between donor age and telomere length in human somatic tissues (Harley et al. 1990; Hastie et al. 1990). In sperm cells, on the other hand, telomeres are longer than the telomeres from somatic cells from the same donor (Cooke and Smith 1986; de Lange et al. 1990), in keeping with the fact that telomerase activity is present in germline cells (Kim et al. 1994; Mantell and Greider 1994; Prowse and Greider 1995). These observations have been interpreted in the context of Hayflick and Moorhead's proposition that the finite doubling capacity of human cells in culture constitutes an *in vitro* model for cellular aging (Hayflick and Moorhead 1961). In this view telomeres represent a mitotic clock that counts the number of divisions that a cell can undergo before its replicative potential is exhausted (Counter et al. 1992). Telomerase has thus become the center of attention of studies of cellular senescence. The interest in this enzyme has recently been rewarded by the demonstration that stable expression of the catalytic subunit in normal human cells is sufficient for apparently unlimited extension of the dividing capacity of these cells (Bodnar et al. 1998; Vaziri and Benchimol 1998). Thus, in at least some circumstances, the presence of a functional telomere appears to be the limiting factor in determining the ability of a cell to sustain cell divisions. Wider claims that telomeres may have a role in determining aging at the organismal level fail to explain the fact that mice are a short-lived species despite having extremely long telomeres (Kipling and Cooke 1990).

Telomere length regulation: a tumor suppressor mechanism?

The implications of the finding that telomeres act as a mitotic clock to prevent unlimited cell division are not restricted to the phenomenon of cellular senescence but extend to tumor biology, where the loss of telomeric function may act as a barrier to malignant transformation (Goldstein 1990; de Lange 1998b). This idea is supported by the observation that most human tumors have high levels of telomerase activity, compared to the absence of telomerase in most normal somatic cell types (Counter et al. 1994; Kim et al. 1994; Meyerson et al. 1997). Studies in knock-out mice for the telomerase RNA, though, have shown that oncogenically transformed cells from these mice form tumors in nude mice at normal frequencies, questioning, at least in this species, the requirement of telomerase activation for tumor formation (Blasco et al. 1997). However, more recent studies have shown that mTR$^{-/-}$ mice show impaired tumorigenesis in the INK4a$^{-/-}$ background compared to wild type (R. Greenberg, C. Greider and R. DePinho, personal communication). Thus, even though differences in tumorigenesis between mice and humans are considerable and underscored by the high frequency at which mouse cells undergo immortalization *in vitro*, telomere shortening could act as a tumor suppressor mechanism in both species. Such mechanism might be more effective in humans due to a higher selective pressure against tumorigenesis in this longer-lived species.

Telomeric DNA binding activities: single-strand binding factors

As discussed above, single-stranded overhangs may be a universal feature of eukaryotic telomeres. Therefore, overhang-binding factors with a protective function may constitute a requirement for all cell types. In addition, end-binding activities may be

needed to facilitate the interaction of telomerase or other processing enzymes with the DNA. End binding activities are described in detail in the introductions to Chapters 1 and 2.

In ciliates, thanks to the presence of millions of telomeres per cell, end-binding activities have been first, and best, characterized (Gottschling and Zakian 1986; Price 1990; Prescott 1994). In these organisms end factors appear to carry out a capping function (Gottschling and Zakian 1986; Price and Cech 1987; Price and Cech 1989) but it is still unclear whether analogous activities exist in organisms that do not present this unususal chromosome fragmentation and duplication process (see introduction of Chapter 1).

In budding yeast, genetic and biochemical screens have identified several genes that may participate in end-binding (Lundblad and Szostak 1989; Lin and Zakian 1994; Lendvay et al. 1996). Two of these genes, EST4/CDC13 and YKU70-80, when impaired, lead to the loss of terminal DNA in the C-strand (Garvik et al. 1995; Gravel et al. 1998). In addition, Cdc13p and Yku70-80p possess single-strand and end-specific DNA binding activity, respectively ((Nugent et al. 1996) and Chapter 2). Both CDC13 and YKU70 interacts genetically with telomerase and with one another (Nugent et al. 1996; Nugent et al. 1998; Polotnianka et al. 1998).

In vertebrates, the evidence for single-strand binding factors at telomeres is more scant. We have identified a factor from frog eggs and chicken erythrocytes that has biochemical characteristics similar to those of the ciliate proteins, but an answer to its role at telomeres awaits the isolation of the gene ((Cardenas et al. 1993) and Chapter 1). More recently, heterogenous nuclear ribonucleoprotein A1 (hnRNP A1, which binds single-stranded telomeric repeats *in vitro*) has been shown to be involved in

telomere length regulation (McKay and Cooke 1992; Ishikawa et al. 1993; LaBranche et al. 1998).

Our understanding of how the terminal part of the telomere is packaged remains very unsatisfactory. However, the capping role of end factors is likely to be a key component in the assembly of a functional telomeric complex. The phenotype of cdc13 and yku70 cells in yeast strongly suggests the existence of one, or more, specific nucleases, whose activity at telomeres needs to be kept in check. In addition, a thorough description of end binding factors may be required for an adequate understanding of telomerase function. EST1 and EST4 may be examples of genes that bridge end-binding to telomerase activity, since they possess the binding requirements for such an activity and display a telomerase-negative phenotype (Lendvay et al. 1996; Nugent et al. 1996; Virta-Pearlman et al. 1996). In addition, it appears that the very terminal repeats in yeast telomeres affect telomerase activity in a different manner compared to more internally located repeats (Krauskopf and Blackburn 1996), suggesting the possibility that repeats at the terminus may be bound by factors other than Rap1p (Repressor activator protein 1, see below).

Telomeric DNA binding activities: double-strand binding factors

Telomeric proteins that bind to telomeric repeats in double-stranded form are the best characterized from a functional point of view. These proteins include Rap1p in *S.cerevisiae* (Shore and Nasmyth 1987), Taz1p in *S.pombe* (Telomere-associated in *Schizosaccharomyces pombe* (Cooper et al. 1997)) and the mammalian proteins TRF1 and TRF2 (Chong et al. 1995; Bilaud et al. 1997; Broccoli et al. 1997b). The DNA

binding characteristic of Rap1p and the TRFs will be described in more detail in Chapter 3, here I will summarize the genetic and functional data on these proteins.

Rap1p is an essential protein that has multiple functions in yeast cells, acting both as an activator and as a repressor of transcription, dependent on local context (see (Shore 1994) for review). The protein is 827 amino acids long, with a centrally located DNA binding domain spanning amino acids 361-596 that is composed of two Myb-type repeats (Konig et al. 1996). The silencing properties of Rap1p are well characterized, and Rap1p-dependent repression of transcription occurs both at the HMR and HML loci and at telomere-proximal regions (Telomere Position Effect, TPE). The C-terminal domain of Rap1p is essential for silencing and it is involved in telomere length regulation (Conrad et al. 1990; Kyrion et al. 1992; Moretti et al. 1994). Several classes of mutations in RAP1 have been identified that affect telomere length. A group of temperature-sensitive mutants (rap^{ts}) with point mutations in the DNA binding domain have shorter telomeres (Lustig et al. 1990). Intragenic suppressors of one of the rap^{ts} mutants (rap1-5) make up a second class of RAP1 alleles ($rap1^{t}$) that have extremely elongated telomeres (Kyrion et al. 1992). All of the $rap1^{t}$ mutations introduce stop codons that result in the elimination of the C-terminal domain. A group of missense mutations in the C-terminal domain also result in telomere elongation ($rap1^{s}$) (Buck and Shore 1995). These results clearly implicate the C-terminal domain in telomere length regulation. Overexpression of this domain results in telomere elongation, possibly by a squelching mechanism, consistent with the idea that Rap1p-interacting factors act as negative regulators of telomere length (Conrad et al. 1990). Two such factors have been identified by two-hybrid screens: Rif1p and Rif2p (Rap1p-Interacting Factors 1 and 2) (Hardy et al. 1992; Wotton and Shore 1997). Mutant strains in either of these two

genes have longer telomeres, and rif1-rif2 double mutants have telomeres as long as the rap1^t strains (Wotton and Shore 1997). Thus, the Rap1p C-terminal domain appears to regulate telomere length via the recruitment of Rif1p and Rif2p at the telomeric complex. Recent experiments in which Rap1p molecules have been tethered to the telomeres by a heterologous DNA binding domain, have shown that the sensing mechanism that regulates telomere length is able to count the number of Rap1p molecules bound to the telomere, independently of their orientation (Marcand et al. 1997).

Similarities to the yeast model for telomere length regulation have emerged from *S.pombe* and mammalian cells. In fission yeast, mutations in the gene for the double-strand telomere binding protein Taz1p result in telomere elongation (Cooper et al. 1997). In mammalian cells, two telomeric proteins have been identified that bind to double stranded TTAGGG repeats, TRF1 and TRF2 (Zhong et al. 1992; Bilaud et al. 1996; Broccoli et al. 1997b). The proteins have a similar domain structure (see the introduction to Chapter 3) and form homodimers (Bianchi et al. 1997; Broccoli et al. 1997b). Truncated versions of the proteins have been generated that affect the ability of the wild type protein to bind to telomeric DNA both *in vitro* and *in vivo* ((Bianchi et al. 1997; van Steensel and de Lange 1997; van Steensel et al. 1998) and Chapter 3). The introduction of the TRF1 dominant negative allele in HT1080 cells results in progressive telomere elongation, indicating that TRF1 acts as negative regulator of telomere length, possibly by inhibiting access of telomerase to the telomere ((van Steensel and de Lange 1997), see the introduction to Chapter 4). Consistent with this model, overexpression of TRF1 or TRF2 in HT1080 cells generates gradual shortening of telomeric DNA ((van Steensel and de Lange 1997) and A. Smogorzewska, B. van Steensel and T. de Lange,

unpublished). Induction of the TRF2 dominant negative allele in HT1080 cells generates a more dramatic phenotype, that includes the formation of chromosome end-to-end fusions and the loss of the terminal G-strand overhangs (van Steensel et al. 1998). Thus, TRF2-mediated overhang protection may be necessary for capping function in mammalian cells, raising the intriguing and unexpected possibility that in this system capping function may be not be carried out by single-strand binding factors.

In addition to a role in the control of telomere length, telomeric proteins may have a function in meiotic cells. This has emerged from studies in fission yeast where mutants for *taz1* and a second gene that affects telomeric silencing (*lot2*) result in loss of the normal clustering of telomeres around the spindle pole body (SPB) in the pre-meiotic stage (Cooper et al. 1998; Nimmo et al. 1998). Taz1p- and Lot2p-mediated clustering of the telomeres at the SPB appears to be important for the homologous pairing of the chromosomes, since both mutants are defective in meiotic recombination. Although fission yeast has a rather unusual prophase I, with "horse-tail" movements of the telomeres-SPB complex that appear to facilitate homologue chromosome pairing (Chikashige et al. 1994), clustering of telomeres around the centrosome is observed in mammalian cells (see (Dernburg et al. 1995) for review). Because of this similarity, and the structural similarity between TRF proteins and Taz1p (see Introduction of Chapter 3), it is possible that an analogous meiotic function is carried out by telomeric proteins at mammalian telomeres (de Lange 1998a).

Not only DNA binding factors: higher order structure of the telomeric complex

The counting mechanism that appears to be at work at yeast and mammalian telomeres raises the question of how the signal from the total mass of telomeric protein bound at each telomere is transmitted to telomerase and the telomere replication machinery. It appears possible that some sort of higher order organization of the DNA is involved.

Non-nucleosomal chromatin structure at telomeres has been documented in ciliates and mammals, as well as in yeast (Blackburn and Chiou 1981; Gottschling and Cech 1984; Wright et al. 1992; Tommerup et al. 1994; Wright and Zakian 1995). However, it is in the latter organism that the most complete evidence for the presence of an altered chromatin structure at telomeres has been produced. As mentioned above, Rap1p is involved in transcriptional repression at telomeres (TPE, see (Shore 1995) for review). TPE is mediated by the assembly at yeast telomeres of a large protein complex that includes Sir2p, Sir3p, and Sir4p (Silent Information Regulators) and histones H3 and H4 (Hecht et al. 1996). Interaction of the C-terminal domain of Rap1p with Sir3p and Sir4p has been demonstrated by two-hybrid assays (Moretti et al. 1994). Targeting of Sir3p and Sir4p to the telomeres by heterologous DNA binding domains is sufficient to establish silencing, suggesting that the role of Rap1p in this process is in recruiting Sir3p and Sir4p (Lustig et al. 1996; Marcand et al. 1996). Sir3p and Sir4p also interact with the N-termini of histones H3 and H4, which are required for silencing (Hecht et al. 1995). The biological significance of TPE in budding yeast has not been determined yet. However, fission yeast also has TPE, even though no homology is detected between the species with regards to the factors involved (Nimmo et al. 1994).

Interestingly, mutations in SIR3 and SIR4 cause telomere shortening (Palladino et al. 1993), possibly because the proteins normally compete with Rif1p and Rif2p for binding to Rap1p.

TPE has not been detected in mammalian cells. However, some evidence indicates that an altered chromatin structure exists in mammals. In mammalian cells, telomeric DNA is packaged in nucleosomes that are smaller that in normal chromatin (Makarov et al. 1993; Tommerup et al. 1994) and micrococcal digestion patterns suggest that terminal telomeric DNA may be packaged in a non-nucleosomal structure (Tommerup et al. 1994). More direct evidence for an unusual chromatin organization of mammalian telomeres comes from EM studies with TRF1, in which telomere-telomere associations are observed *in vitro* ((Griffith et al. 1998), see Chapter 3).

An additional level of complexity in mammals, likely to influence chromatin organization, is introduced by the ongoing identification of TRF1- and TRF2-interacting factors (S. Smith, A. Schmidt and T. de Lange, in press; and B. Li and T. de Lange, unpublished). The characterization of these factors and the identification of the additional ones that, most likely, still remain unidentified, will prove necessary for our understanding of the functioning of the telomeric complex.

Objectives

The multiple functions of the vertebrate telomeric complex require the assistance of a multitude of protein factors, only some of which probably will turn out to be stable components of the complex. However, factors that bind to telomeric DNA are likely to be central to our understanding of the telomeric complex as they provide the scaffold for regulating the access by additional regulatory components. For this reason we have

chosen to focus our attention on the identification and characterization of DNA binding activities at vertebrate telomeres.

In the first two Chapters I address the issue of the presence at vertebrate telomeres of capping factors with single-standed DNA binding activity. As mentioned above, in vertebrates these end factors remain elusive. In Chapter 1 I describe the identification, characterization and partial purification of a putative telomere end factor from avian cells. In Chapter 2 I characterize the DNA binding properties of human protein Ku, whose yeast homologue is part of the telomeric complex.

In Chapter 3 I present a characterization of the DNA binding properties of the mammalian telomeric proteins TRF1 and TRF2. Understanding how TRF1 and TRF2 assemble on telomeric DNA may offer insight into how these telomeric proteins negatively affect telomere length regulation.

Finally, in Chapter 4 I test *in vitro* the hypothesis that TRF1 and TRF2 directly inhibit telomerase access to the DNA terminus.

Chapter 1

In search of a vertebrate telomere end binding activity

SUMMARY

Telomere end binding activities have been biochemically characterized in ciliated protozoa. Candidate end factors have also been identified genetically in yeast. Previously we had identified a DNA binding activity from *Xenopous laevis* oocytes (Xenopus Telomere End Factor, XTEF) that has nucleic acid binding characteristics reminiscent of ciliate telomeric proteins. XTEF binds to single-stranded telomeric repeats of the form 5'-$(TTAGGG)_2$-3' when present at a 3' terminus but not when they are located at a 5' terminus or internally. XTEF binding is sequence specific and facilitated by adjacent duplex DNA. Importantly, XTEF does not bind the RNA sequence 5'-$(UUAGGG)_2$-3'. We have identified a similar, and possibly identical, activity from chicken erythrocytes (ATEF, Avian TEF) and report a protocol for the partial purification of this avian protein.

INTRODUCTION

Telomeres protect chromosome ends from nucleolytic attack and terminal fusions (Muller 1938; McClintock 1942; Sandell and Zakian 1993; van Steensel et al. 1998) presumably through the formation of a non-nucleosomal protein complex at the telomere, which has been documented in ciliates, yeast and mammals (Gottschling and Cech 1984; Price 1990; Wright et al. 1992; Tommerup et al. 1994). The precise composition of the telomeric complex has not been resolved yet, but two types of DNA binding activities are likely to be involved: double-stranded binding factors and end binding factors. The former class includes well characterized examples from budding yeast, fission yeast, and mammals (Shore and Nasmyth 1987; Chong et al. 1995; Konig et al. 1996; Broccoli et al. 1997b; Cooper et al. 1998). The identification of members of the latter class has presented a more serious challenge, presumably because of the low abundance of terminal binding activities. This problem has been successfully circumvented by taking advantage of the unusual chromosome biology of ciliates. These organisms assign replicative function to a germinal micronucleus and transcriptional function to a somatic macronucleus. The macronucleus is highly polygenomic and contains multiple copies of fragmented chromosomes. About 200 minichromosomes of average length of about 600 kb exist in the holotrichous ciliate *Tetrahymena*, in addition to about 10^4 copies of a 20 kb long minichromosome containing the rDNA genes. The hypotrichous ciliates *Euplotes* and *Oxytricha* each carry about $2x10^4$ 2 kb long fragments per nucleus (see (Prescott 1994) for review).

In *Oxytricha nova* about $5x10^7$ telomeric ends are present in the macronucleus. The fact that most (if not all) of these ends has identical conformation has allowed the

direct determination of the terminal DNA sequence (Klobutcher et al. 1981), which has the following configuration:

NNNGGGGTTTTGGGGTTTTGGGGTTTTGGGGTTTTGGGG3'
NNNCCCCAAAACCCCAAAACCCC5'

This terminal DNA is specifically bound by a heterodimeric factor (α/β) that protects the guanines in the overhang from methylation by DMS (Gottschling and Zakian 1986; Price and Cech 1987; Gray et al. 1991). The heterodimer is composed of a 56 (α) and a 41 (β) Kd subunit (Hicke et al. 1990; Gray et al. 1991) and has great specificity of binding for the *Oxytricha* telomere overhang (Raghuraman et al. 1989). The α subunit by itself is able to bind DNA but the interaction is greatly stabilized by DNA-dependent heterodimerization with the β subunit (Fang and Cech 1993b; Fang et al. 1993). Perhaps the most striking features of α/β binding to DNA is the extreme stability of the complex. Likely due to hydrophobic interactions, the binding is stable to incubation in 2.5 M NaCl or 6 M CsCl (Gottschling and Zakian 1986; Price and Cech 1987; Price and Cech 1989). Even more remarkably, the dissociation of the ternary complex is very slow, with a half life of about 100 hours (Fang et al. 1993). Such tenacious mode of binding is what might be expected for a telomere terminus factor with a protective function. In fact, α/β protects macronuclear DNA from nuclease digestion *in vitro* (Gottschling and Zakian 1986). In addition to the *Oxytricha* heterodimer two α homologs have been identified in *Euplotes crassus* and one in *Stylonicha mytilus* (Price 1990; Fang and Cech 1991; Wang et al. 1992). Interestingly, one of the *Euplotes* proteins appears to be involved in telomere DNA replication rather than in the capping function (Carlson et al. 1997; Fan and Price 1997).

To date, it remains unclear if analogous activities exist in organisms other than ciliates and the question remains open as to whether these proteins are unique to this

group of organisms due to their peculiar chromosomal biology. Nevertheless, candidates for telomere end factors have been identified in organisms such as *Chlamydomonas* (with a *S. cerevisiae* homolog), *Tetrahymena,* and budding yeast (Petracek et al. 1994; Konkel et al. 1995; Sheng et al. 1995). Lin and Zakian screened λ expression libraries with a labeled telomeric G-strand oligonucleotide and identified two *S. cerevisiae* genes, both of which contained RNA recognition motifs (RRMs) but failed, when deleted, to cause any observable telomeric phenotype (Lin and Zakian 1994). Two other candidates in yeast, EST1 and EST4/CDC13, have emerged from screens for senescence phenotypes caused by telomere attrition (Lundblad and Szostak 1989; Lendvay et al. 1996). Both of these genes have been shown to encode proteins that display single-stranded G-strand specific DNA binding (Lin and Zkian 1996; Nugent et al. 1996; Virta-Pearlman et al. 1996). Est1p in particular appears to require a free 3' terminus for DNA binding activity, although it is able to bind non-specifically to RNA with higher affinity (Virta-Pearlman et al. 1996). Whether these proteins represent true capping factors or rather, for example, replication or telomerase associated factors, remains to be determined. Interestingly, a component of the yeast telomeric complex with DNA end binding activity but no sequence specificity for telomeric sequences has been identified in Yku70-80p, a factor involved in double-strand-break repair (see Chapter 2).

In vertebrates, a protein that binds to G-G base pairs in a proposed hairpin terminal structure has been described (Gualberto et al. 1992). In addition, several efforts to purify a telomere end factor have resulted in the isolation of components of the heterogeneous nuclear ribonucleoprotein family (hnRNPs) (McKay and Cooke 1992; Ishikawa et al. 1993). These proteins have specific DNA binding activity for telomeric G-strand DNA but their affinity for RNA of the same sequence is higher, suggesting that

the binding to telomeric DNA *in vitro* may be fortuitous. Nevertheless, one of these proteins (hnRNP A1) has recently been implicated in telomere length regulation in mouse cells (LaBranche et al. 1998).

Since no vertebrate homolog to any of these proteins exists in the database and the rate of evolution of telomeric proteins appear to be high (Wang et al. 1992; Broccoli et al. 1997a; Cooper et al. 1997) we have undertaken a biochemical approach to the identification of a vertebrate end binding activity. We have identified an activity from *Xenopus* oocytes (XTEF) that resembles the telomere end factors of hypotrichous ciliates (Cardenas et al. 1993). XTEF binds specifically to a double-stranded oligonucleotide that contains a 5'-$(TTAGGG)_2$-3' overhang. The presence of an adjacent duplex enhances binding and a free 3' terminus is required. Importantly, XTEF does not bind the RNA sequence 5'-$(UUAGGG)_2$-3'. In addition, XTEF binding is highly salt resistant. Thus, XTEF possesses the biochemical characteristics of telomere terminus end factors as described in the ciliate system. We describe the identification of an XTEF-like activity in chicken erythrocytes and its partial purification.

RESULTS

Characterization of a putative telomere end binding activity from chicken erythrocytes (ATEF)

Because of the difficulty in obtaining large amounts of *Xenopus* oocytes for biochemical purification, we searched for XTEF-like activity in various other cell types. Band-shif assays performed with nuclear extracts from chicken erythrocytes revealed a DNA binding activity (Avian Telomere End Factor, ATEF) that specifically recognized a telomeric overhang. Binding experiments were performed with a labelled oligonucleotide containing a duplex moiety flanked by the overhang 5'-$(TTAGGG)_2$-3' (oligonucleotide PTE, Fig.1.1) and the specificity of binding was examined by performing competitions with DNA molecules that contained either mutations in the single-stranded overhang or had different configurations at the terminus (Figs. 1.1 and 1.2). The competition assays indicated that ATEF had binding specificity essentially identical to XTEF (Cardenas et al. 1993). ATEF, like XTEF, had high binding affinity for single-stranded $(TTAGGG)_2$ when the sequence was at a 3' terminus (oligonucleotides PTE and G2DNA, Figs. 1.1 and 1.2a). When the same single-stranded sequence was located at a 5' or internally in the DNA the affinity was much lower (oligonucleotides R.duplex+3'ds and 5'repeats, Figs. 1.1 and 1.2a). In addition, the RNA equivalent of G2DNA (G2RNA) did not compete for ATEF binding (Fig. 1.2a). Finally, the binding to the 3' overhang was sequence specific, with the 3 guanines apparently being the most critical residues (Figs. 1.1 and 1.2b).

Even though the amount of ATEF extracted from chicken erythrocytes nuclei is small (about 10-30 molecules/nucleus, as estimated by quantitative gel shift) compared to the amount extractable from frog eggs ($3x10^7$ molecules/cell), the possibility to

Fig. 1.1 **DNA substrates used for characterization of the candidate avian telomere end factor (ATEF)**

Asterisks indicate positions at which a labeled residue was present after fill-in reaction with Klenow polymerase.


PTE
5'AAACTCGACCAGTACAGCTACTTAGGGTTAGGG
3'***GAGCTGGTCATGTCGATG
R.duplex+3'ds
5'AAACTCGACCAGTACAGCTACTTAGGGTTAGGGCAGCTGAGATCTCCC**
3'***GAGCTGGTCATGTCGATG GTCGACTCTAGAGGGAA
5'Repeats
5'TTAGGGTTAGGGCAGCTGAGATCTCCC**
3'GTCGACTCTAGAGGGAA
G2DNA
5'AAAACTCGACTAGTGCATCGACTTAGGGTTAGGG
G2RNA
5'AAAACUCGACUAGUGCAUCGACUUAGGGUUAGGG
(TTTGGG)2tail
5'AAAACTCGACTTAGGGTTAGGGTTTGGGTTTGGG
3'***TGAGCTGAATCCCAATCCC
(TTACCC)2tail
5'AAAACTCGACTTAGGGTTAGGGTTACCCTTACCC
3'***TGAGCTGAATCCCAATCCC
(AAAGGG)2tail
5'AAAACTCGACTTAGGGTTAGGGAAAGGGAAAGGG
3'***TGAGCTGAATCCCAATCCC

Fig. 1.1

easily obtain large amounts of chicken blood made erythrocytes a preferable source of material for biochemical purification.

Partial purification of ATEF

As a critical step in the purification scheme we developed an affinity column based on a duplex oligonucleotide containing a biotinylated residue at one end and a single-stranded overhang of the sequence 5'-$(TTAGGG)_2$-3' at the other end. The oligonucleotide was coupled to an agarose-streptavidine matrix. ATEF displayed very high affinity for such a column and in fact it proved very difficult to recover activity from it. High salt (2 M KCl) elution conditions were insufficient for recovery, whereas urea and guanidinum chloride appeared to denature the protein irreversibly. We found that elution with a carbonate buffer at pH 10.7 allowed recovery of activity with a yield that varied from about 5 to 30%, probably due to partial inactivation of the protein. For these and other reasons the affinity column was chosen as the last step in the purification scheme (see below).

The necessity to operate at the smallest possible volumes in the last steps of the purification, in order to minimize losses, conflicted with the early observation that the loading of large amounts of protein on the affinity column resulted in very poor recovery. To circumvent this problem we attempted to achieve a partial purification of the factor that would allow efficient loading and elution at the last step. A large scale purification was attempted from about 10^{13} nuclei derived from 8.5 l of chicken blood. The 10g of total protein in the extracts were applied to six subsequent chromatographic steps: heparin-sepharose, phosphocellulose, TMAE-sepharose, Sephacryl-300, CM Sepharose and single-stranded DNA cellulose (Fig.1.3). Columns were ulted with linear

Fig. 1.2 **Competition assay for DNA binding specificity of ATEF**

Band-shift assays were performed with labeled oligonucleotide PTE (Fig. 1.1). Samples were run on a native 5% polyacrylamide gel. Competitor DNAs are as indicated at top of each lane (see Fig. 1.1 for DNA sequence of oligonucleotides). Molar excess of competitor is 3, 12 and 48x in each set of competition assays. The ATEF-DNA complex is indicated. Other, non-specific, DNA binding activities were observed.

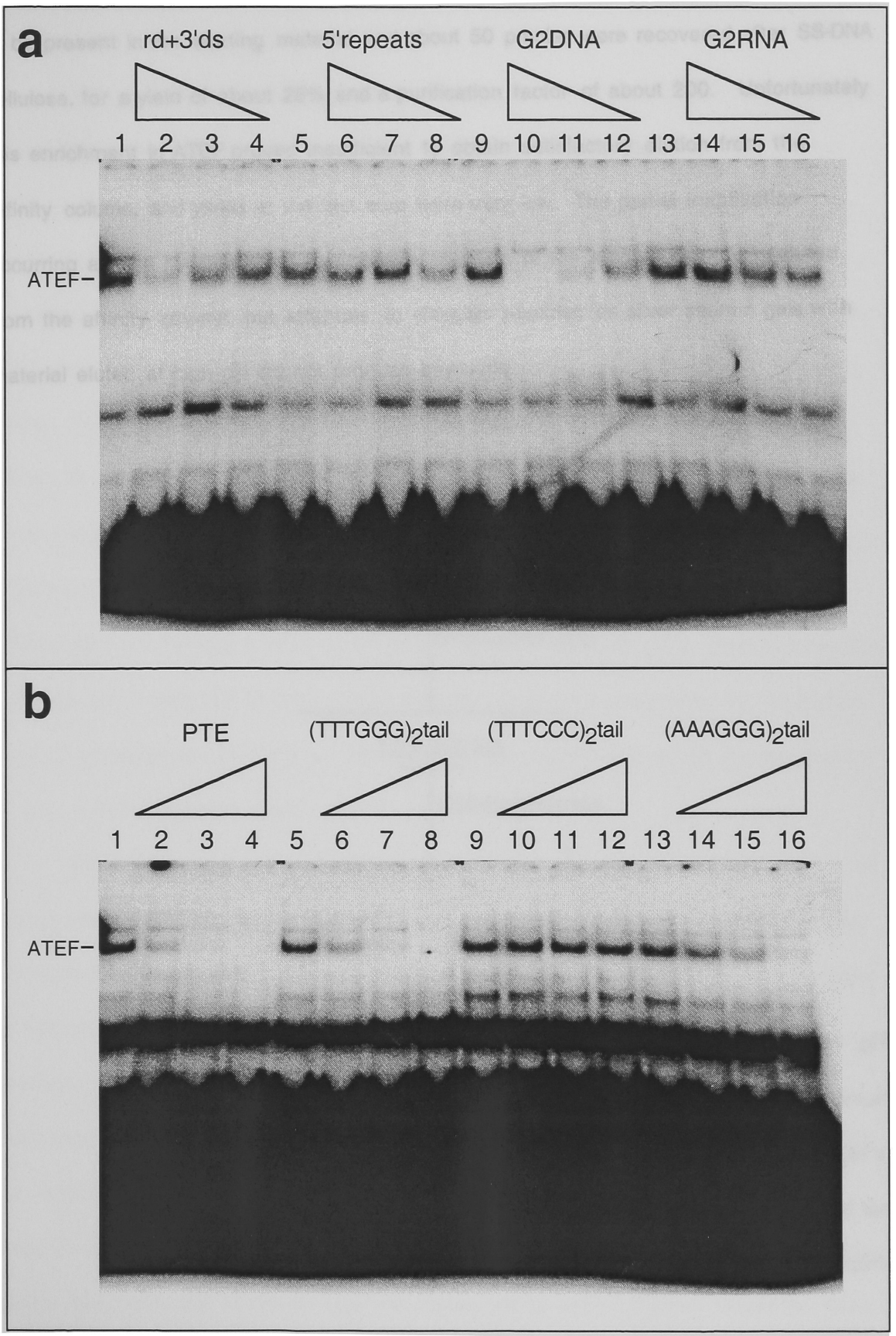

a
rd+3'ds
5'repeats
G2DNA
G2RNA
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
ATEF
b
PTE
(TTTGGG)2tail
(TTTCCC)2tail
(AAAGGG)2tail
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
ATEF

Fig. 1.2

salt gradients ranging from 0.1 to 1 M KCl. About 200 pmoles of ATEF were estimated to be present in the starting material and about 50 pmoles were recovered after SS-DNA cellulose, for a yield of about 25% and a purification factor of about 200. Unfortunately this enrichment in ATEF proved insufficient to obtain satisfactory elution from the affinity column, and yields at the last step were very low. The partial inactivation occurring at high pH prevented an accurate estimate of the amount of protein eluted from the affinity column, but attempts to visualize peptides on silver stained gels with material eluted at high pH did not produce any band.

Fig. 1.3 **Scheme of purification protocol**

Numbers 0.1 and 1 indicate KCl concentration, in molarity. Each diagonal bar is representative of the salt gradient used for elution of each column and the bar below the gradient indicates the approximate position in the gradient at which ATEF activity was recovered.

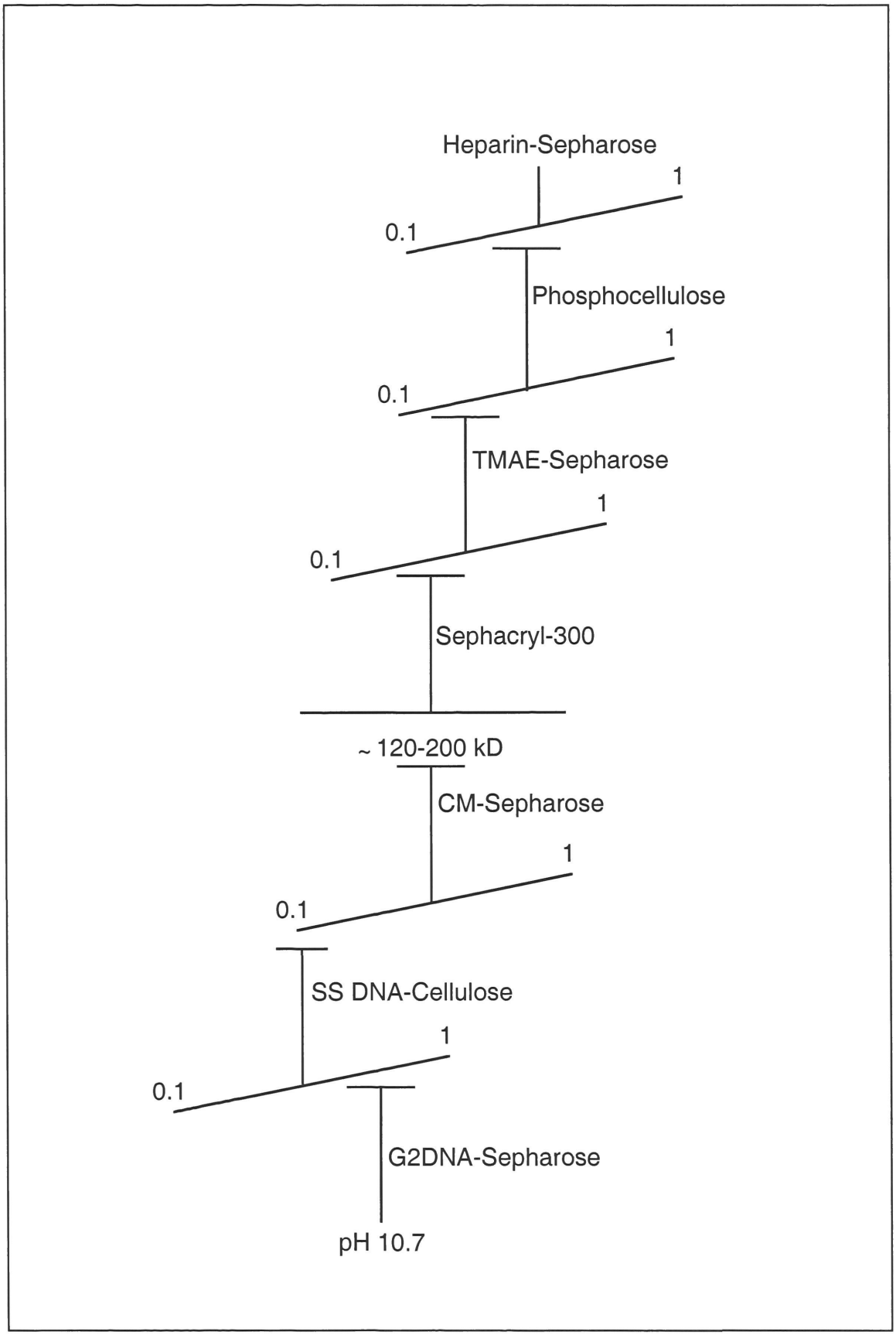

Heparin-Sepharose
1
0.1
Phosphocellulose
1
0.1
TMAE-Sepharose
1
0.1
Sephacryl-300
~ 120-200 kD
CM-Sepharose
1
0.1
SS DNA-Cellulose
1
0.1
G2DNA-Sepharose
pH 10.7

Fig. 1.3

DISCUSSION

With the development of an affinity purification step for ATEF we have identified a possible route to the biochemical purification of the protein. Unfortunately further work is required for the identification of additional significant purification steps. We have not been able to identify conditions that would allow >3 fold purification with any of various ion exchangers, sizing resins or hydrophobic materials. Only a single-stranded DNA cellulose resin offered adequate results (about 10 fold purification with a yield of 90%). Furthermore, we were forced by the limited amount of starting material to cut losses throughout the procedure and therefore to compromise purification. Perhaps a more successful approach would limit drastically the chromatographic steps to a phosphocellulose column, followed by SS-DNA cellulose and a much larger (50-100x) affinity column. Possibly a much larger affinity column would not suffer from overloading problem, and losses due to large elution volumes would be compensated by the higher yield of a streamlined procedure. Nevertheless, the problems posed by the purification of such a low abundance protein appear still daunting.

Since this project was undertaken our understanding of the terminal structure of vertebrate (in fact, mammalian) telomeres has progressed significantly. Evidence has emerged that indicates the presence of long G-strand overhangs (130-260 nucleotides) at half, and possibly all, mammalian telomeres (Makarov et al. 1997; McElligott and Wellinger 1997; Wright et al. 1997). These overhangs are much longer than the overhangs found at the telomeres of hypotrichous ciliates. Therefore it is possible that the requirements for capping factors are much different at mammalian telomeres. Indeed if an α/β-like end factor is present at vertebrate telomeres, it now seems unlikely that all the protective functions could be carried out by such a protein, unless the factor

has the ability to bind internally on the overhang without requiring a free 3' terminus. Thus, the presence of a free 3' end for binding should not be a strict requirement for such candidate activities and searches for telomere capping proteins should not be limited by this requirement.

It is interesting that two end binding activities that have recently been implicated in telomere biology (Ku and hnRNP A1) show no specificity for telomeric DNA but, rather, appear to carry out at the telomere only part of their cellular role. Thus the possibility exists that capping functions may be carried out by non-exclusively telomeric proteins at the telomere. The ciliate paradigm at vertebrate telomeres has thus been replaced by a possible scenario in which multiple factors may be needed to take care of the long overhangs and factors involved in double-strand-break repair and splicing may perform some of the function of telomere-specific end binding activities. Nevertheless, at some level unique telomeric functions may have to be specified (for example, recruitment of telomerase) that require the presence of telomere-specific end binding factors. ATEF/XTEF, because of its DNA binding specificity, remains a good candidate for such an activity.

Chapter 2

Ku binds telomeric DNA *in vitro*

SUMMARY

Ku is a heterodimeric protein with high binding affinity for ends, nicks and gaps in double-stranded DNA. Both in mammalian cells and in budding yeast Ku plays a role in non-homologous end joining in the double-strand-break (DSB) repair pathway. However, Ku has a more significant role in DNA repair in mammalian cells compared to yeast, where a homology-dependent pathway is the predominant one. Recently Ku has been shown to be a likely component of the telomeric complex in yeast, suggesting the possibility of a similar role for Ku at mammalian telomeres. However, long single-stranded G-rich overhangs are continuously present at mammalian but not at yeast telomeres. These overhangs have the potential to fold *in vitro* into G-G base paired conformations, such as G-quartets, that might prevent Ku from recognizing telomeric ends and thus offer a mechanism to sequester the telomere from the prevalent DSB repair pathway in mammals. We show here that Ku binds to mammalian telomeric DNA ends *in vitro* and that G-quartet conformations are unable to prevent Ku from binding with high affinity to the DNA. Our results indicate that the DNA binding characteristics of Ku are consistent with its direct interaction with telomeric DNA in mammalian cells and its proposed role as a telomere end factor.

INTRODUCTION

Ku is a heterodimer of 70 and 80 kD subunits that was originally identified as an autoimmune antigen and subsequently found to be involved in V(D)J recombination and double-strand-break (DSB) repair in mammalian systems((Mimori and Hardin 1986) see (Jackson 1996; Lieber et al. 1997) for review, (Liang et al. 1996; Nussenzweig et al. 1996; Zhu et al. 1996; Gu et al. 1997; Jhappan et al. 1997)). Ku has sequence independent affinity for double-stranded DNA ends, both blunt and 3'- or 5'-recessed, and for internal nicks or gaps (Mimori et al. 1986; de Vries et al. 1989; Blier et al. 1993; Falzon et al. 1993). Binding to single-stranded DNA occurs with lower affinity (Mimori and Hardin 1986; de Vries et al. 1989). Sequence specific DNA and RNA binding has also been reported (Giffin et al. 1996; Yoo and Dynan 1998) but its *in vivo* significance has not been established. The mechanism of binding remains largely unknown, but actual DNA ends do not seem to be required, since Ku can bind dumbbell-shaped molecules and circular DNA containing single-stranded bubbles (Falzon et al. 1993). Interestingly, Ku is able to slide internally into the duplex DNA after initial binding to the end (de Vries et al. 1989) and internalized Ku molecules are resistant to high salt conditions after DNA circularization, suggesting the possibility of a doughnut shaped heterodimer (Paillard and Strauss 1991).

Recently, Ku homologs have been identified in budding yeast (YKU70/HDF1 and YKU80/HDF2 (Feldmann and Winnacker 1993; Boulton and Jackson 1996a; Feldmann et al. 1996)). In this organism, contrary to mammalian cells, a Rad52-dependent homologous recombination pathway is responsible for the majority of the DSB repair events (Friedberg et al. 1995). However, in a rad52 background, a second pathway

becomes apparent that is Ku-dependent and involves joining of non-homologous ends (Boulton and Jackson 1996b; Milne et al. 1996; Siede et al. 1996; Tsukamoto et al. 1996; Tsukamoto et al. 1997a).

In addition to its role in repair, YKU70-80 plays a role at telomeres. Both yku70 and yku80 cells are partially defective in telomere maintenance, with telomeres in mutant cells shrinking to about 70% of their normal length (Boulton and Jackson 1996a; Porter et al. 1996; Nugent et al. 1998). Yku70-80p may be an integral component of the yeast telomeric complex. First, YKU70 and YKU80 have an effect on telomere position effect (TPE), a phenomenon in which genes located near telomeres are transcriptionally repressed. TPE in yeast telomeres is dependent on a multiprotein complex that includes the Sir2p, Sir3p and Sir4p (see (Grunstein 1997) for review). Loss of Yku70-80p has a dramatic effect on TPE, equivalent to impairment of SIR2, SIR3 or SIR4 (Boulton and Jackson 1998; Nugent et al. 1998). Second, although the nuclear localization of Yku70/80p has not been determined directly, in yku70 cells the subnuclear localization of telomeres may be altered (Laroche et al. 1998). Finally, crosslinking studies indicate that Yku80p is bound to telomeric DNA *in vivo* and loss of YKU70-80 function generates strand-specific loss of telomeric DNA (Gravel et al. 1998). How the dual role of Ku in yeast as a telomere end factor and as a DSB repair protein can be reconciled, remains to be determined.

The DNA binding activity of Ku suggests that Ku may bind directly to telomeric DNA. However, because of the interaction of Yku70p with Sir4p in yeast two-hybrid assays (Tsukamoto et al. 1997b), it is also possible that the interaction of Yku70-80p with the telomere is mediated by protein-protein interactions. In addition, the effect of single-stranded telomeric tails on Ku binding to double-stranded DNA remains unknown,

an issue of interest in the light of the recognized ability of telomeric G-rich strands to adopt G-G base paired conformations *in vitro* (see (Henderson 1995) for review). Indeed, telomeric G-strand DNA *in vitro* readily forms, under physiological conditions, a structure in which 4 guanine residues from 4 different telomeric repeats are arranged in a planar conformation (G-quartet or G-DNA) that is stabilized by Hoogsteen-type hydrogen bonds. For any particular telomeric sequence typically 3 or 4 of these planar arrangements are stacked onto each other and are stabilized by Na^+ or K^+ ions loosely coordinated at the center. Several types of related structures, all belonging to the G-quartet family, have been described, with variability residing on the relative orientation of the 4 strands (parallel or antiparallel) and the loops connecting them (diagonal or lateral) (Sen and Gilbert 1988; Sundquist and Klug 1989; Williamson et al. 1989; Kang et al. 1992)).

Mammalian telomeres terminate in long (130-270 bp) single-stranded G-rich overhangs (Makarov et al. 1997; McElligott and Wellinger 1997; Wright et al. 1997), whereas long (>20 bp) overhangs occur at yeast telomeres only briefly in late S-phase (Dionne and Wellinger 1996; Wellinger et al. 1996). Thus, differences in the requirements for telomere end binding activities might exist in the two organisms. In particular, the possibility exists that Ku in mammalian cells is limited to a role in DSB repair and that the G-rich overhangs at mammalian telomeres might be sufficient to sequester telomeric ends from this repair pathway. Our results indicate that the ability of Ku to bind to telomeric DNA ends in different conformations is consistent with a direct interaction of Ku with mammalian telomeric DNA and its proposed role as a telomeric protein.

Ku binds to telomeric DNA ends *in vitro*

Given the end binding activity of Ku and its suggested role at yeast telomeres, it was of interest to investigate the ability of Ku to bind to telomeric DNA *in vitro*. Oligonucleotides were synthesized terminating in the human telomeric TTAGGG repeats in different configurations, both 3' protruding and blunt ended (Fig. 2.1a; T1, T2 and T3). Human Ku was purified to apparent homogeneity from HeLa cells nuclear extracts (de Vries et al. 1989) (Fig. 2.1b). The affinity of Ku for the various substrates in band-shift assays was then compared to the affinity for an oligonucleotide of similar size but of random sequence (Fig. 2.1a; R1). Two representative binding experiments performed with the control DNA and the three telomeric DNAs as competitors are shown in Fig. 2.2. In these and in other experiments (data not shown) the telomeric oligonucleotides competed efficiently for Ku binding to either the telomeric DNA T1 (Fig. 2.2a) or the control DNA R1 (Fig. 2.2b). The telomeric substrates, in fact, proved to be slightly better competitors than the random sequence DNA. T3 in particular performed as the best competitor. This result is in agreement with the known preference of Ku for longer double-stranded regions (Falzon et al. 1993). In summary, these results show that DNAs terminating with telomeric sequence are able to bind Ku with an affinity comparable to that displayed by non-telomeric DNA ends.

Fig. 2.1a **DNA oligonucleotides used in binding and competition experiments with Ku**

Telomeric sequences are indicated in bold.

Fig. 2.1b **Silver stained SDS-PAGE gel of purified Ku**

a

		length of duplex	length of (each) single strand
R1	TTAAGGACATGCCCGGGCATGTCC CCTGTACGGGCCCGTACAGGAATT	20	-
R2	ACTCGACCAGTACAGCTACGTGAGTGTGATG GAGCTGGTCATGTCGATGCACTCACACTACA	30	-
R3	AAACATCACACTCACGTAGCTGTACTGGTCGAG	-	33
T1	CCCTGTAGCATCGACGGGTTAGGGTTAGGG GGGATTGGGATTGGGACATCGTAGCTGCCC	18	12
T2	CTAACCGAACACTGACGCTAGGGTTAGGGTTAGGGT GATTGGGATTGGGATTGGCTTGTGACTGCGATCCCA	24	12
T3	CTAACCCTAACCGAACACTGACGCTAGGGTTAGGGT GATTGGGATTGGCTTGTGACTGCGATCCCAATCCCA	36	36
Q1	TACTCGACCAGTAGACTCACGTGAGAATTCCTA$G_4T_4G_4T_4G_4T_4G_4$ $G_4T_4G_4T_4G_4T_4G_4$ATGAGCTGGTCATCTGAGTGCACTCTTAAGGAT	33	28
Q2	$T_4G_4T_4G_4T_4G_4T_4G_4$	-	32

b

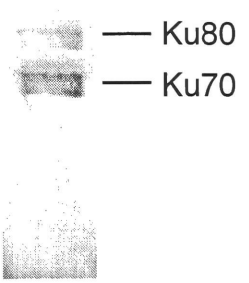


Fig. 2.1

Fig. 2.2 **Competition assays for Ku binding with telomeric and non-telomeric DNA oligonucleotides**

Labelled oligonucleotides T1 (a) or R1 (b) were present at a concentration of about 1.5 nM (T1) and 0.5 nM (R2) and incubated in the presence of DNA competitors and about 0.2 nM (a) or 0.05 nM (b) purified Ku.

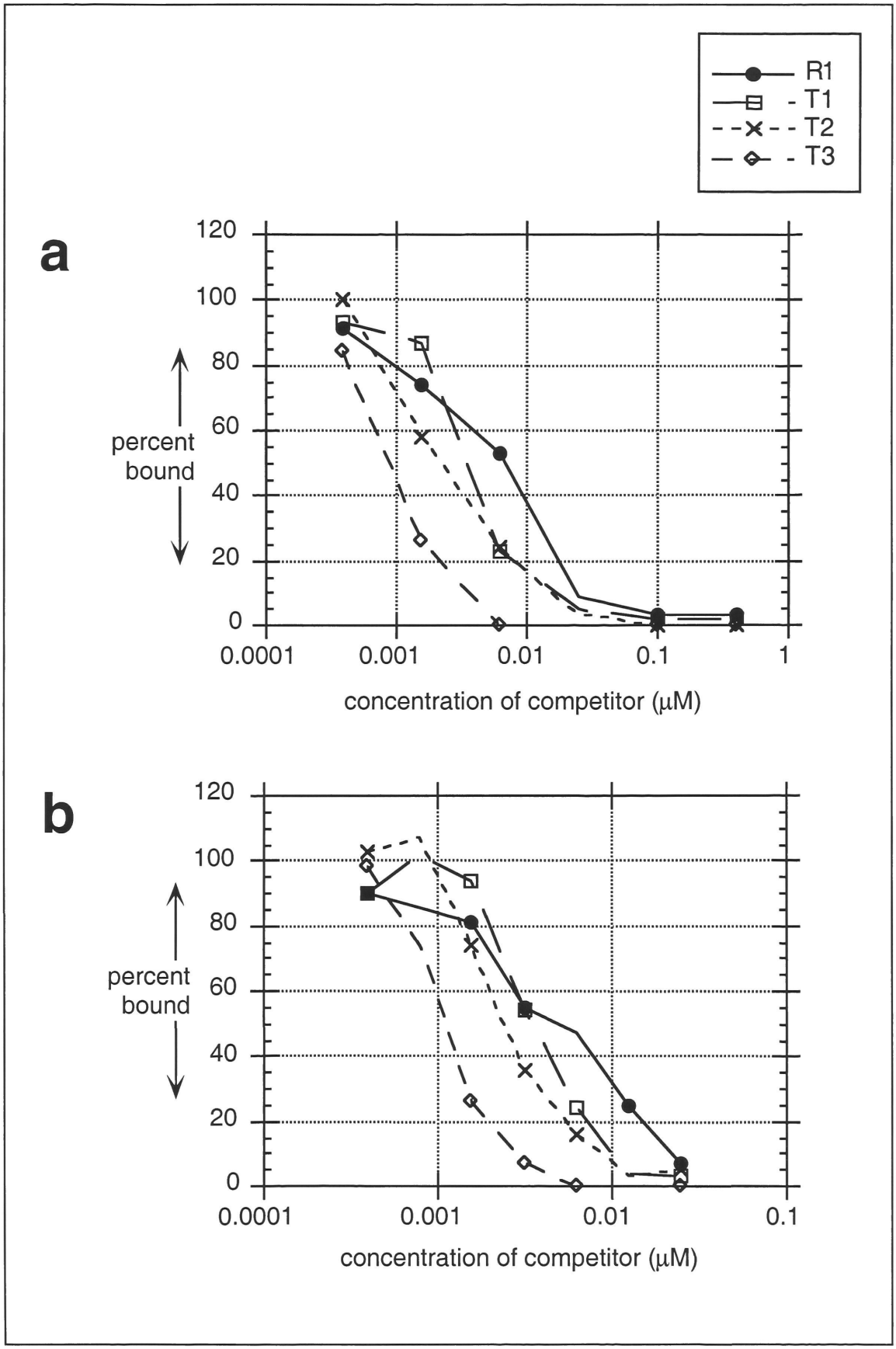

R1
T1
T2
T3
a
120
100
80
60
40
20
0
percent bound
0.0001
0.001
0.01
0.1
1
concentration of competitor (μM)
b
120
100
80
60
40
20
0
percent bound
0.0001
0.001
0.01
0.1
concentration of competitor (μM)

Fig. 2.2

Terminally located G-quartet structures do not prevent Ku from binding to DNA ends

Since single-stranded telomeric sequences are known to adopt G-G base paired structures *in vitro* (see (Henderson 1995) for review, (Sen and Gilbert 1988; Sundquist and Klug 1989; Williamson et al. 1989; Kang et al. 1992)) we next determined whether Ku can bind to DNA molecules terminating in telomeric sequences folded in a G-quartet conformation (G-DNA). To this end we constructed duplex DNA molecules bearing 3.5 copies of the telomeric repeat from *Oxytricha nova* (TTTTGGGG; Fig. 2.1a). Although human telomeric sequences have been shown to fold into G-quartets (Wang and Patel 1993; Balagurumoorthy and Brahmachari 1994; Murchie and Lilley 1994), the *Oxytricha* telomeric repeat was chosen because it is the one that more readily adopts an intramolecular G-quartet conformation *in vitro* (Williamson et al. 1989; Kang et al. 1992; Smith and Feigon 1992; Smith and Feigon 1993; Wang and Patel 1995). Because intramolecular G-quartets are dependent on the presence of K^+ or Na^+, but are not formed in the presence of Li^+ (Williamson et al. 1989), the annealed oligonucleotides were incubated either in the absence of any cation or in the presence of either K^+ or Li^+. When K^+ was present the annealed oligonucleotide displayed an increased migration rate (Fig. 2.3, compare lane 18 with lane 17, dot versus asterisk), consistent with the folding of the two single-stranded tails into a G-quartet conformation (Williamson et al. 1989; Zahler et al. 1991). Although incubation in Li^+ gave sometimes rise to a double band (Fig. 2.3 lane 19, bar), the mobility of the oligonucleotide was slower compared to the mobility observed after incubation in K^+ and was comparable to the mobility of the DNA incubated in TE without ions (Fig. 2.3, compare lanes 17, 18 and 19; also compare lanes 11, 14, 15 with lanes 6, 9, 10 and 1, 4, 5). In addition, when we employed a chemical

Fig. 2.3 **Binding of Ku to DNA molecules terminating in G-DNA**

Labelled Q1 at a concentration of 60 pM was incubated in the presence of 0.06 (lanes 5,10,15), 0.2 (lanes 4, 9, 14), 0.6 (lanes 3, 8, 13) and 1.7 nM Ku (lanes 2, 7, 12). Reaction were in TE (Tris pH 8.0, EDTA 0.1 mM), with no salt (lanes 1-5), 50 mM KCl (lanes 6-10) or 50 mM LiCl (lanes 11-15). No protein and only 3% glycerol was added to samples 16-19 prior to loading on a 5% native polyacrylamide gel. The asterisk refers to the migration rate of the DNA in the absence of salts or in the presence of Li+. The dot indicates the faster mobility of the probe observed in the presence of K+. The bar points out the slow mobility double band sometimes observed in the presence of Li^+. Both bands in gels represent Ku-DNA complexes, containing either one or two Ku heterodimers.

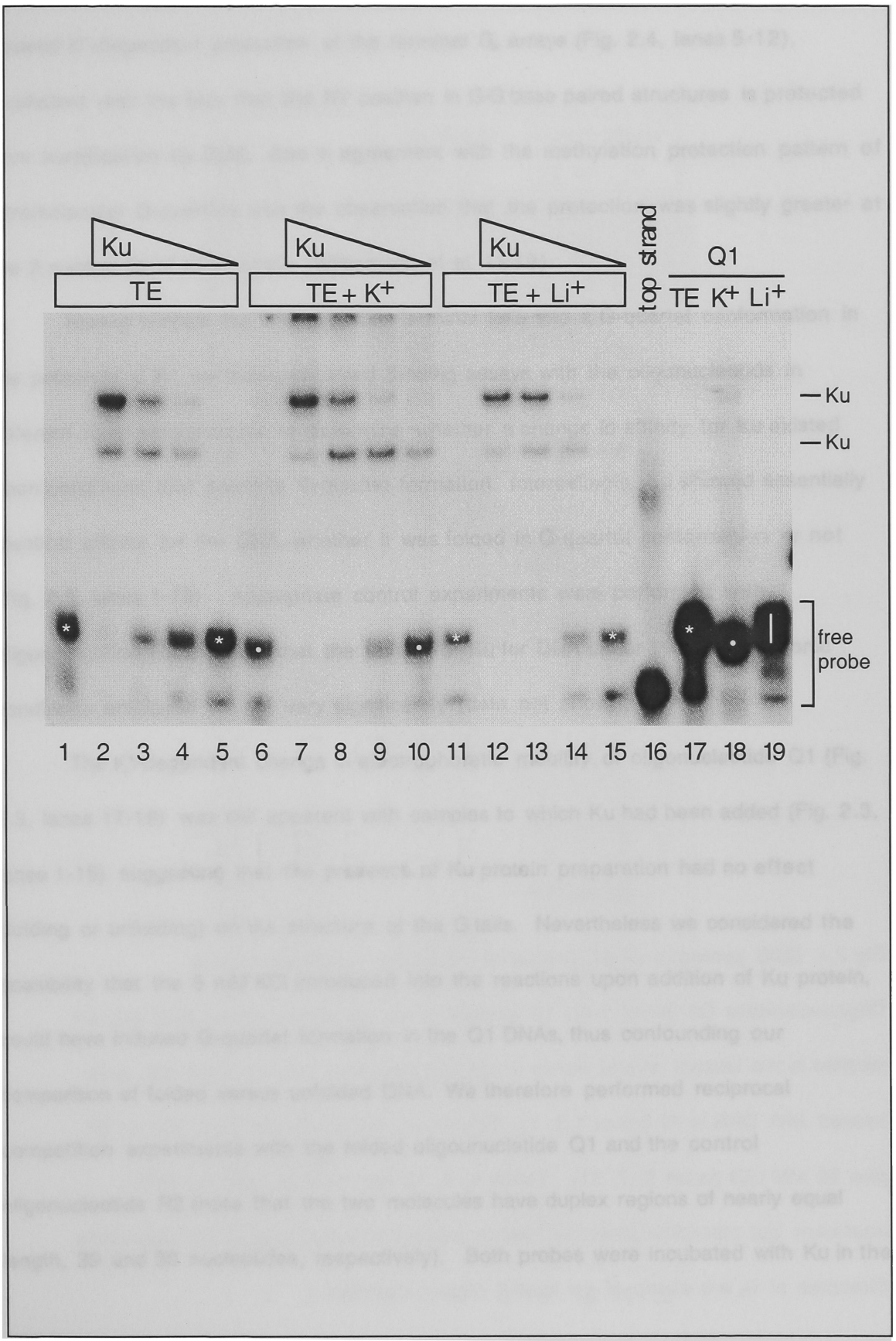

Ku
Ku
Ku
top strand
Q1
TE
TE + K+
TE + Li+
TE K+ Li+
Ku
Ku
free probe
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

Fig. 2.3

protection assay to further verify the structure of the terminal tails, both strands showed K^+-dependent protection of the terminal G_4 arrays (Fig. 2.4, lanes 5-12), consistent with the fact that the N7 position in G-G base paired structures is protected from modification by DMS. Also in agreement with the methylation protection pattern of intramolecular G-quartets was the observation that the protection was slightly greater at the 2 central Gs of each arrays (Williamson et al. 1989).

Having verified the folding of the terminal tails into a G-quartet conformation in the presence of K^+, we then performed binding assays with the oligonucleotide in different ionic environments to determine whether a change in affinity for Ku existed upon conditions that promote G-quartet formation. Interestingly, Ku showed essentially identical affinity for the DNA, whether it was folded in G-quartet conformation or not (Fig. 2.3, lanes 1-15). Appropriate control experiments were performed with oligonucleotide R2 to verify that the affinity of Ku for DNA under the different ionic conditions employed did not vary significantly (data not shown).

The K^+-dependent change in electrophoretic mobility of oligonucleotide Q1 (Fig. 2.3, lanes 17-19) was still apparent with samples to which Ku had been added (Fig. 2.3, lanes 1-15) suggesting that the presence of Ku protein preparation had no effect (folding or unfolding) on the structure of the G tails. Nevertheless we considered the possibility that the 5 mM KCl introduced into the reactions upon addition of Ku protein, could have induced G-quartet formation in the Q1 DNAs, thus confounding our comparison of folded versus unfolded DNA. We therefore performed reciprocal competition experiments with the folded oligounucletide Q1 and the control oligonucleotide R2 (note that the two molecules have duplex regions of nearly equal length, 33 and 30 nucleotides, respectively). Both probes were incubated with Ku in the

Fig 2.4 **DMS protection of G-quartet forming oligonucleotides**

Oligonucleotides Q2 (lanes 1-4), Q1 labelled in the top strand (lanes 5-8) and Q1 labelled in the bottom strand (lanes 9-12) at a concentration of 1.5-4 nM were treated with DMS in TE (lanes 1,5, 9), TE plus 50 mM KCl (lanes 2, 6, 10) or TE plus 50 mM LiCl (lanes 3, 7, 11). Lanes 4, 8, 12 are controls, with no DMS treatment but piperidine cleavage. Bands represent G positions in the sequences. Stretches of G_4 are at top of gel in all 3 oligonucleotides.

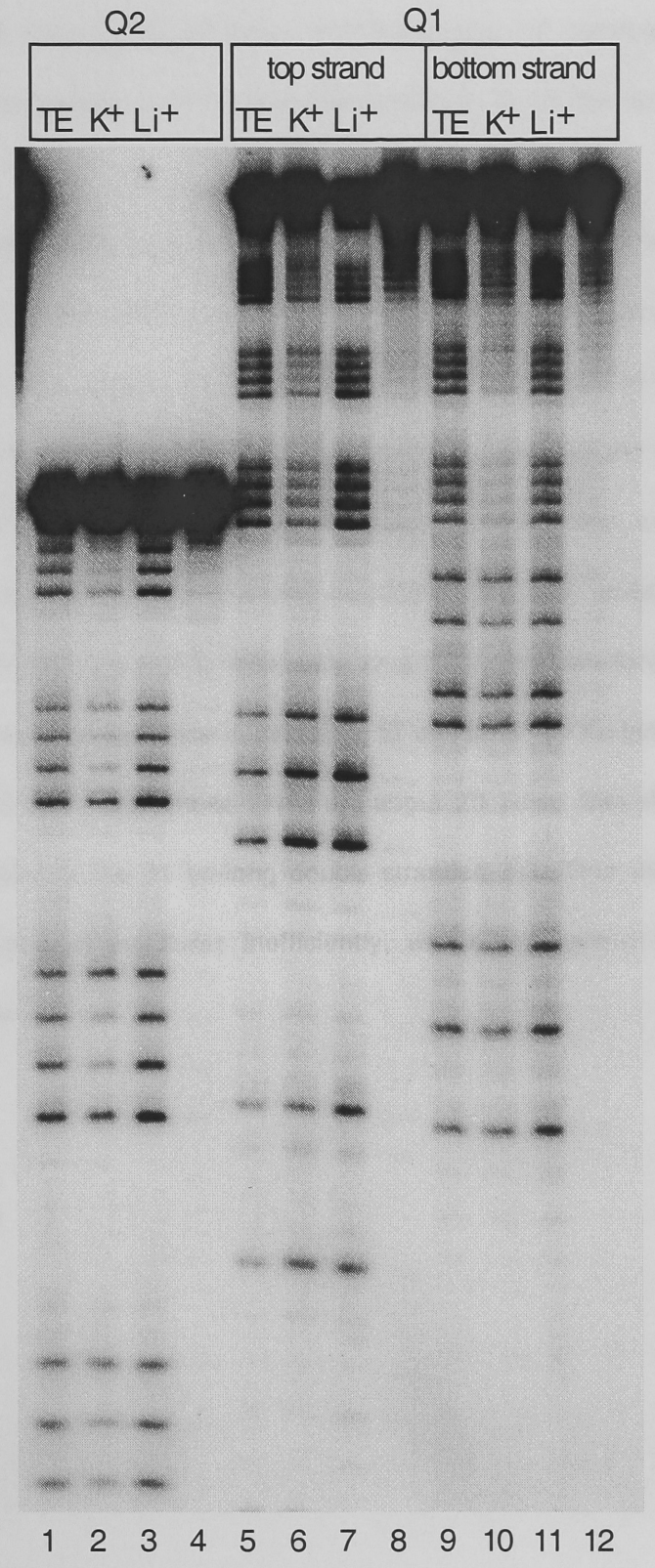

Q2
Q1
top strand
bottom strand
TE K+ Li+
TE K+ Li+
TE K+ Li+
1 2 3 4 5 6 7 8 9 10 11 12

Fig. 2.4

presence of K^+ and a molar excess of either cold competitor DNA (Fig. 2.5a). Both DNAs competed to the same extent for Ku binding indicating that Ku binds to either substrate with very similar affinities, thus supporting our conclusion that terminally located G-quartets do not affect the binding of Ku to the duplex.

Ku binds G-quartets with low affinity compared to double-stranded DNA

To see if Ku has the ability to bind to G-quartet structures *per se* we performed competition experiments with an oligonucleotide entirely composed of four copies of the *Oxytricha* telomeric repeat (Fig. 2.1a; Q2). This molecule displays the expected K^+-dependent ability to adopt an intramolecular G-quartet conformation as judged by gel elelectrophoresis (data not shown) and DMS protection (Fig. 2.4, lanes 1-4). Interestingly, Q2 behaves practically identically to a 33 nucleotides-long single-stranded oligonucleotide of random sequence in its ability to compete for Ku binding (Fig. 2.5b, lanes 10-17, and 19-26). Both these DNAs are about 20 times less effective as competitors compared to the 31 bp-long double-stranded R2. Thus these results show that Ku binds to G-quartet structures inefficiently, with affinity similar to the one for single-stranded DNA.

Fig. 2.5a **Comparison of Ku DNA binding affinities for G-DNA capped DNA and double-stranded DNA**

Labelled oligonucleotides R2 (lanes 16-30) and Q1 (lanes 1-15) were present at 250 pM and the Ku concentration was about 200 pM. Binding reactions were carried out in TE with 50 mM KCl. The amount of competitor (in molar excess) is indicated at the top of each lane.

Fig. 2.5b **Comparison of Ku DNA binding affinities for G-DNA, double-stranded DNA and single-stranded DNA**

Labelled oligonucleotide R2 was present at 250 pM and the Ku concentration was about 200 pM. Binding reactions were carried out in TE with 50 mM KCl.
The amount of competitor (in molar excess) is indicated at the top of each lane.

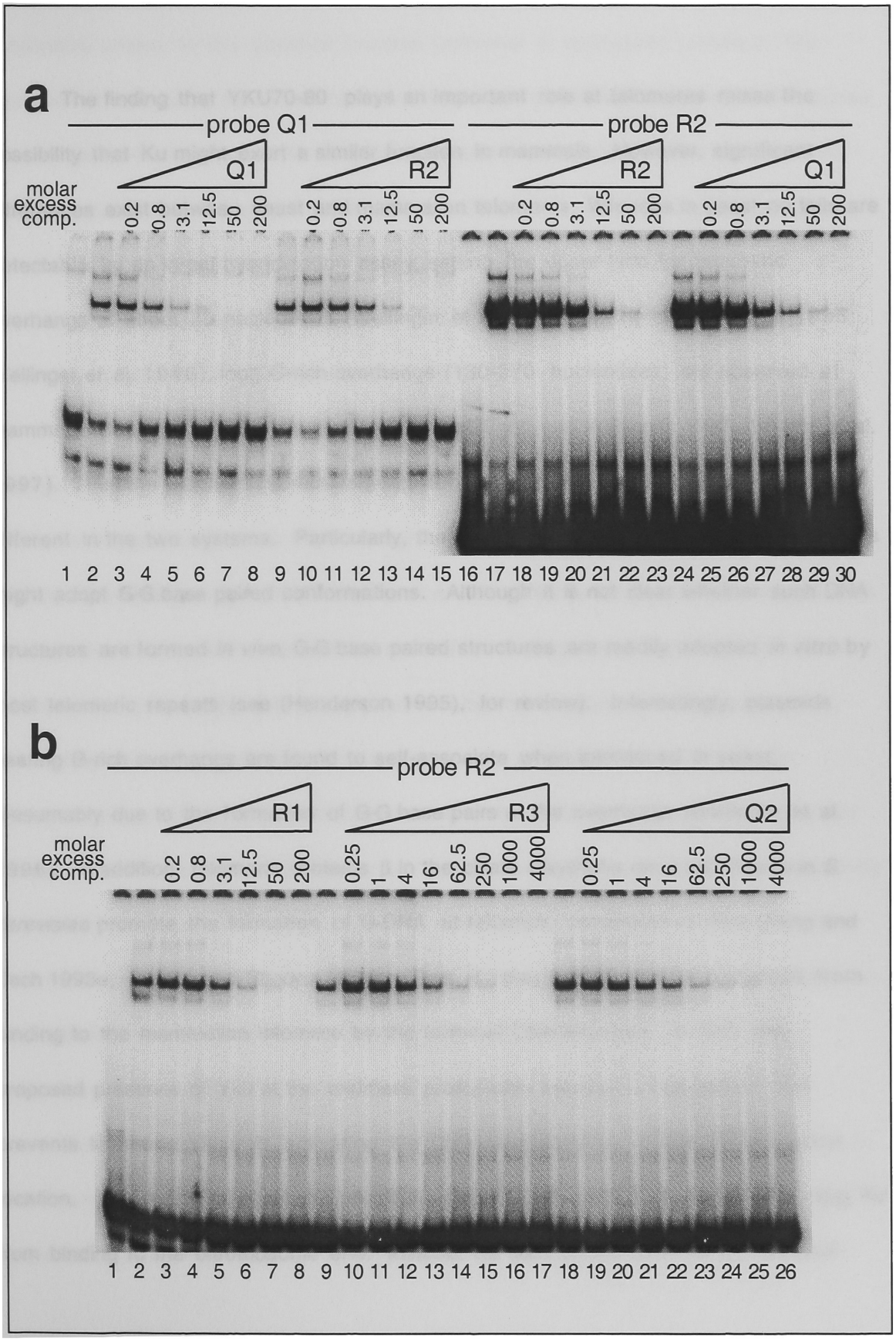

a
probe Q1
probe R2
molar excess comp.
Q1
R2
R2
Q1
0.2
0.8
3.1
12.5
50
200
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
b
probe R2
molar excess comp.
R1
R3
Q2
0.2
0.8
3.1
12.5
50
200
0.25
1
4
16
62.5
250
1000
4000
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26

Fig. 2.5

DISCUSSION

The finding that YKU70-80 plays an important role at telomeres raises the possibility that Ku might exert a similar function in mammals. However, significant differences exist between yeast and mammalian telomeres. Whereas in yeast no tails are detectable by an in-gel hybridization assay, setting the upper limit for telomeric overhangs at about 20 nucleotides (Wellinger et al. 1993; Dionne and Wellinger 1996; Wellinger et al. 1996), long G-rich overhangs (130-270 nucleotides) are observed at mammalian telomeres (Makarov et al. 1997; McElligott and Wellinger 1997; Wright et al. 1997). Thus the biochemical requirements for binding to chromosome ends may be different in the two systems. Particularly, the G-rich tails found at mammalian telomeres might adopt G-G base paired conformations. Although it is not clear whether such DNA structures are formed *in vivo*, G-G base paired structures are readily adopted *in vitro* by most telomeric repeats (see (Henderson 1995), for review). Interestingly, plasmids bearing G-rich overhangs are found to self-associate when introduced in yeast, presumably due to the formation of G-G base pairs at the overhangs (Wellinger et al. 1996). In addition, telomeric proteins β in the ciliate *Oxytricha nova* and Rap1p in *S. cerevisiae* promote the formation of G-DNA at telomeric sequences *in vitro* (Fang and Cech 1993a; Giraldo and Rhodes 1994). Thus, Ku may conceivably be prevented from binding to the mammalian telomere by the terminal DNA structure. In fact, the proposed presence of Yku at the telomere presumably requires a mechanism that prevents telomeric Yku from activating the DSB repair pathway at that chromosomal location. Our results rule out the possibility that G-DNA might function in preventing Ku from binding to the chromosome end. Instead, we have shown that the biochemical

characteristics of Ku are consistent with its presence at mammalian telomeres. Clearly, a definitive answer to this question requires functional or cytological evidence. Our attempts at identifying Ku at telomeres in human and rodent cell lines have not resulted in proof for co-localization of Ku with the telomeric protein TRF1 (A Bianchi and T. de Lange, unpublished). However, these experiments are not conclusive, because a signal at the telomere might be masked by the abundance of Ku throughout the nucleus and because of the expected scarcity of an end factor at the telomere, which might be below the detection limits of immunofluorescence techniques.

The mechanism of Ku binding to DNA is not known. The search for a common requirement in the various DNA substrates that are bound with high affinity by Ku has led to the suggestion that Ku might recognize the transition from single to double-stranded DNA (Falzon et al. 1993; Rathmell and Chu 1994). This conclusion is based on the fact that free 3' or 5' ends are not needed for binding and is supported to some extent by the observation that Ku appears to bind more tightly to oligonucleotides terminating with an AT rich sequence as opposed to ones terminating in a G-rich sequence, suggesting that possibly the melting of the terminal base pairs provides the substrate for Ku recognition (Falzon et al. 1993). In addition, a Ku DNA-helicase activity that could help in generating such substrate has been reported (Tuteja et al. 1994). However, this model for DNA binding is not consistent with our finding that Ku is able to bind with high affinity to inter-strand terminally crosslinked (with psoralen) DNA molecules (A. Bianchi and T. de Lange, unpublished).

In the case of the G-DNA containing substrates used in this work (which all have free 3' and 5' ends) it appears likely that the recognition happens at either the free 3' or 5' end (or in the T-loop region). With recognition occurring at the 3' end or in the T-

loop only, stabilization of Ku binding on the duplex would appear likely to require unfolding of the G-quartet with lower on-rate and affinity as a consequence. Since this is not observed, we favor the hypothesis that it is the 5' free end that is recognized. Even though our experiments do not address the state of the folded overhang after Ku binding, we argue that an unfolding step would be likely to lower the affinity of Ku for the DNA and we therefore suggest that Ku is able to recognize the junction between double-stranded DNA and G-quartet without the necessity to unfold the structure.

Chapter 3

Characterization of the DNA binding activity of the human telomeric proteins TRF1 and TRF2

SUMMARY

The mammalian telomeric proteins TRF1 and TRF2 bind to TTAGGG repeats in double-stranded form *in vitro* and localize to telomeres *in vivo,* both during interphase and in mitosis. Each protein contains one Myb-type repeat at the C-terminus. We have conducted a detailed analysis of the DNA binding mode of TRF1 and shown that two Myb domains in the dimer are required for efficient binding. Analysis of DNA binding sites selected by a PCR-mediated approach, has allowed the determination of the consensus binding site for an isolated Myb domain. Surprisingly, we found that the dimeric protein contacts two of these minimal binding sites in a manner that is not sensitive to their relative orientation and spacing. We suggest that TRF1 contacts these bipartite binding sites by using the two Myb domains in the dimer independently but simultaneously. We propose a model for TRF1 binding that predicts the existence of a flexible region located between the highly conserved dimerization and DNA binding domains, that would allow free relative rotation of the Myb domains. This novel binding mechanism is different from the one employed by the yeast telomeric protein Rap1p and could possibly be of more common nature among telomeric proteins, given the fact that TRF2 and fission yeast Taz1p have a domain organization similar to TRF1.

In addition, we describe experiments that suggest that TRF1 may impose a higher order structure on telomeric DNA.

Finally, we report on a preliminary characterization of the DNA binding activity of TRF2.

INTRODUCTION

TRF1 and TRF2 are the only known protein components of mammalian telomeres. Both proteins appear to be localized at all telomeres during interphase and in mitosis (Chong et al. 1995; Ludérus et al. 1996; Bilaud et al. 1997; Broccoli et al. 1997b; van Steensel and de Lange 1997; van Steensel et al. 1998). Human TRF1 was first identified in nuclear extracts from HeLa cells as a low-abundance activity capable of binding in a specific manner to arrays of TTAGGG repeats *in vitro* (Zhong et al. 1992). TRF1 mRNA is present in all tissues (Chong et al. 1995) and homologs have been isolated from mouse and chicken ((Broccoli et al. 1997a) and C. Price, D. Broccoli and T. de Lange, unpublished). TRF2 was identified as a TRF1 homologue in the database (Bilaud et al. 1996; Bilaud et al. 1997; Broccoli et al. 1997b). Both TRF1 and TRF2 act as negative regulators of telomere length, possibly by inhibiting the interaction of telomerase with telomere termini ((van Steensel and de Lange 1997) and A. Smogorzewska, B. van Steensel and T. de Lange, unpublished). Similarly, duplex telomeric DNA-binding activities in budding and fission yeast have been implicated as negative regulators of telomere length (McEachern and Blackburn 1995; Krauskopf and Blackburn 1996; Cooper et al. 1997; Marcand et al. 1997). In addition, TRF2 has a role in stabilizing G-strand overhangs and preventing end-to-end chromosome fusions (van Steensel et al. 1998).

In vitro, TRF1 binds to arrays of duplex TTAGGG repeats, irrespective of the presence of a DNA terminus (Zhong et al. 1992). Single-stranded DNA is not an effective TRF1 substrate and neither are heterologous telomeric sequences, such as double-stranded arrays of TTGGGG, TTAGGC, TTTAGGG, TTAGGGGG and TAGGG repeats

(Zhong et al. 1992; Hanish et al. 1994; Chong et al. 1995). This sequence specificity of TRF1 matches the sequence requirements for *de novo* telomere formation in human cells, suggesting that the protein is involved in this process (Hanish et al. 1994).

Comparison of the human and mouse TRF1 and TRF2 sequences has allowed the identification of several domains in the proteins, based on their degree of conservation (Fig. 3.1a) (Bilaud et al. 1997; Broccoli et al. 1997a; Broccoli et al. 1997b). TRF1 and TRF2 bear two regions of relative high homology. The first is located at the C-terminus and consists of one Myb-type repeat, which is responsible for making the contacts with DNA (see below). The second conserved region starts near the N-terminus and extends to amino acids 245-264. This domain is unique to TRF proteins and is responsible for homodimerization (Bianchi et al. 1997). The region between the Myb and dimerization domains is very poorly conserved but in all cases contains putative nuclear localization signals. TRF1 and TRF2 differ dramatically at the N-terminal domain, which is acidic in TRF1 and basic in TRF2, pointing to different functional roles for the two proteins.

The Myb–related sequence found at the C-terminus of the TRF proteins takes its name from having been first identified in c-Myb, a transcripitional activator with oncogenic potential, and encodes a DNA binding domain of the helix-turn-helix (HTH) type. The presence of one or more tandem copies of these repeats defines a large class of Myb-related proteins that generally function in the regulation of cell growth and differentiation. Members of this class have been discovered in mammals, *Drosophila*, yeast, *Dictyostelium* and appear to be particularly abundant in plants (see (Lipsick 1996) for a general review, and (Martin and Paz-Ares 1997), for a review of plant Myb proteins).

Fig. 3.1a **Domain organization of TRF1 and TRF2**

From (Broccoli et al 1997a). Percent identities between mouse and human TRF1 and TRF2, and between human TRF1 and TRF2, are indicated.

Fig. 3.1b **Sequence alignment of the Myb-repeats of telomeric proteins from yeast and mammals, and of Tbf1p**

Adapted from (Konig and Rhodes 1997). The sequences of the Myb repeats from all Myb-related telomeric proteins, plus the Myb-repeat from the Tbf1 gene from *S.cerevisiae* (ScTbf1), and its *S.pombe* homolog (SpTbf1), are shown. The single Myb-repeats from the mammalian TRFs are shown, as well as both the repeats from both *S.cerevisiae* and its close relative *K.lactis*. Only the crystal structure of Rap1p from *S.cerevisiae* is known, all other helix assignements are inferred from homology.

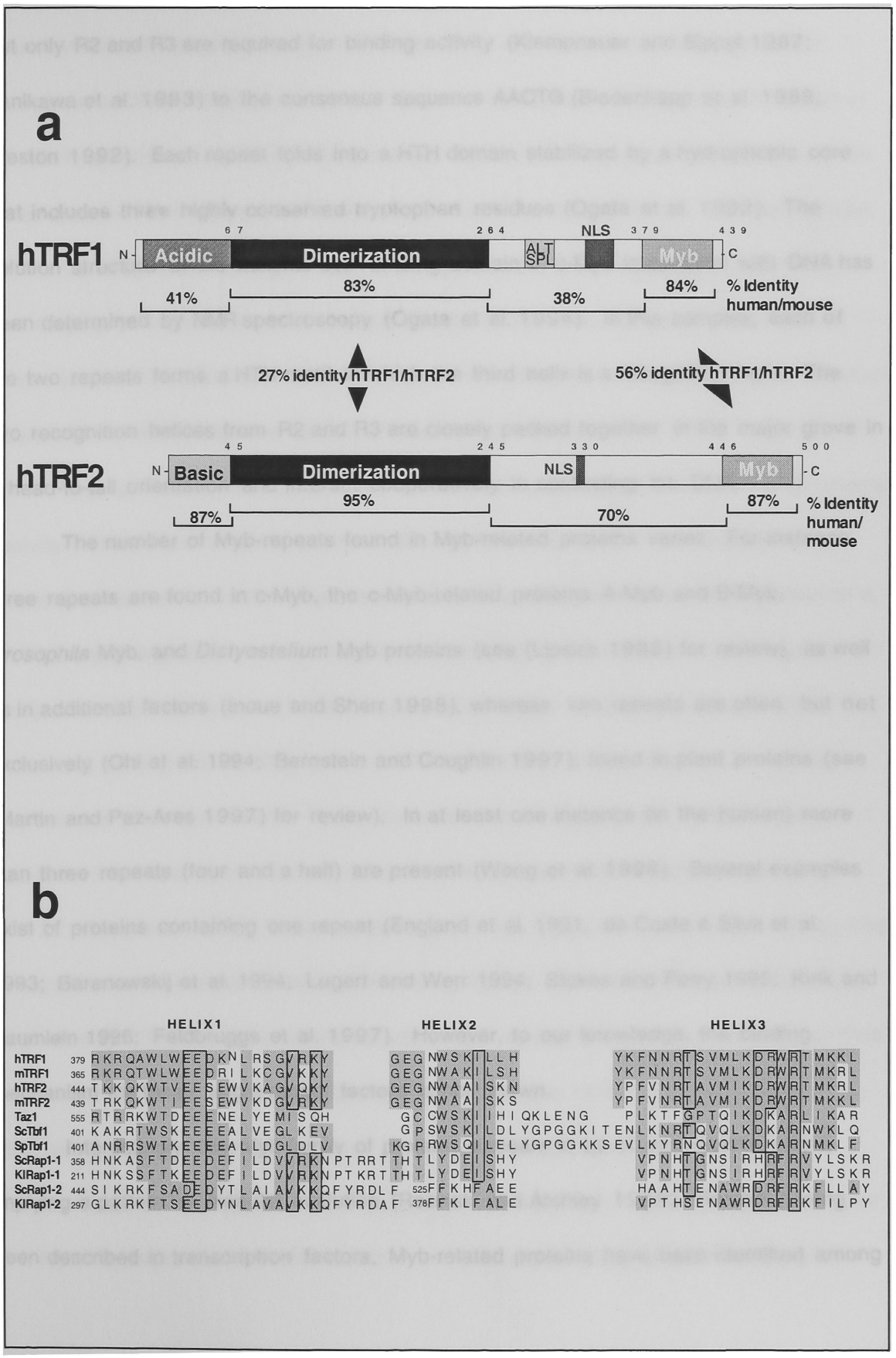

a
hTRF1
N
Acidic
Dimerization
ALT SPL
NLS
Myb
C
67
264
379
439
41%
83%
38%
84%
% Identity human/mouse
27% identity hTRF1/hTRF2
56% identity hTRF1/hTRF2
hTRF2
N
Basic
Dimerization
NLS
Myb
C
45
245
330
446
500
87%
95%
70%
87%
% Identity human/ mouse
b
HELIX1
HELIX2
HELIX3
hTRF1 379 RKRQAWLWEEDKNLRSGVRKY GEGNWSKILLH YKFNNRTSVMLKDRWRTMKKL
mTRF1 365 RKRQTWLWEEDRILKCGVKKY GEGNWAKILSH YKFNNRTSVMLKDRWRTMKRL
hTRF2 444 TKKQKWTVEESEWVKAGVQKY GEGNWAAISKN YPFVNRTAVMIKDRWRTMKRL
mTRF2 439 TRKQKWTIEESEWVKDGVRKY GEGNWAAISKS YPFVNRTAVMIKDRWRTMKKL
Taz1 555 RTRRKWTDEEEENELYEMISQH GCCWSKIIHIQKLENG PLKTFGPTQIKDKARLIKAR
ScTbf1 401 KAKRTWSKEEEEALVEGLKEV GPSWSKILDLYGPGGKITENLKNRTQVQLKDKARNWKLQ
SpTbf1 401 ANRRSWTKEEEEALLDGLDLV KGPRWSQILELYGPGGKKSEVLKYRNQVQLKDKARNMKLF
ScRap1-1 358 HNKASFTDEEDEFILDVVRKNPTRRTTHTLYDEISHY VPNHTGNSIRHRFRVYLSKR
KlRap1-1 211 HNKSSFTKEEDEFILDVVRKNPTKRTTHTLYDEISHY VPNHTGNSIRHRFRVYLSKR
ScRap1-2 444 SIKRKFSADEDYTLAIAVKKQFYRDLF 525FFKHFAEE HAAHTENAWRDRFRKFLLAY
KlRap1-2 297 GLKRKFTSEEDYNLAVAVKKQFYRDAF 378FFKLFALE VPTHSENAWRDRFRKFILPY

Fig. 3.1

In c-Myb the DNA binding domain is composed of three Myb-repeats, R1-3, but only R2 and R3 are required for binding activity (Klempnauer and Sippel 1987; Tanikawa et al. 1993) to the consensus sequence AACTG (Biedenkapp et al. 1988; Weston 1992). Each repeat folds into a HTH domain stabilized by a hydrophobic core that includes three highly conserved tryptophan residues (Ogata et al. 1992). The solution structure of the minimal DNA binding domain of c-Myb in complex with DNA has been determined by NMR spectroscopy (Ogata et al. 1994). In this complex, each of the two repeats forms a HTH motif in which the third helix is a recognition helix. The two recognition helices from R2 and R3 are closely packed together in the major grove in a head-to-tail orientation and interact cooperatively in contacting the DNA.

The number of Myb-repeats found in Myb-related proteins varies. For instance, three repeats are found in c-Myb, the c-Myb-related proteins A-Myb and B-Myb, *Drosophila* Myb, and *Dictyostelium* Myb proteins (see (Lipsick 1996) for review), as well as in additional factors (Inoue and Sherr 1998), whereas two repeats are often, but not exclusively (Ohi et al. 1994; Bernstein and Coughlin 1997), found in plant proteins (see (Martin and Paz-Ares 1997) for review). In at least one instance (in the human) more than three repeats (four and a half) are present (Wong et al. 1998). Several examples exist of proteins containing one repeat (England et al. 1991; da Costa e Silva et al. 1993; Baranowskij et al. 1994; Lugert and Werr 1994; Stokes and Perry 1995; Kirik and Baumlein 1996; Feldbrugge et al. 1997). However, to our knowledge, the binding mechanism for these single-repeat factors is not known.

Interestingly, the Myb family of proteins appears to be of polyphyletic origin, implying a pattern of modular evolution (Rosinski and Atchley 1998). Besides having been described in transcription factors, Myb-related proteins have been identified among

telomere binding factors. In fact, all the 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; Konig et al. 1996; Bilaud et al. 1997; Broccoli et al. 1997b; Cooper et al. 1997), with the possible exception of the yeast protein Tel2p (Runge and Zakian 1996; Kota and Runge 1998). The Myb repeats of TRF1 and TRF2 are closely related to the repeat of Taz1p (Fig. 3.1b). Also highly homologous to these repeats is the Myb domain of the essential yeast gene TBF1, whose product can bind to TTAGGG repeats *in vitro* but has no known telomeric function (Liu and Tye 1991; Brigati et al. 1993; Bilaud et al. 1996).

Rap1p, from *S.cerevisiae*, is the best characterized duplex telomeric DNA binding activity. The protein binds to telomeric DNA both *in vitro* and *in vivo* (Berman et al. 1986; Buchman et al. 1988a; Buchman et al. 1988b; Longtine et al. 1989; Conrad et al. 1990), and it regulates telomere length by recruiting to the telomere a critical mass of Rif1p and Rif2p (Hardy et al. 1992; Marcand et al. 1997; Wotton and Shore 1997). The consensus binding site for Rap1p is 5'-GGTGTGTGGGTGT-3' (Buchman et al. 1988a; Graham and Chambers 1994). Rap1p sites occur on average every 18 bp on yeast telomeric DNA, and all of them can be simultaneously occupied *in vitro* (Gilson et al. 1993). The protein is 827 amino acids long, but the region from amino acid 361 to 596 is sufficient to define the minimal DNA-binding domain (DBD) (Henry et al. 1990; Gilson et al. 1993). This domain contains two Myb-type repeats (Larson et al. 1994) that fold into two HTH motifs, as revealed by the crystal structure of the complex between the DBD and the consensus site (Konig et al. 1996). The arrangement of the two Myb-repeats in the tertiary structure of Rap1p is entirely different from the one encountered in c-Myb: in Rap1p the two HTH motifs 1 and 2 bind separately to two tandemly

repeated elements (5'-GGGTGT-3' and 5'-GGTGT-3', respectively). Therefore, in Rap1p two intramolecularly and tandemly arranged Myb-motifs contact two tandemly arrayed sites on the DNA. Interestingly, the HTH motifs of Rap1p present a similarity to the structure of homeodomains in that they have N-teminal arms that make contacts with DNA in the minor groove, in addition to the major groove contacts of their recognition helices (Gehring et al. 1994; Konig et al. 1996; Konig and Rhodes 1997). The presence of the N-teminal arm in the Rap1p HTH domains allows each of these motifs to recognize a 5-6 bp site, whereas in c-Myb two HTH motifs are needed to specify contacts to a single pentameric site. The presence of a similar N-terminal arm in the TRF1 Myb-motif has been proposed, based on its amino acid sequence (Konig et al. 1998).

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.

We have undertaken a biochemical characterization of the binding mode of TRF1 and have addressed two main issues. First, given the fact that the known examples of Myb-protein structures reveal that two repeats are used in binding to DNA, we have investigated how single-repeat Myb-proteins in general, and TRF1 in particular, may achieve efficient DNA-binding. Second, to better understand how the human telomeric complex is assembled, we have investigated the precise nature of the TRF1 site on the

DNA. In addition we have obtained evidence for a higher level of organization of TRF1-bound DNA that may have profound implications for telomere function. A preliminary analysis of the binding of TRF2 to DNA is also included and discussed.

RESULTS

TRF1 binds to DNA as a dimer

Since TRF1 bears only one Myb-repeat, we addressed the possibility that the protein, like Rap1p and c-Myb, might use two Myb domains for DNA binding, bringing them together by homotypic dimerization. To test this idea, we constructed a fusion protein consisting of green fluorescent protein (GFP, MW of 26 kD) attached to the N-terminus of full-length TRF1. Both the fusion protein and wild type TRF1 were expressed *in vitro* by coupled transcription/translation in cell lysates derived from rabbit reticulocytes. This reaction gave rise to protein products of the expected molecular weight (Fig. 3.2b, lanes 2 and 3). Both proteins generated complexes with telomeric DNA in a band-shift assay and, as expected, the fusion protein produced a higher mobility complex (Fig. 3.2a, lane 2) compared to the wild type TRF1 (Fig. 3.2a, lane 3). When the two TRF1 derivatives where co-expressed (Fig. 3.2b, lanes 4 and 5) and assayed by band-shift, the two protein-DNA complexes were still apparent and, in addition, a novel complex of intermediate mobility was observed (Fig. 3.2a, lanes 4 and 5). Since the migration of the complexes obtained with either TRF1 derivative appears to be inversely proportional to mass, the simplest interpretation for the appearance of the new complex is that it contains intermediate protein mass, and therefore both polypeptides (Hope and Struhl 1987). In this view the slowest migrating complex is made of homodimers of the fusion protein, and the fastest migrating complex is made of homodimers of wild type TRF1.

In these band-shift experiments we have never observed a complex that could represent TRF1 monomers. Furthermore, when assaying co-expressed variants of TRF1

Fig. 3.2 **TRF1 binds to DNA as a dimer**

(a) shows a band-shift analysis of *in vitro* translated proteins. Expression of TRF1 derivatives from rabbit reticulocyte lysates was achieved in the presence of [^{35}S]methionine. The deduced protein composition of the protein-DNA complex is indicated at the left of the gel. The proteins present in each binding reaction are shown directly below each lane, as resolved by SDS-PAGE (b).

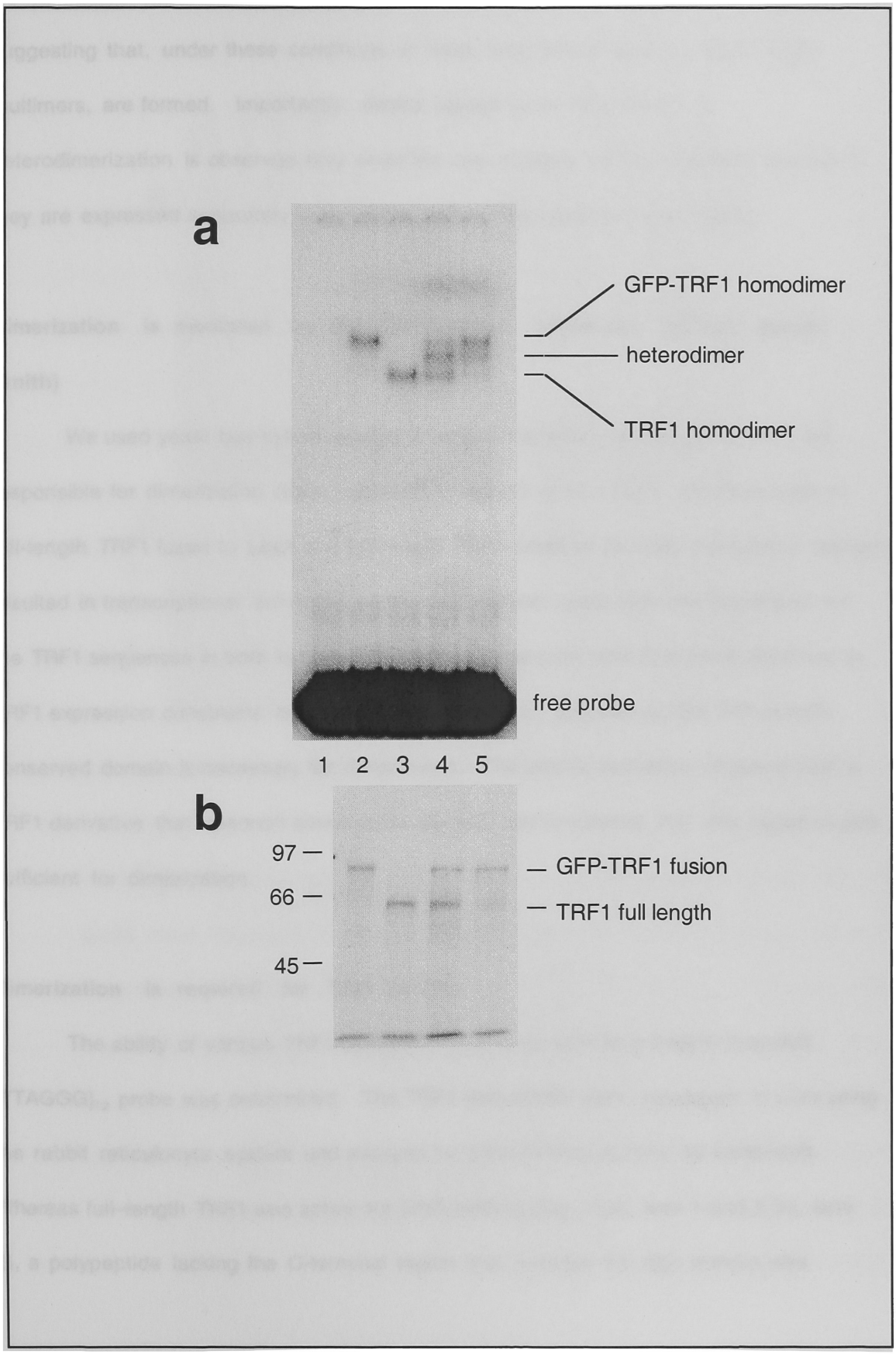

a
GFP-TRF1 homodimer
heterodimer
TRF1 homodimer
free probe
1 2 3 4 5
b
97
66
45
GFP-TRF1 fusion
TRF1 full length

Fig. 3.2

we have never observed the appearance of more than one band of intermediate mobility, suggesting that, under these conditions at least, only dimers, and not higher order multimers, are formed. Importantly, dimers appear to be very stable, as heterodimerization is observed only when the two proteins are co-translated and not if they are expressed separately and subsequently mixed (Bianchi et al. 1997).

Dimerization is mediated by the TRF-specific conserved domain (Susan Smith)

We used yeast two-hybrid assays to determine which sequences in TRF1 are responsible for dimerization (data published in Bianchi et al. 1997). Co-expression of full-length TRF1 fused to LexA and full-length TRF1 fused to the GAL4 activation domain resulted in transcriptional activation of the *lacZ* reporter gene that was dependent on the TRF1 sequences in both hybrids. A series of N-terminal and C-terminal deletions in TRF1 expression constructs determined that the region spanned by the TRF-specific conserved domain is necessary for dimerizaton. The strong activation obtained with a TRF1 derivative that spanned amino acids 66-263, demonstrated that this region is also sufficient for dimerization.

Dimerization is required for DNA binding

The ability of various TRF1 deletion mutants to bind to a double-stranded $[TTAGGG]_{12}$ probe was determined. The TRF1 derivatives were expressed *in vitro* using the rabbit reticulocyte system and assayed for DNA binding activity by band-shift. Whereas full–length TRF1 was active for DNA binding (Fig. 3.3a, lane 1 and 3.3d, lane 2), a polypeptide lacking the C-terminal region that includes the Myb domain was

Fig. 3.3 **Deletion mapping of TRF1 sequences required for DNA binding activity**

TRF1 derivatives produced in the rabbit reticulocyte system and examined by band-shift assays for DNA binding activity are shown in (f) in diagram form, along with a summary of the results of the DNA binding assays. Band-shift assays are shown in (a) and (d). The proteins assayed are shown in SDS-PAGE gels in (b), (c) and (e).

Proteins used in band-shift assays were produced with cold amino acids, while proteins produced for visualization in SDS-PAGE gels were labeled with [^{35}S]methionine.

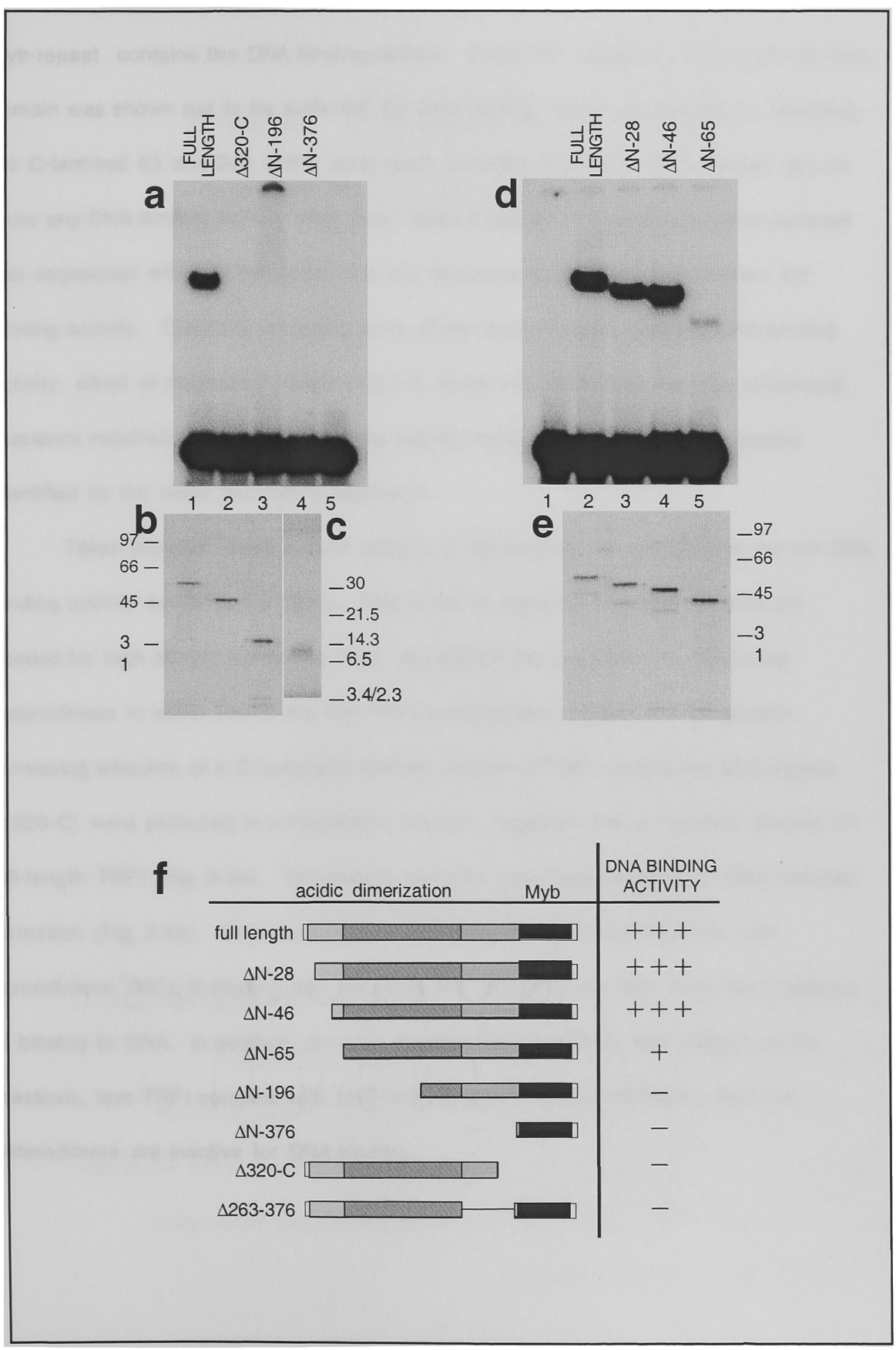

a
FULL LENGTH
Δ320-C
ΔN-196
ΔN-376
1 2 3 4 5
b
97
66
45
3
1
c
30
21.5
14.3
6.5
3.4/2.3
d
FULL LENGTH
ΔN-28
ΔN-46
ΔN-65
1 2 3 4 5
e
97
66
45
3
1
f
acidic dimerization
Myb
DNA BINDING ACTIVITY
full length
ΔN-28
ΔN-46
ΔN-65
ΔN-196
ΔN-376
Δ320-C
Δ263-376
+++
+++
+++
+
−
−
−
−

Fig. 3.3

defective for DNA binding activity (Fig. 3.3a, lane 2), consistent with the fact that the Myb-repeat contains the DNA binding domain. Under the conditions employed, the Myb domain was shown not to be sufficient for DNA binding since two constructs spanning the C-terminal 63 and 243 amino acids (both including the entire Myb-domain) did not show any DNA binding activity (Fig. 3.3a, lanes 3 and 4). These results demonstrated that sequences either in the acidic or in the dimerization domain are important for binding activity. Deletions removing parts of the acidic domain retained DNA binding activity, albeit at diminished levels (Fig.3.3, lanes 3-5), indicating that the N-terminal sequence requirements for DNA binding activity reside in the dimerization domain identified by the yeast two-hybrid approach.

Taken together these results point to a requirement for dimerization for full DNA binding activity by TRF1 (Fig. 3.3f). This might be because two Myb domains are needed for high affinity binding to DNA. We tested this possibility by producing heterodimers in which one of the two TRF1 polypeptides lacked the Myb domain. Increasing amounts of a C-terminally deleted version of TRF1 lacking the Myb–repeat (Δ320-C) were produced in a translation reaction together with a constant amount of full-length TRF1 (Fig. 3.4b). The protein mixtures were then assayed for DNA complex formation (Fig. 3.4a). Only one complex was observed and it co-migrated with homodimeric TRF1, indicating that heterodimers with only one Myb motif are incapable of binding to DNA. In addition, as more deleted TRF1 (Δ320-C) was present in the reactions, less TRF1 complex with DNA was observed, further indicating that the heterodimers are inactive for DNA binding.

Fig. 3.4 **TRF1 needs two Myb domains in order to bind DNA efficiently** Full-length TRF1 and a TRF1 truncated allele lacking the Myb domain were expressed *in vitro* using a rabbit reticulocyte lysate. [^{35}S]methionine-labelled protein products are shown in an SDS-PAGE gel depicted in (b). (a) shows band-shift assays with the protein mixtures shown immediately below each corresponding lane in the SDS-PAGE gel.

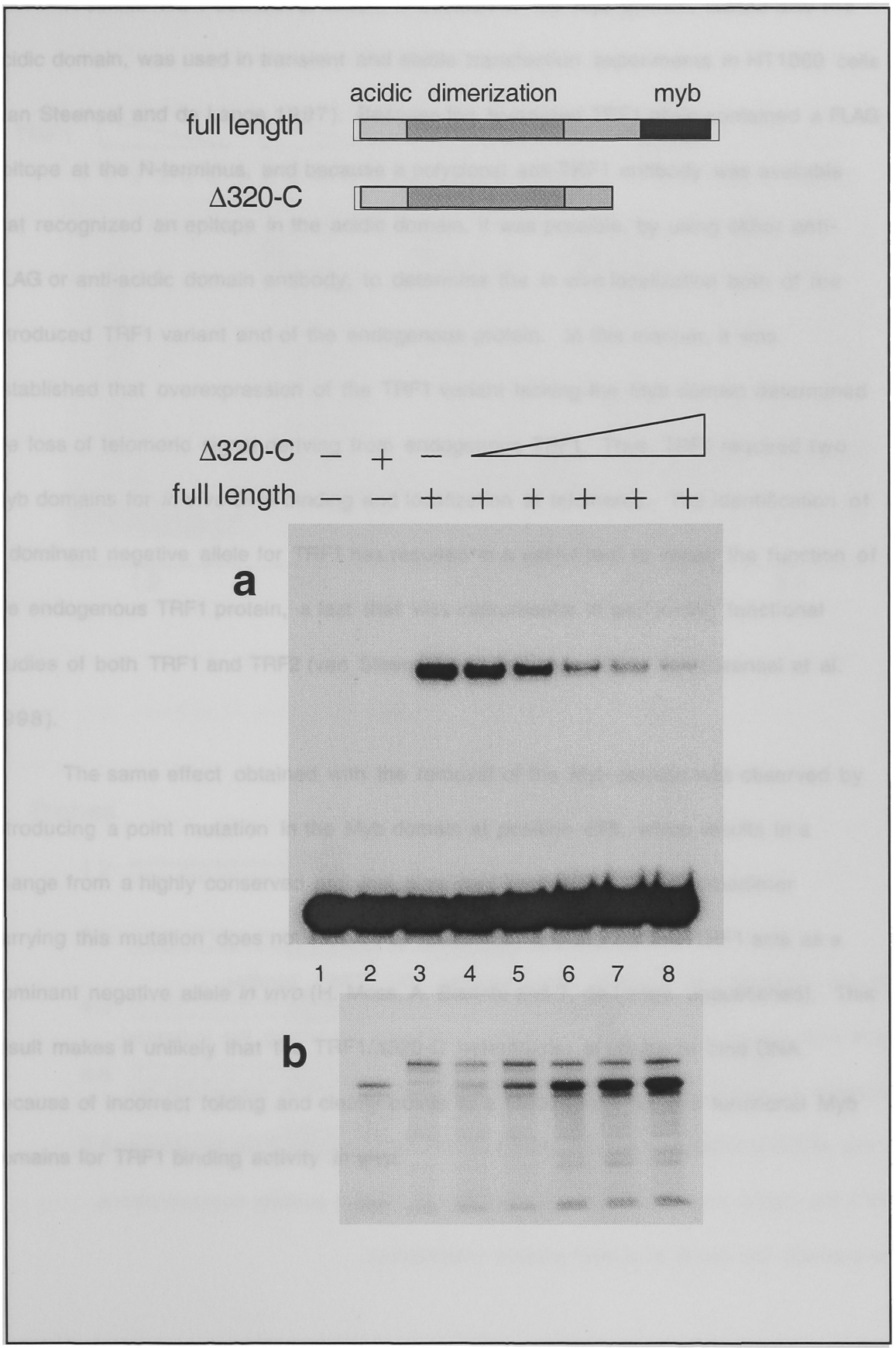

acidic
dimerization
myb
full length
Δ320-C
Δ320-C – + –
full length – – + + + + + +
a
1 2 3 4 5 6 7 8
b

Fig. 3.4

A similar TRF1 derivative, which, in addition to the Myb domain, lacked also the acidic domain, was used in transient and stable transfection experiments in HT1080 cells (van Steensel and de Lange 1997). Because the truncated TRF1 allele contained a FLAG epitope at the N-terminus, and because a polyclonal anti-TRF1 antibody was available that recognized an epitope in the acidic domain, it was possible, by using either anti-FLAG or anti-acidic domain antibody, to determine the *in vivo* localization both of the introduced TRF1 variant and of the endogenous protein. In this manner, it was established that overexpression of the TRF1 variant lacking the Myb domain determined the loss of telomeric signal deriving from endogenous TRF1. Thus, TRF1 required two Myb domains for *in vivo* DNA binding and localization at telomeres. The identification of a dominant negative allele for TRF1 has resulted in a useful tool to impair the function of the endogenous TRF1 protein, a fact that was instrumental in performing functional studies of both TRF1 and TRF2 (van Steensel and de Lange 1997; van Steensel et al. 1998).

The same effect obtained with the removal of the Myb domain was observed by introducing a point mutation in the Myb domain at position 425, which results in a change from a highly conserved arginine to valine (Fig. 3.1). A TRF1 homodimer carrying this mutation does not localize to telomeres, and the mutant TRF1 acts as a dominant negative allele *in vivo* (H. Moss, A. Bianchi and T. de Lange, unpublished). This result makes it unlikely that the TRF1/Δ320-C heterodimer is unable to bind DNA because of incorrect folding and clearly points to a requirement for two functional Myb domains for TRF1 binding activity *in vivo*.

Fig. 3.5 **Analysis of spacing requirements for TRF1 binding**

Band-shift assays with TRF1 on restriction fragments containing the indicated telomeric sequences were performed and analyzed on a native acrylamide gel. In this as in subsequent band-shift experiments with pure baculovirus-expressed TRF1, the higher mobility complexes observed at highest protein concentrations are probably the result of protein-protein interactions.

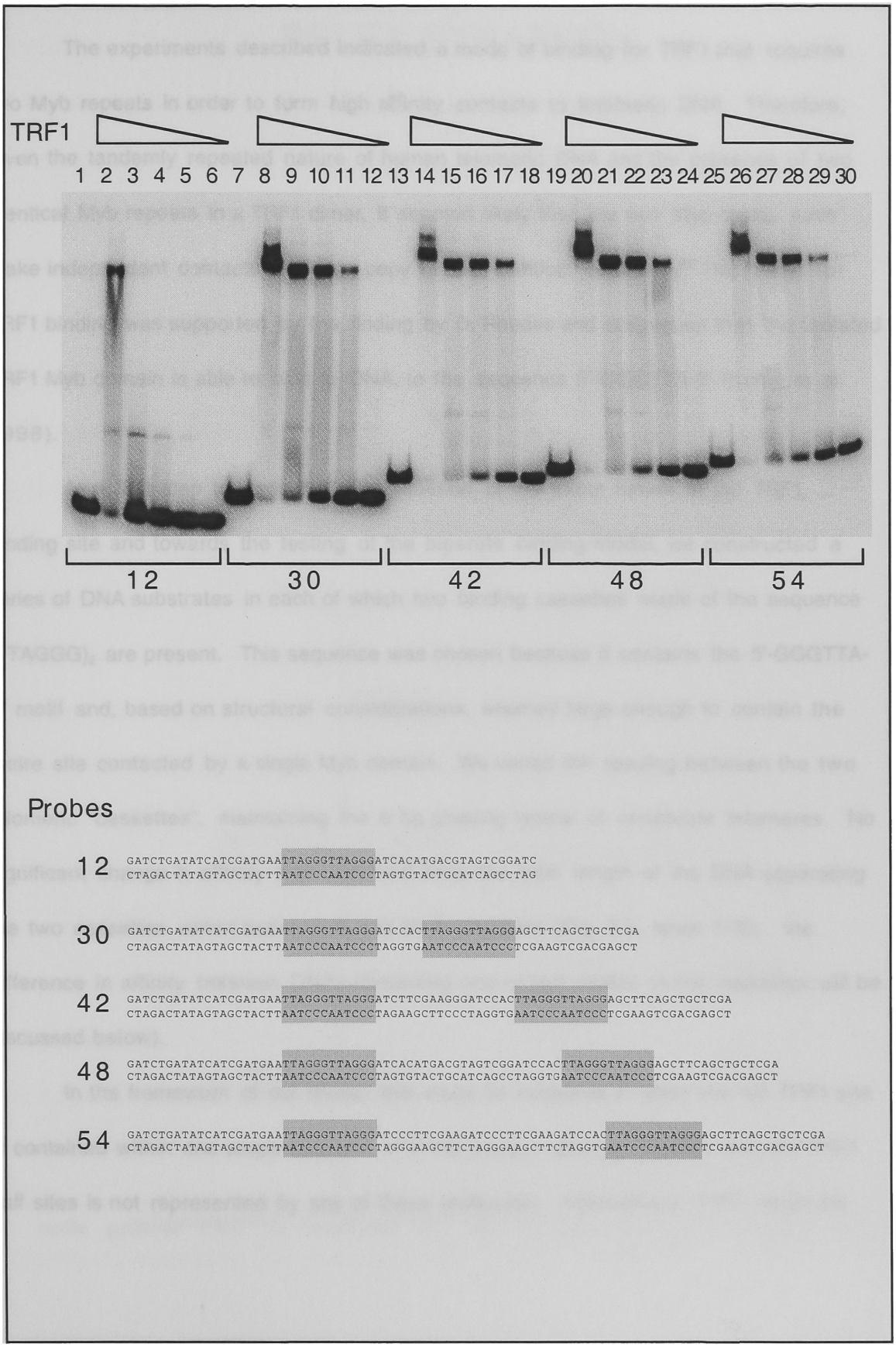

TRF1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
12
30
42
48
54
Probes
12
GATCTGATATCATCGATGAATTAGGGTTAGGGATCACATGACGTAGTCGGATC
CTAGACTATAGTAGCTACTTAATCCCAATCCCTAGTGTACTGCATCAGCCTAG
30
GATCTGATATCATCGATGAATTAGGGTTAGGGATCCACTTAGGGTTAGGGAGCTTCAGCTGCTCGA
CTAGACTATAGTAGCTACTTAATCCCAATCCCTAGGTGAATCCCAATCCCTCGAAGTCGACGAGCT
42
GATCTGATATCATCGATGAATTAGGGTTAGGGATCTTCGAAGGGATCCACTTAGGGTTAGGGAGCTTCAGCTGCTCGA
CTAGACTATAGTAGCTACTTAATCCCAATCCCTAGAAGCTTCCCTAGGTGAATCCCAATCCCTCGAAGTCGACGAGCT
48
GATCTGATATCATCGATGAATTAGGGTTAGGGATCACATGACGTAGTCGGATCCACTTAGGGTTAGGGAGCTTCAGCTGCTCGA
CTAGACTATAGTAGCTACTTAATCCCAATCCCTAGTGTACTGCATCAGCCTAGGTGAATCCCAATCCCTCGAAGTCGACGAGCT
54
GATCTGATATCATCGATGAATTAGGGTTAGGGATCCCTTCGAAGATCCCTTCGAAGATCCACTTAGGGTTAGGGAGCTTCAGCTGCTCGA
CTAGACTATAGTAGCTACTTAATCCCAATCCCTAGGGAAGCTTCTAGGGAAGCTTCTAGGTGAATCCCAATCCCTCGAAGTCGACGAGCT

Fig. 3.5

The size of the TRF1 binding site

The experiments described indicated a mode of binding for TRF1 that requires two Myb repeats in order to form high affinity contacts to telomeric DNA. Therefore, given the tandemly repeated nature of human telomeric DNA and the presence of two identical Myb repeats in a TRF1 dimer, it seemed likely that the two Myb motifs each make independent contacts with one copy of two identical DNA sites. This model for TRF1 binding was supported by the finding by D. Rhodes and colleagues that the isolated TRF1 Myb domain is able to bind to DNA, to the sequence 5'-GGGTTA-3' (Konig et al. 1998).

As a first step towards the determination of the exact nature of the TRF1 binding site and towards the testing of the bipartite binding model, we constructed a series of DNA substrates in each of which two binding cassettes made of the sequence $(TTAGGG)_2$ are present. This sequence was chosen because it contains the 5'-GGGTTA-3' motif and, based on structural considerations, seemed large enough to contain the entire site contacted by a single Myb domain. We varied the spacing between the two telomeric "cassettes", maintaining the 6 bp phasing typical of vertebrate telomeres. No significant change in affinity was observed when the total length of the DNA separating the two cassettes varied between 6 and 30 nucleotides (Fig. 3.5, lanes 7-30; the difference in affinity between DNAs containing one or two copies of the cassettes will be discussed below).

In the framework of our model, this could be explained if either the full TRF1 site is contained within one single cassette or if the correct spacing between the two TRF1 half sites is not represented by any of these molecules. Alternatively, TRF1 might be

Fig. 3.6 **Schematic of the strategy for the selection of TRF1 binding sites**

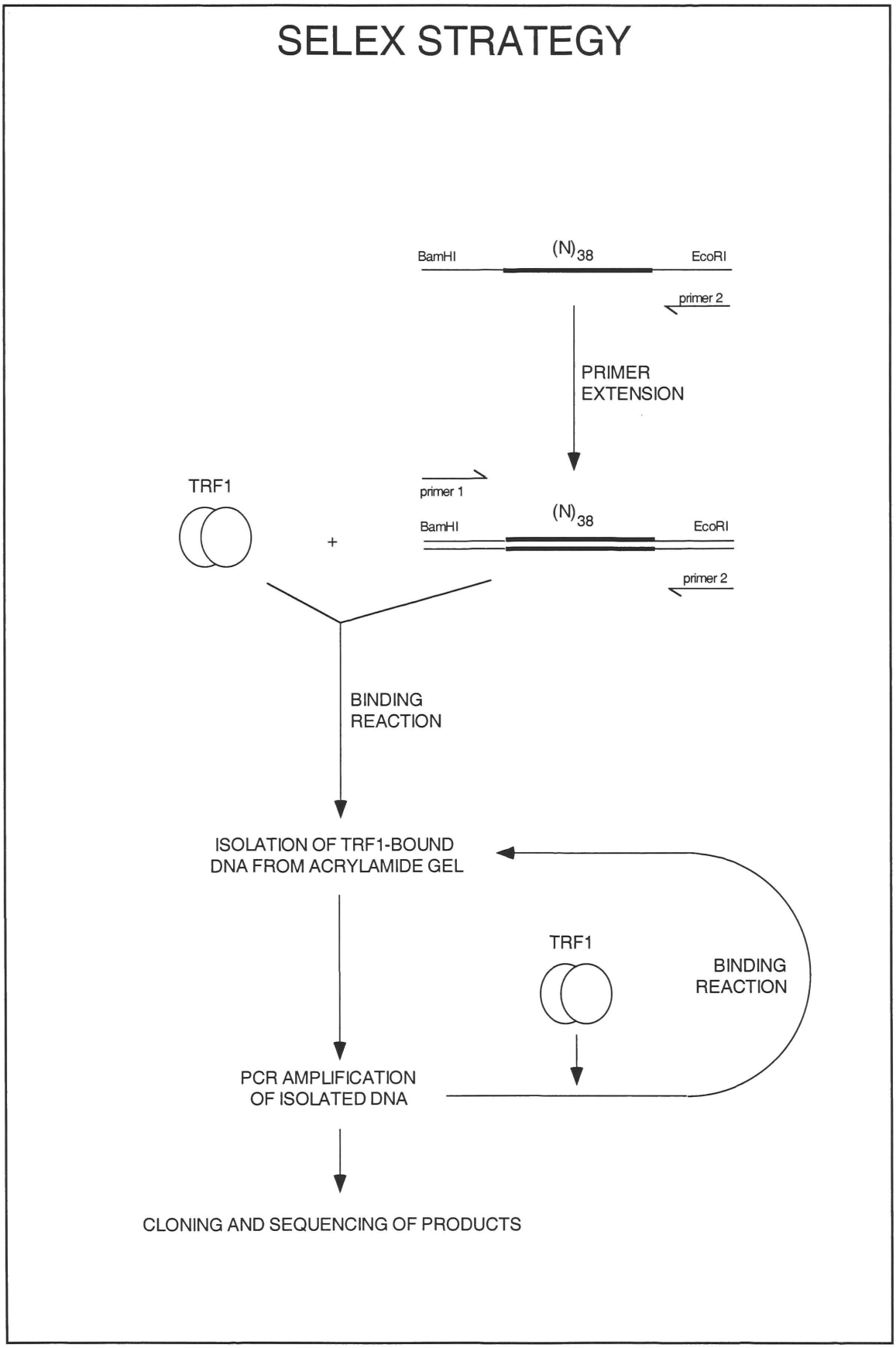

SELEX STRATEGY
BamHI
(N)38
EcoRI
primer 2
PRIMER
EXTENSION
TRF1
primer 1
(N)38
BamHI
EcoRI
+
primer 2
BINDING
REACTION
ISOLATION OF TRF1-BOUND
DNA FROM ACRYLAMIDE GEL
TRF1
BINDING
REACTION
PCR AMPLIFICATION
OF ISOLATED DNA
CLONING AND SEQUENCING OF PRODUCTS

Fig. 3.6

able to bind by contacting both the half sites in a manner that is not significantly affected by the relative distance of the sites. In any case, these results suggested that, if TRF1 binds to a site with definite spacing, this site is smaller than 36 bp or larger than 54 (in our experiments, a 36 bp site would be represented by the two "cassettes" separated by 12 bp – a substrate that was not tested-, whereas a 54 bp site would have a $\geq$36 bp spacer between the cassettes). A site bigger than 54 bp seems unlikely in light of the size of the protein and can be ruled out from band-shift and EM binding experiments on arrays of 12 repeats (Bianchi et al. 1997; Griffith et al. 1998). Furthermore, from EM studies, we can exclude a scenario in which the DNA is looped around a dimer of TRF1 that bears the two Myb motifs at opposite ends of the dimer (Griffith et al. 1998). Additionally, even though none of the experiments here described takes into account nucleosomal DNA, we have also not observed enhanced affinity of TRF1 for DNA wrapped around a nucleosome (an arrangement that can bring together on one side of the nucleosome octamer two sites separated by more than 50 bp - D. Rhodes, A, Bianchi and T. de Lange, unpublished).

Identification of the TRF1 binding site by SELEX

To determine the precise configuration of the TRF1 binding site, we undertook a PCR-based approach. The advantage of such an approach is that no a priori information about sequence specificity or mode of binding is required. It is however necessary to be sure that the size of the randomized oligonucleotide is large enough to contain the entire binding site. As discussed above, the binding experiments described in Fig. 3.5 show that if the TRF1 binding site is made of two identical half sites with strict spacing requirements, these must then encompass a piece of DNA smaller than 36 bp.

Fig. 3.7 **List of DNA sequences recovered by TRF1 binding after seven rounds of selection**

DNA sequences containing two proximally located copies of the minimal TRF1 binding site (5'-AGGGTT-3') are shown in (a), whereas (b) shows sequences containing two non-adjacent copies of the site.

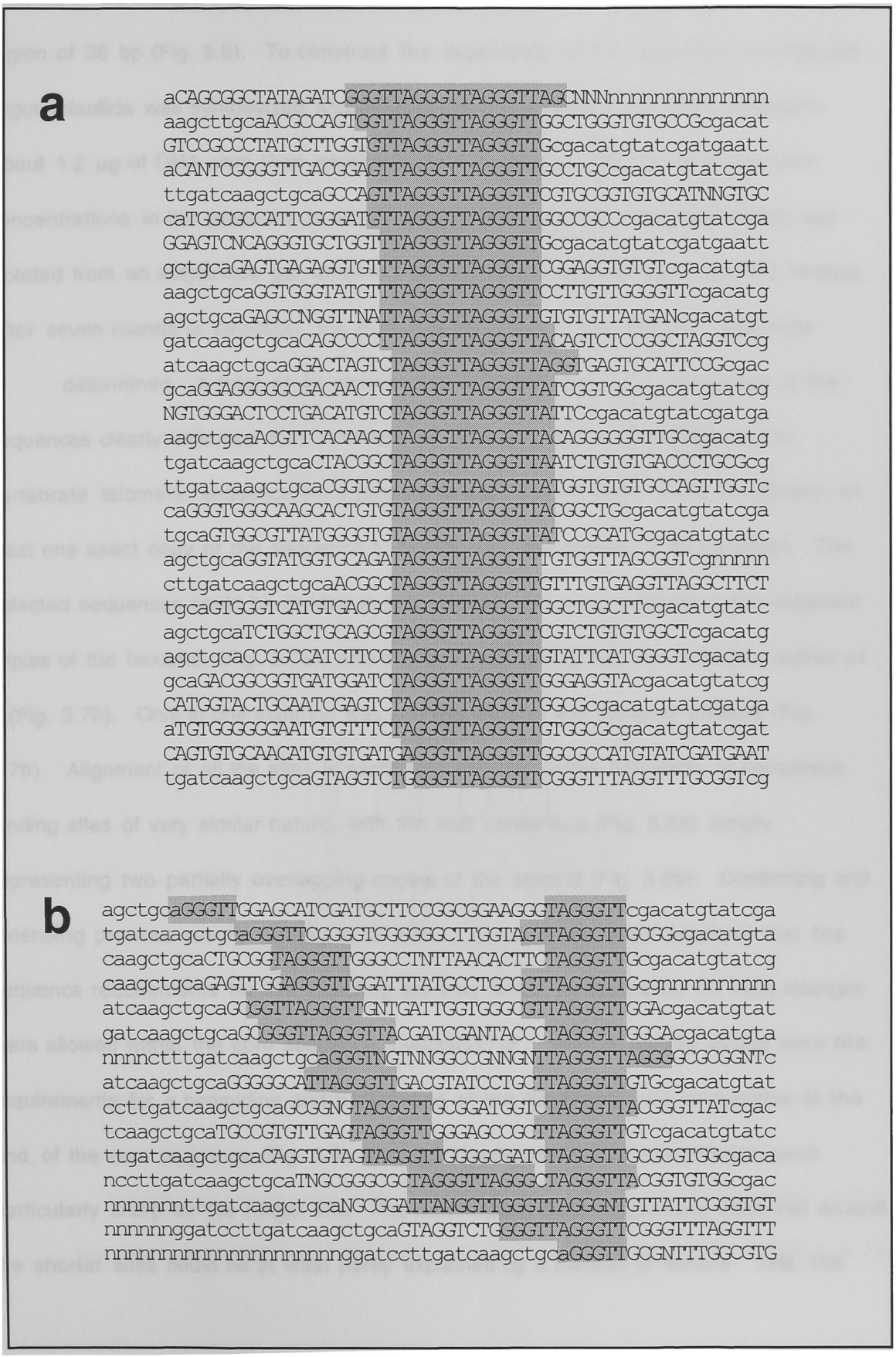
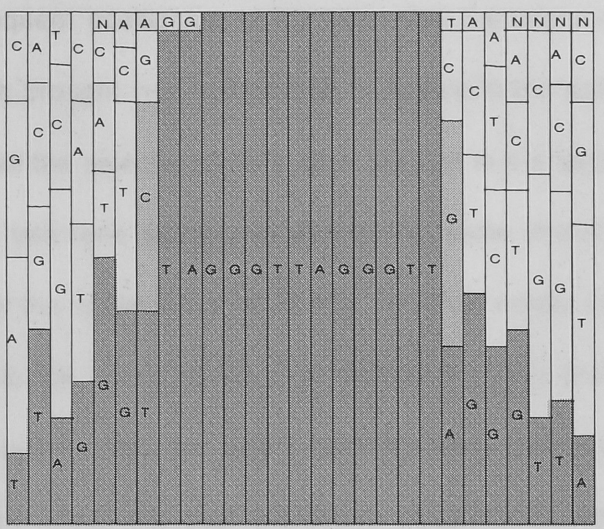

a
aCAGCGGCTATAGATCGGGTTAGGGTTAGGGTTAGCNNNnnnnnnnnnnnnn
aagcttgcaACGCCAGTGGTTAGGGTTAGGGTTGGCTGGGTGTGCCGcgacat
GTCCGCCCTATGCTTGGTGTTAGGGTTAGGGTTGcgacatgtatcgatgaatt
aCANTCGGGGTTGACGGAGTTAGGGTTAGGGTTGCCTGCcgacatgtatcgat
ttgatcaagctgcaGCCAGTTAGGGTTAGGGTTCGTGCGGTGTGCATNNGTGC
caTGGCGCCATTCGGGATGTTAGGGTTAGGGTTGGCCGCCcgacatgtatcga
GGAGTCNCAGGGTGCTGGTTTAGGGTTAGGGTTGcgacatgtatcgatgaatt
gctgcaGAGTGAGAGGTGTTTAGGGTTAGGGTTCGGAGGTGTGTcgacatgta
aagctgcaGGTGGGTATGTTTAGGGTTAGGGTTCCTTGTTGGGGTTcgacatg
agctgcaGAGCCNGGTTNATTAGGGTTAGGGTTGTGTGTTATGANcgacatgt
gatcaagctgcaCAGCCCCTTAGGGTTAGGGTTACAGTCTCCGGCTAGGTCcg
atcaagctgcaGGACTAGTCTAGGGTTAGGGTTAGGTGAGTGCATTCCGcgac
gcaGGAGGGGGCGACAACTGTAGGGTTAGGGTTATCGGTGGcgacatgtatcg
NGTGGGACTCCTGACATGTCTAGGGTTAGGGTTATTCcgacatgtatcgatga
aagctgcaACGTTCACAAGCTAGGGTTAGGGTTACAGGGGGGTTGCcgacatg
tgatcaagctgcaCTACGGCTAGGGTTAGGGTTAATCTGTGTGACCCTGCGcg
ttgatcaagctgcaCGGTGCTAGGGTTAGGGTTATGGTGTGTGCCAGTTGGTc
caGGGTGGGCAAGCACTGTGTAGGGTTAGGGTTACGGCTGcgacatgtatcga
tgcaGTGGCGTTATGGGGTGTAGGGTTAGGGTTATCCGCATGcgacatgtatc
agctgcaGGTATGGTGCAGATAGGGTTAGGGTTTGTGGTTAGCGGTcgnnnnn
cttgatcaagctgcaACGGCTAGGGTTAGGGTTGTTTGTGAGGTTAGGCTTCT
tgcaGTGGGTCATGTGACGCTAGGGTTAGGGTTGGCTGGCTTcgacatgtatc
agctgcaTCTGGCTGCAGCGTAGGGTTAGGGTTCGTCTGTGTGGCTcgacatg
agctgcaCCGGCCATCTTCCTAGGGTTAGGGTTGTATTCATGGGGTcgacatg
gcaGACGGCGGTGATGGATCTAGGGTTAGGGTTGGGAGGTAcgacatgtatcg
CATGGTACTGCAATCGAGTCTAGGGTTAGGGTTGGCGcgacatgtatcgatga
aTGTGGGGGGAATGTGTTTCTAGGGTTAGGGTTGTGGCGcgacatgtatcgat
CAGTGTGCAACATGTGTGATGAGGGTTAGGGTTGGCGCCATGTATCGATGAAT
tgatcaagctgcaGTAGGTCTGGGGTTAGGGTTCGGGTTTAGGTTTGCGGTcg
b
agctgcaGGGTTGGAGCATCGATGCTTCCGGCGGAAGGGTAGGGTTcgacatgtatcga
tgatcaagctgcaGGGTTCGGGGTGGGGGGCTTGGTAGTTAGGGTTGCGGcgacatgta
caagctgcaCTGCGGTAGGGTTGGGCCTNTTAACACTTCTAGGGTTGcgacatgtatcg
caagctgcaGAGTTGGAGGGTTGGATTTATGCCTGCCGTTAGGGTTGcgnnnnnnnnnn
atcaagctgcaGCGGTTAGGGTTGNTGATTGGTGGGCGTTAGGGTTGGAcgacatgtat
gatcaagctgcaGCGGGTTAGGGTTACGATCGANTACCCTAGGGTTGGCAcgacatgta
nnnnctttgatcaagctgcaGGGTNGTNNGGCCGNNGNTTAGGGTTAGGGGCGCGGNTc
atcaagctgcaGGGGGCATTAGGGTTGACGTATCCTGCTTAGGGTTGTGcgacatgtat
ccttgatcaagctgcaGCGGNGTAGGGTTGCGGATGGTCTAGGGTTACGGGTTATcgac
tcaagctgcaTGCCGTGTTGAGTAGGGTTGGGAGCCGCTTAGGGTTGTcgacatgtatc
ttgatcaagctgcaCAGGTGTAGTAGGGTTGGGGCGATCTAGGGTTGCGCGTGGcgaca
ncctttgatcaagctgcaTNGCGGGCGCTAGGGTTAGGGCTAGGGTTACGGTGTGGcgac
nnnnnnnttgatcaagctgcaNGCGGATTANGGTTGGGTTAGGGNTGTTATTCGGGTGT
nnnnnnggatccttgatcaagctgcaGTAGGTCTGGGGTTAGGGTTCGGGTTTAGGTTT
nnnnnnnnnnnnnnnnnnnnnggatccttgatcaagctgcaGGGTTGCGNTTTGGCGTG

Fig. 3.7

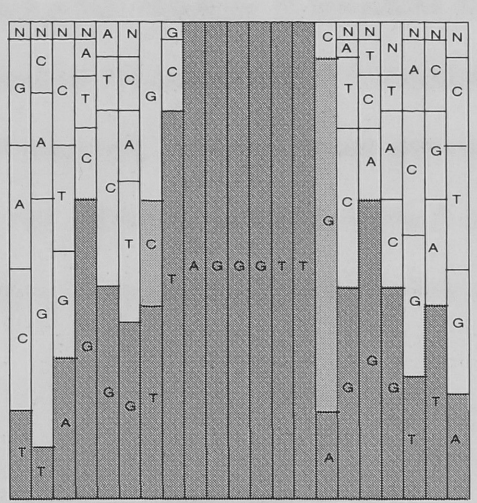
Therefore, we designed a degenerate oligonucleotide library with a central randomized region of 38 bp (Fig. 3.6). To construct the degenerate library, a partially randomized oligonucleotide was synthesized and rendered double-stranded by primer extension. About 1.2 μg of DNA were then allowed to bind to purified TRF1 at various protein concentrations in the presence of ϕX174 DNA as a competitor. The bound DNA was isolated from an acrylamide gel, amplified by PCR, and subjected again to TRF1 binding. After seven rounds of selection, the PCR products were cloned and their sequence determined. A total of 43 clones were analyzed (Fig. 3.7). Inspection of the sequences clearly indicated that stretches of DNA with perfect homology to the vertebrate telomeric sequence were present in each one of them. More specifically, at least one exact copy of the sequence 5'-AGGGTT-3' was present in all instances. The selected sequences could be divided in two classes: one class containing two adjacent copies of the hexamer (Fig. 3.7a), and one class containing two non-adjacent copies of it (Fig. 3.7b). Only in one instance was a single copy of the hexamer present (Fig. 3.7b). Alignment of all the sites in each class resulted in the derivation of consensus binding sites of very similar nature, with the first consensus (Fig. 3.8a) simply representing two partially overlapping copies of the second (Fig. 3.8b). Confirming and extending previous results (Zhong et al. 1992), 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 at the beginning, and for a purine at the end, of the core sequence (Fig. 3.8a and b). The boundaries of conservation were particularly sharp for the longer site. The non-random base frequencies observed around the shorter sites could be at least partly explained by a number of factors: first, the

Fig. 3.8 **Alignement of the two classes of sites recovered by SELEX and definition of the consensus binding site for the TRF1 Myb domain**

a



b



c

base frequencies in synthesized library oligonucleotide	G	0.450
	A	0.138
	T	0.222
	C	0.185

Fig. 3.8

unequal base frequencies in the starting DNA, which is G-rich (see below); second, the presence of the second site, which further contributes to deviate base frequencies from randomness; finally, the frequent positioning of the sites near the edges of the randomized sequence, which brought non-randomized sequence in the analyzed stretch. Remarkably, not only were all the selected double sites present in the same orientation, but, more in general, all the telomeric sequences showed the same strand polarity with respect to the orientation in the oligonucleotide library. In other words, in all the clones the G-rich strand belonged to the strand of the oligonucleotide library that had been chemically syntesized. We believe that this bias in strand polarity does not reflect an effect of the non-randomized sequences from the library on selection, but rather is a consequence of a bias in the base composition of the synthesized oligonucleotide (Fig. 3.8c) which turns out to have a high frequency of G residues (0.45 instead of the expected 0.25; as determined by the sequencing of 5 unselected clones (containing 190 bp, see Experimental Procedures). Identical results were obtained with a second oligonucleotide library that had a different sequence on the 5' side of the randomized sequence. Attempts to obtain a less biased library from other suppliers were unsuccessful.

TRF1 binds to DNA by engaging both Myb domains

In good agreement with the 5'-GGGTTA-3'core site characterized by Rhodes and coworkers for binding by the isolated Myb domain of TRF1 (Konig et al. 1998), we believe that the 5'-YTAGGGTTR-3' site represents the sequence recognized by a single Myb domain of TRF1 (see below). The finding that in all cases examined except one, two copies of the 5'-AGGGTT-3' sequence were present, is in agreement with a model

Fig. 3.9 **Band-shift analysis of binding of TRF1 and Myb to DNA molecules containg one or two copies of the minimal binding site**

Binding assays were performed under standard TRF1 conditions, with 0.1% NP40. Poli(dI·dC) competitor (100 ng/sample) was added to the TRF1 reactions. Gel was run in 0.5xTBE at 4°C. Quantitation of reactions shown in (a) were obtained by Phosphorimager and are shown in (b). DNA probes "half" and "full" were restriction fragments containing the indicated telomeric sequence, and were 111 and 117 bp long, respectively. Protein concentrations varied in two-fold steps. Quantitation was performed by taking all the complexes into account as bound DNA (thus both complexes in the case of the Myb domain).

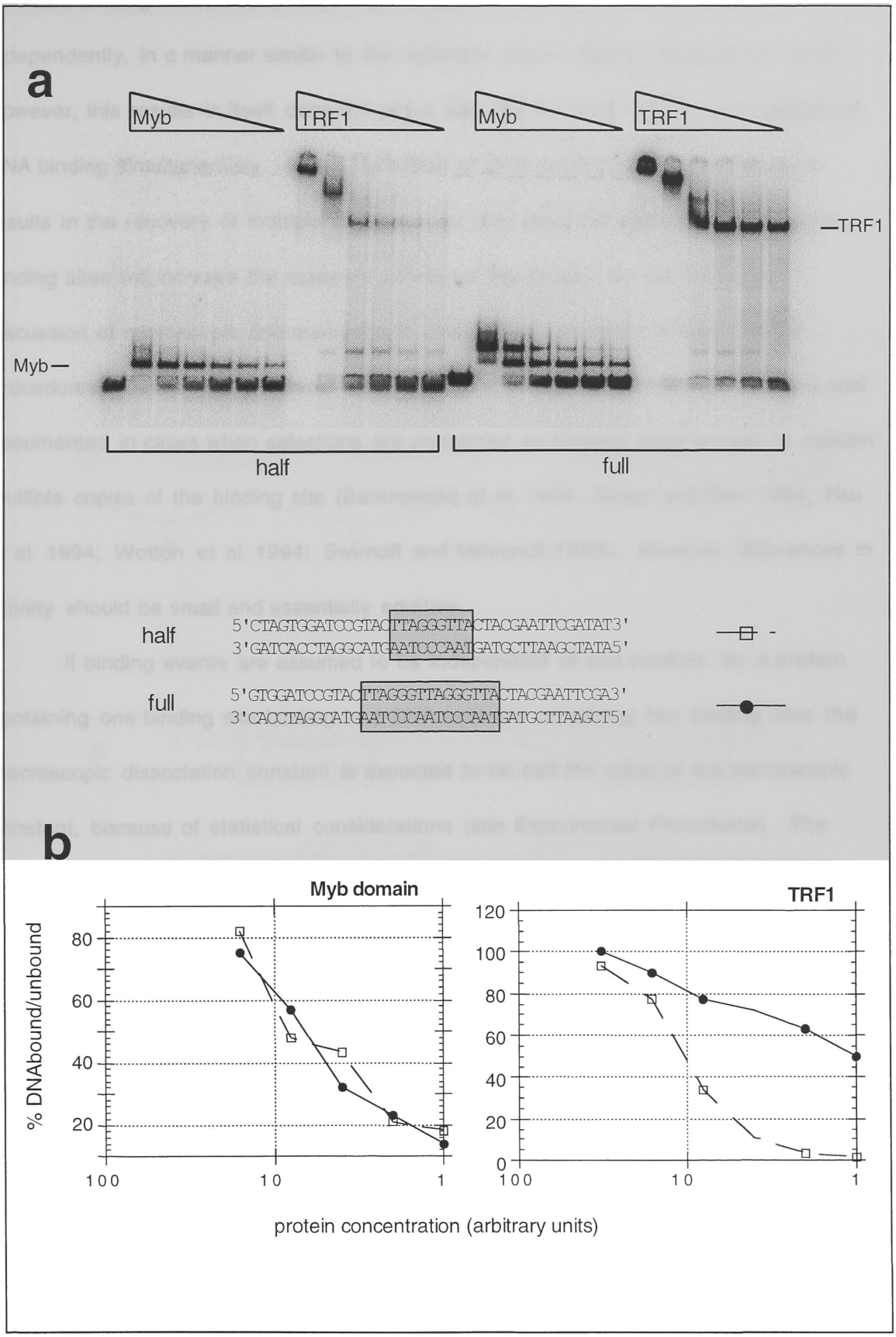

a
Myb
TRF1
Myb
TRF1
—TRF1
Myb—
half
full
half
5'CTAGTGGATCCGTACTTAGGGTTACTACGAATTCGATAT3'
3'GATCACCTAGGCATGAATCCCAATGATGCTTAAGCTATA5'
full
5'GTGGATCCGTACTTAGGGTTAGGGTTACTACGAATTCGA3'
3'CACCTAGGCATGAATCCCAATCCCAATGATGCTTAAGCT5'
b
Myb domain
TRF1
% DNAbound/unbound
80
60
40
20
120
100
80
60
40
20
0
100
10
1
100
10
1
protein concentration (arbitrary units)

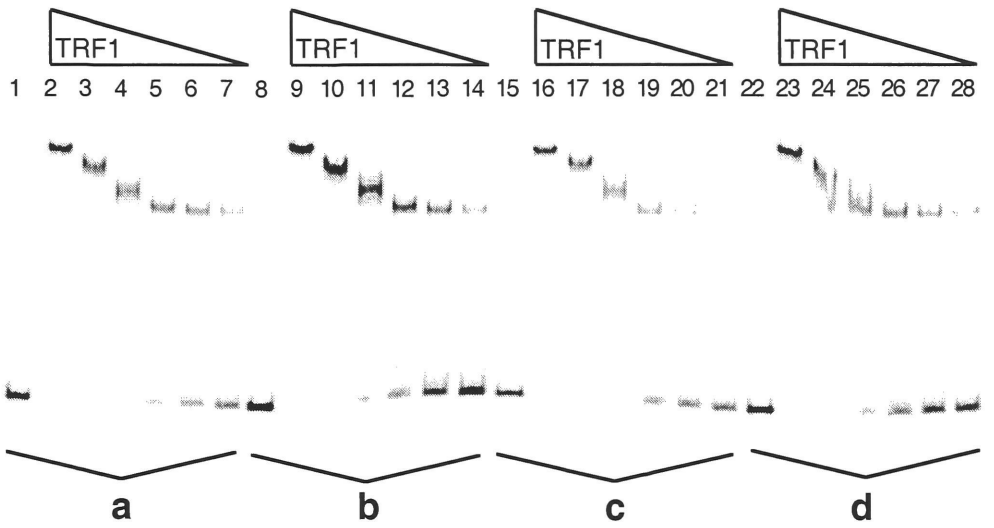
Fig. 3.9

for TRF1 binding in which each Myb motif in the TRF1 dimer contacts one site independently, in a manner similar to the telomeric protein Rap1p (Konig et al. 1996). However, this results in itself does not prove that the two Myb domains are engaged in DNA binding simultaneously. In fact, selection of DNA binding sites by SELEX often results in the recovery of multiple copies of the site, since the presence of additional binding sites will increase the apparent affinity of the protein for the DNA (see discussion of microscopic and macroscopic dissociation constants in Experimental Procedures) (Cantor and Schimmel 1980; Wilson et al. 1993). This phenomenon is well documented in cases when selections are performed on libraries large enough to contain multiple copies of the binding site (Baranowskij et al. 1994; Brown and Baer 1994; Hsu et al. 1994; Wotton et al. 1994; Swirnoff and Milbrandt 1995). However, differences in affinity should be small and essentially additive.

If binding events are assumed to be independent of one another, for a protein containing one binding site binding to a DNA molecule containing two binding sites the macroscopic dissociation constant is expected to be half the value of the microscopic constant, because of statistical considerations (see Experimental Procedures). The same protein will instead have identical microscopic and macroscopic dissociation constants when binding to a DNA molecule containing one site. Therefore the Myb domain is predicted to have Kd values differing by two-fold for DNAs bearing either one or two minimal binding sites. On the other hand, a protein with two binding sites will have a macroscopic dissociation constant for the first binding event to a DNA containing one binding site that is equal to half the value of the microscopic constant, and for binding to a DNA with two binding sites, a value that is equal to one fourth of the microscopic constant. In the case that TRF1 binds to DNA by engaging only one Myb

Fig. 3.10 **Binding assay of two representative sequences from each of the two classes of SELEX TRF1 binding sites**

Band-shift assays were done under standar TRF1 conditions. DNA fragments used for binding were 170 bp long. Diagram indicates in bold matches to TRF1 minimal consensus site. Area potentially occupied by consensus is boxed. Protein concentrations varied in 1.5-fold steps.



a

5'GCGG**TAGGGTTG**GGCNTCTTAACACGT**CTAGGGTTG**CGAC3'
3'CGCC**ATCCCAAC**CCGNAGAATTGTGCA**GATCCCAAC**GCTG5'

b

5'AGGTGTAG**TAGGGTTG**GGGCGAT**CTAGGGTTG**CGCGTGGC3'
3'TCCACATC**ATCCCAAC**CCCGCTA**GATCCCAAC**GCGCACCG5'

c

5'TCTGGCTGCAGCG**TAGGGTTAGGGTT**CGCTGTGTGGCTC3'
3'AGACCGACGTCGC**ATCCCAATCCCAA**GCGACACACCGAG5'

d

5'CCGGCCATCTTC**CTAGGGTTAGGGTTG**TATTCATGGGGT3'
3'GGCCGGTAGAAG**GATCCCAATCCCAAC**ATAAGTACCCCA5'

Fig. 3.10

domain, then, for both the dimeric protein and the isolated Myb domain, a two-fold increase in affinity should be expected for a DNA containing two minimal binding sites compared to a DNA that contains a single site. On the other hand, if both Myb domains are engaged simultaneously in binding, a cooperative effect is to be expected for TRF1 with an increase larger than two-fold for binding to the single- and double-site DNAs. Therefore, to discriminate between the two alternative possibilities that the full TRF1 site might be represented by one or two copies of the minimal binding site we compared the relative affinities of TRF1 for DNAs containing either one or two copies of the minimal binding sites (Fig. 3.9a, and data not shown). A difference in affinity of about 10 fold was observed for TRF1 binding to the sites, with the DNA containing two copies being the better substrate, as expected (Fig. 3.9b). At high TRF1 concentrations a larger gel-shift complex is observed, probably due to protein-protein interactions (see also Fig. 3.5).

On the other hand, comparison of the binding affinities of one molecule of the Myb domain alone for the two DNAs showed essentially no change in affinity (Fig. 3.9b). With the Myb domain, as expected, one protein molecule was bound to the probe containing one minimal binding site, and two protein molecules were bound to the probe containing two minimal binding sites, confirming that this site represents the site recognized by a single Myb domain. Precise quantitation of the gels was hindered by smearing, and this may account for the similar affinities observed for Myb with both probes, instead of the two-fold difference theoretically expected. Alternatively, the two binding events of Myb on this DNA may not be entirely independent. However, the differences in TRF1 binding for the two sites were dramatic. A cooperative effect of TRF1 binding to a site containing two repeats was observed, that was not entirely

Fig. 3.11 **Comparison of TRF1 binding sites carrying two copies of the minimal site in either direct or inverted orientation**

TRF1 binding assays were done in standard TRF1 conditions. Protein concentrations varied in three fold steps. Size of DNA fragments was 80 bp.

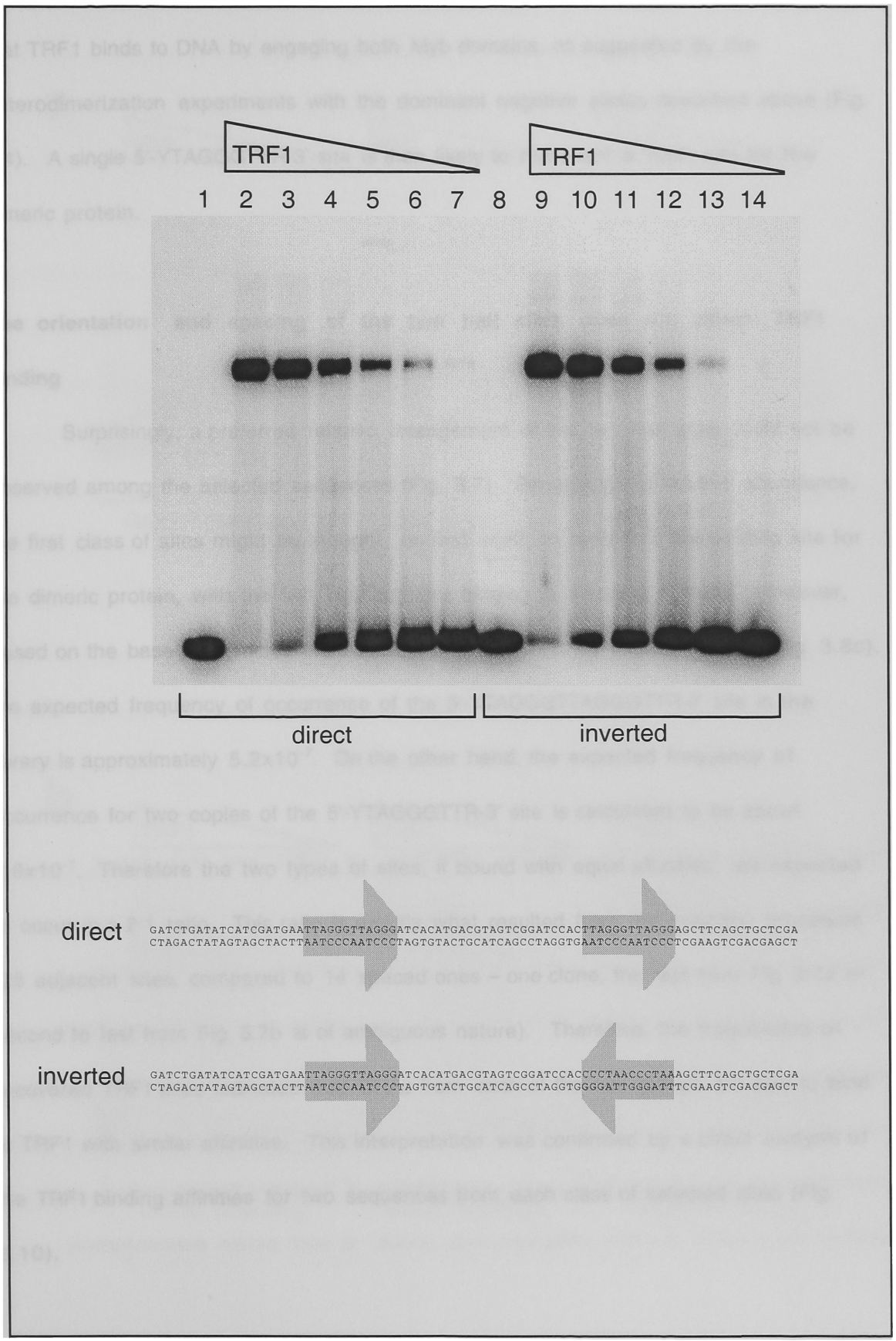

TRF1
TRF1
1 2 3 4 5 6 7 8 9 10 11 12 13 14
direct
inverted
direct
GATCTGATATCATCGATGAATTAGGGTTAGGGATCACATGACGTAGTCGGATCCACTTAGGGTTAGGGAGCTTCAGCTGCTCGA
CTAGACTATAGTAGCTACTTAATCCCAATCCCTAGTGTACTGCATCAGCCTAGGTGAATCCCAATCCCTCGAAGTCGACGAGCT
inverted
GATCTGATATCATCGATGAATTAGGGTTAGGGATCACATGACGTAGTCGGATCCACCCCTAACCCTAAAGCTTCAGCTGCTCGA
CTAGACTATAGTAGCTACTTAATCCCAATCCCTAGTGTACTGCATCAGCCTAGGTGGGGATTGGGATTTCGAAGTCGACGAGCT

Fig. 3.11

attributable to simple multimerization of the binding site. This result supported the idea that TRF1 binds to DNA by engaging both Myb domains, as suggested by the heterodimerization experiments with the dominant negative alleles described above (Fig. 3.4). A single 5'-YTAGGGTTR-3' site is then likely to represent a "half" site for the dimeric protein.

The orientation and spacing of the two half sites does not affect TRF1 binding

Surprisingly, a preferred relative arrangement of the two half sites could not be observed among the selected sequences (Fig. 3.7). Because of its relative abundance, the first class of sites might be thought, on first sight, to represent the binding site for the dimeric protein, with the two Myb domains binding to adjacent repeats. However, based on the base frequencies encountered in the unselected oligonucleotide (Fig. 3.8c), the expected frequency of occurrence of the 5'-YTAGGGTTAGGGTTR-3' site in the library is approximately 5.2×10^{-7}. On the other hand, the expected frequency of occurrence for two copies of the 5'-YTAGGGTTR-3' site is calculated to be about 2.6×10^{-7}. Therefore the two types of sites, if bound with equal affinities, are expected to occur in a 2:1 ratio. This ratio is exactly what resulted from the selection procedure (28 adjacent sites, compared to 14 spaced ones – one clone, the last from Fig. 3.7a or second to last from Fig. 3.7b is of ambiguous nature). Therefore, the frequencies of recovered TRF1 sites indicated that DNAs from each of the two groups are able to bind to TRF1 with similar affinities. This interpretation was confirmed by a direct analysis of the TRF1 binding affinities for two sequences from each class of selected sites (Fig. 3.10).

Fig. 3.12 **Comparison of the binding affinities of TRF1 and Myb for arrays of TTAGGG repeats**

Binding assays were performed under standard TRF1 conditions, with 0.1% NP40. Sonicated *E.coli* DNA competitor (100 ng/sample) was added to the TRF1 reactions. Gel was run in 0.5xTBE at 4°C. DNA probes with 3.5, 6.5 and 12.5 repeats were restriction fragments 88, 106 and 142 bp long, respectively. Protein concentrations varied in two-fold steps. Quantitation of binding affinities was based on the percent of probe unbaound (free probe) at each potein concentration.

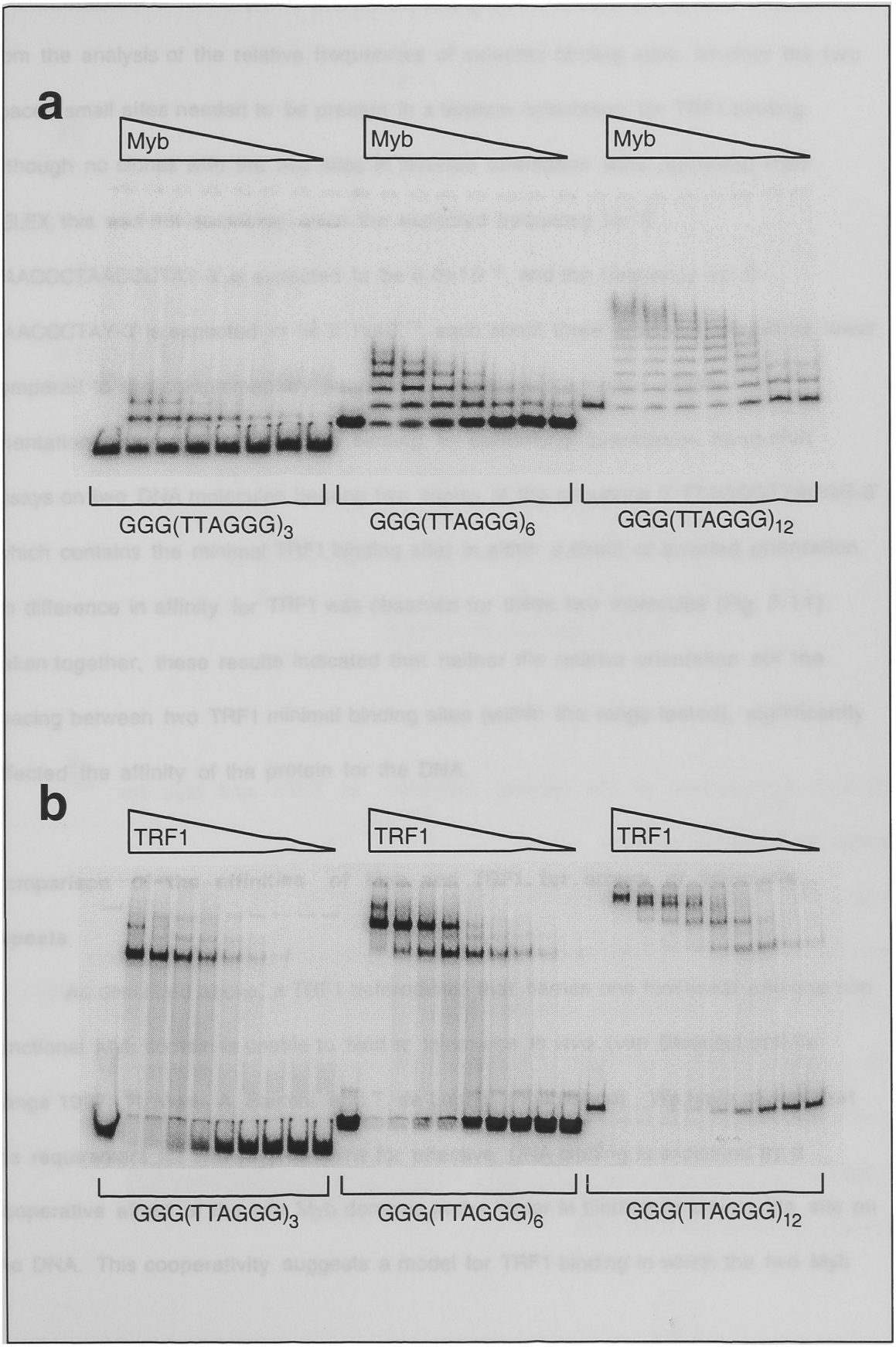

a
Myb
Myb
Myb
GGG(TTAGGG)3
GGG(TTAGGG)6
GGG(TTAGGG)12
b
TRF1
TRF1
TRF1
GGG(TTAGGG)3
GGG(TTAGGG)6
GGG(TTAGGG)12

Fig. 3.12

Because of the strand bias in base composition we were not able to determine, from the analysis of the relative frequencies of selected binding sites, whether the two spaced small sites needed to be present in a tandem orientation for TRF1 binding. Although no clones with the two sites in inverted orientation were recovered from SELEX, this was not surprising, since the expected frequency for 5'-RAACCCTAACCCTAY-3' is expected to be $6.0x10^{-10}$, and the frequency for 5'-RAACCCTAY-3' is expected to be $2.1x10^{-10}$, each about three orders of magnitude lower compared to the complementary sequences. We therefore tested whether the orientation of the half sites affects binding, by performing quantitative band-shift assays on two DNA molecules bearing two copies of the sequence 5'-TTAGGGTTAGGG-3' (which contains the minimal TRF1 binding site) in either a direct or inverted orientation. No difference in affinity for TRF1 was observed for these two molecules (Fig. 3.11). Taken together, these results indicated that neither the relative orientation nor the spacing between two TRF1 minimal binding sites (within the range tested), significantly affected the affinity of the protein for the DNA.

Comparison of the affinities of Myb and TRF1 for arrays of telomeric repeats

As described above, a TRF1 heterodimer that carries one functional and one non functional Myb domain is unable to bind to telomeres *in vivo* (van Steensel and de Lange 1997; H. Moss, A. Bianchi and T. de Lange, unpublished). We have shown that the requirement for two Myb domains for effective DNA binding is explained by a cooperative effect of the two Myb domains in the dimer in binding to a bi-partite site on the DNA. This cooperativity suggests a model for TRF1 binding in which the two Myb

Fig. 3.13 **Comparison of the binding affinities of TRF1 and Myb for arrays of TTAGGG repeats**

In (a) a estimates of the apparent dissociation constants for TRF1 and Myb as a function of telomeric repeats are shown. The data used are from the gels shown in Fig. 3.12. In (b) competition experiments for binding to TRF1 or Myb complex with a plasmid DNA containing 27 TTAGGG repeats are shown. Increments in competitor amount are in two-fold steps. The amount of total competitor DNA in the reactions was kept constant at 1 μg with the addition of appropriate amounts of sonicated *E.coli* DNA.

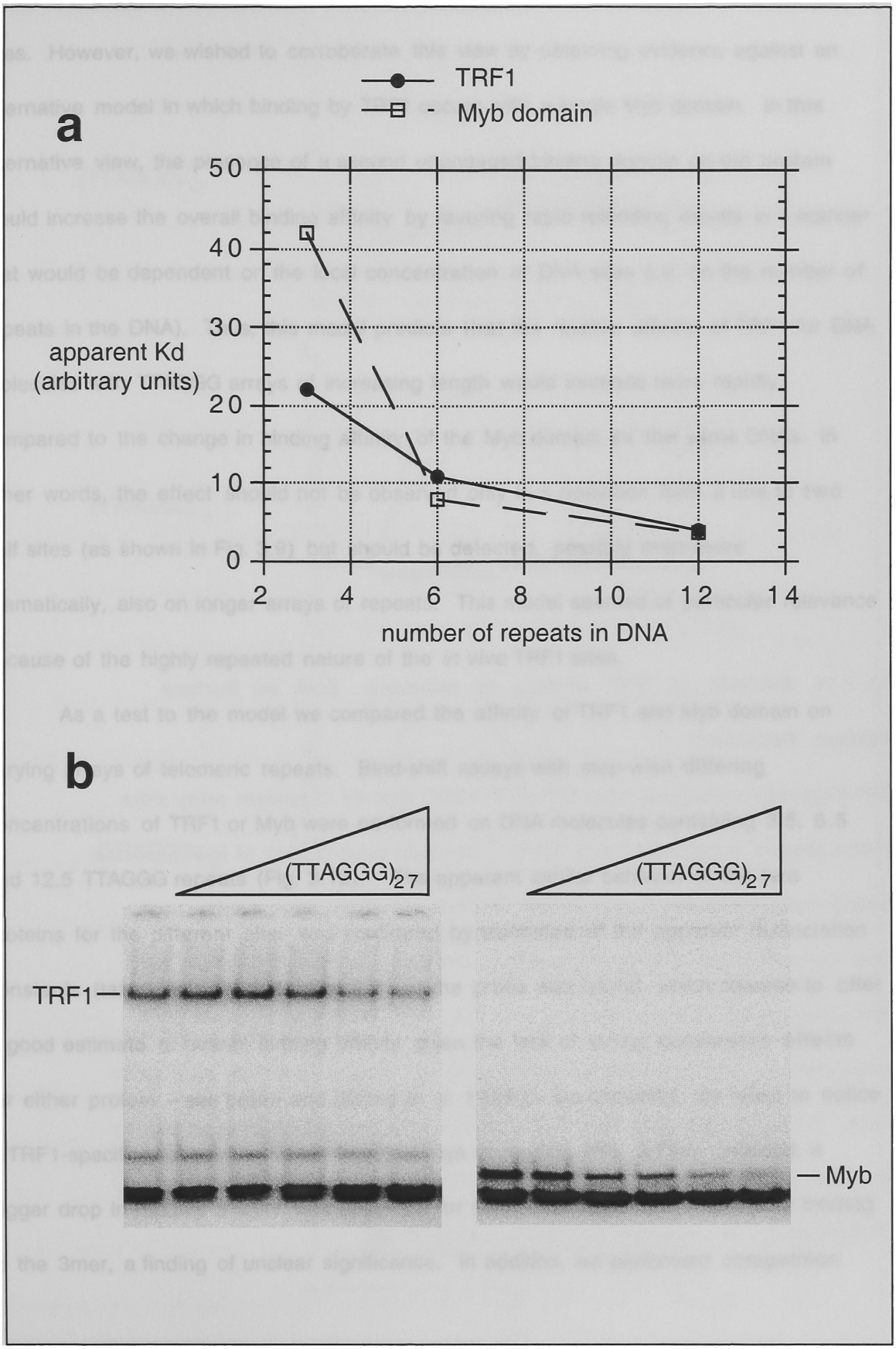

a
TRF1
Myb domain
apparent Kd
(arbitrary units)
50
40
30
20
10
0
2
4
6
8
10
12
14
number of repeats in DNA
b
(TTAGGG)27
(TTAGGG)27
TRF1
Myb

Fig. 3.13

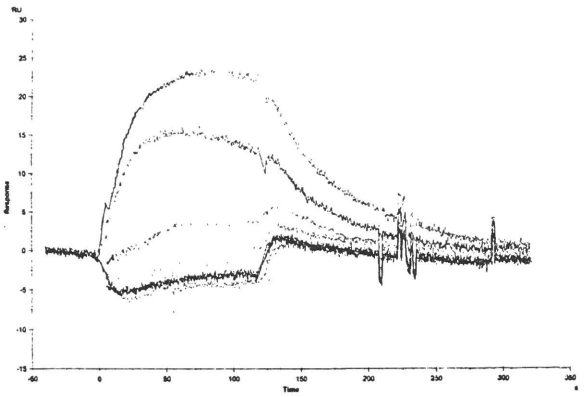
domains are engaged independently but simultaneously in DNA binding to the two half sites. However, we wished to corroborate this view by obtaining evidence against an alternative model in which binding by TRF1 occurs with a single Myb domain. In this alternative view, the presence of a second unengaged binding domain on the protein would increase the overall binding affinity by favoring rapid rebinding events in a manner that would be dependent on the local concentration of DNA sites (i.e. on the number of repeats in the DNA). Thus, this model predicts that the relative affinity of TRF1 for DNA molecules with TTAGGG arrays of increasing length would increase more rapidly compared to the change in binding affinity of the Myb domain for the same DNAs. In other words, the effect should not be observed only in a transition form a one to two half sites (as shown in Fig. 3.9) but should be detected, possibly even more dramatically, also on longer arrays of repeats. This model seemed of particular relevance because of the highly repeated nature of the *in vivo* TRF1 sites.

As a test to the model we compared the affinity of TRF1 and Myb domain on varying arrays of telomeric repeats. Bind-shift assays with step-wise differing concentrations of TRF1 or Myb were performed on DNA molecules containing 3.5, 6.5 and 12.5 TTAGGG repeats (Fig. 3.12). The apparent similar behavior of the two proteins for the different sites was confirmed by estimates of the apparent dissociation constants (taken as the point when 50% of the probe was bound, which seemed to offer a good estimate of overall binding affinity given the lack of strong cooperative effects for either protein – see below and (Konig et al. 1998)). As expected, we failed to notice a TRF1-specific decrease in Kd for longer arrays of repeats (Fig. 3.13a). Instead, a bigger drop in relative affinity was observed for Myb, essentially due to the poor binding to the 3mer, a finding of unclear significance. In addition, we performed competition

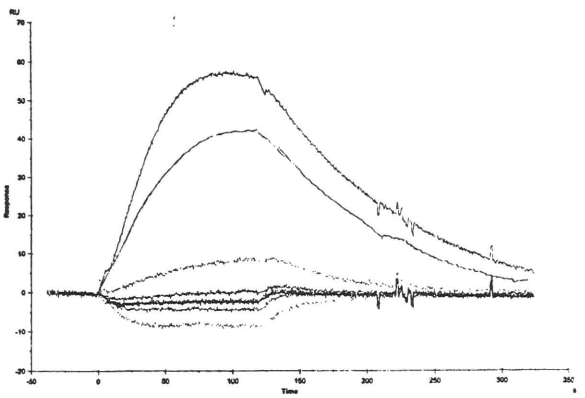
Fig. 3.14 **Analysis of TRF1 binding to telomeric DNA by Surface Plasmon Resonance**

DNA fragments containing from 2 to 5 TTAGGG repeats in tandem array were used in Biacore experiments with TRF1. The DNAs used in each of four separate binding experiments are shown at left. In each one of the four graphs, each curve represents one TRF1 concentration. From top to bottom, whithin each panel, binding curves were obtained with 250, 125, 62.5 and 31.2 nM TRF1 concentration. Time 0 on the X axis indicates the bieginning of the association phase. Time 120 (seconds) marks the beginning of the dissociation phase. In the Y axis the highest indicated number is 30, 70, 160 and 120 for the (TTAGGG)2,3,4 and 5 repeat substrates, respectively.

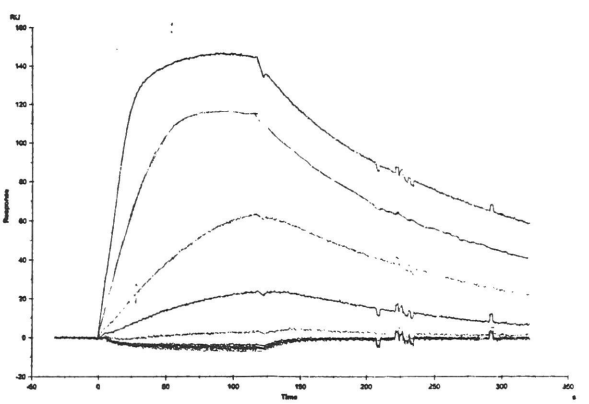
$(TTAGGG)_2$



$(TTAGGG)_3$



$(TTAGGG)_4$



$(TTAGGG)_5$

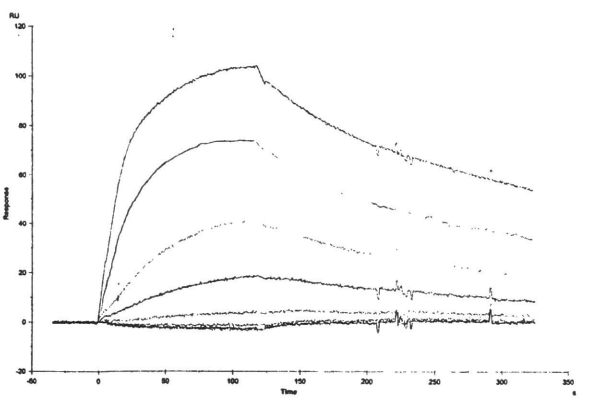


Fig. 3.14

experiments with a DNA substrate containing 27 repeats against binding to a DNA fragment containing the half site, both for TRF1 and Myb (Fig. 3.13b). Also in this case we failed to observe preferential binding of TRF1 to longer arrays of repeats, as indicated by the similar degree of inhibition of binding by the 27mer on TRF1 and on Myb. In summary, we found no evidence for a TRF1-specific (compared to Myb alone) selective preference for long arrays of telomeric repeats that might support an alternative mode of DNA binding.

Analysis of TRF1 binding to DNA by Surface Plasmon Resonance

Since the TRF1 off-rate is extremely fast and unmeasurable by band-shift experiments in which complex dissociation is measured by following the ability of an excess of cold competitor DNA to bind to bound TRF1 (T. de Lange, unpublished) we decided to employ Surface Plasmon Resonance to study the kinetics of TRF1 binding to DNA. Biacore technology allows the kinetics of association and dissociation of macromolecules to be followed in real time. In this approach, biotinylated DNA is bound to a streptavidin-coated microchip contained in a microcell. Buffer solution is then flushed through the cell and (at time 0), the buffer is replaced by the protein solution. After an appropriate amount of time (120 seconds, in our case, which define the association phase), the injection of protein is stopped and replaced by buffer. This marks the beginning of the dissociation phase. The experimental data are then fitted to theoretical curves. We were unable to obtain quantitative estimates of kinetic parameters because the experimental curves poorly fitted theoretical models (Fig. 3.14). However, it was evident that dissociation of TRF1 became increasily slower on longer arrays of repeats. The effect was essentially independent of protein concentration, and the calculated stoichiometry of binding (which ranged from 0.2 to

Fig. 3.15 **Estimation of the half-life of TRF1-DNA complex by Surface Plasmon Resonance**

The data from the Biacore experiments form Fig. 3.14 were analyzed by plotting the percent of signal (RUs) retained as a function of time, for the four sets of DNAs, at each protein concentration. The number next to each curve, indicates the number of TTAGGG repeats present in that particular DNA. The protein concentration present in the binding to each set of DNAs is indicated in the upper right corner of each graph.

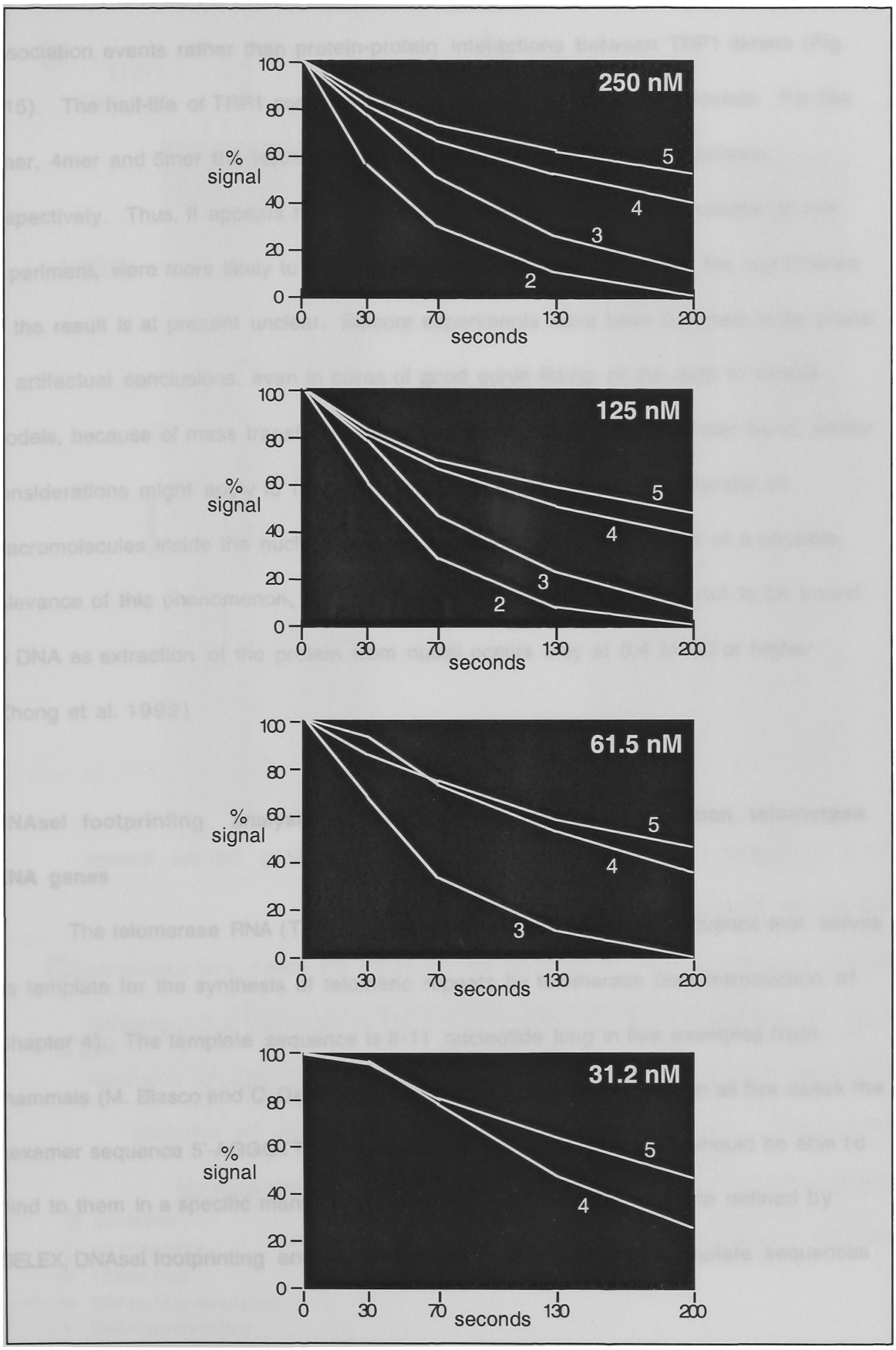

250 nM
125 nM
61.5 nM
31.2 nM
% signal
seconds
0
20
40
60
80
100
0
30
70
130
200
2
3
4
5

Fig. 3.15

1.2 protein molecule per DNA molecule) suggested that we were observing protein-DNA association events rather than protein-protein interactions between TRF1 dimers (Fig. 3.15). The half-life of TRF1 complexes on $(TTAGGG)_2$ was about 30 seconds. For the 3mer, 4mer and 5mer the value increased to about 70, 130 and 200 seconds, respectively. Thus, it appears that the rebinding events, under the conditions of the experiment, were more likely to occur on multimerized sites. However, the significance of the result is at present unclear. Biacore experiments have been reported to be prone to artifactual conclusions, even in cases of good curve fitting of the data to simple models, because of mass transfer limitations (Schuck 1997). On the other hand, similar considerations might apply to the *in vivo* situation given the extreme density of macromolecules inside the nucleus (Garner and Burg 1994). In support of a possible relevance of this phenomenon, we note that little if any TRF1 appears not to be bound to DNA as extraction of the protein from nuclei occurs only at 0.4 M KCl or higher (Zhong et al. 1992)

DNAseI footprinting analysis of TRF1 on the mouse and human telomerase RNA genes

The telomerase RNA (TER) genes contain a short telomeric sequence that serves as template for the synthesis of telomeric repeats by telomerase (see Introduction of Chapter 4). The template sequence is 8-11 nucleotide long in five examples from mammals (M. Blasco and C. Greider, personal communication). Since in all five cases the hexamer sequence 5'-AGGGTT-3' is present, we predicted that TRF1 should be able to bind to them in a specific manner. In agreement with the TRF1 half site defined by SELEX, DNAseI footprinting analysis on the human and mouse TER template sequences

Fig. 3.16 **DNAse I footprinting analysis of TRF1 binding to the human and mouse telomerase RNA template sequences**

DNAse I footprints of TRF1 on hTER (a) and mTER (b) are shown, and summarized in (c). Binding and DNAsel reactins were performed as described in Experimental Procedures. Maxam and Gilbert sequencing reactis were performed for sequence determination. The thick black bars in (a), (b) and (c) indicate the template regions in the RNAs. The hatched area in the template sequences in (c) highlights the core sequence of the minimal TRF1 binding site.

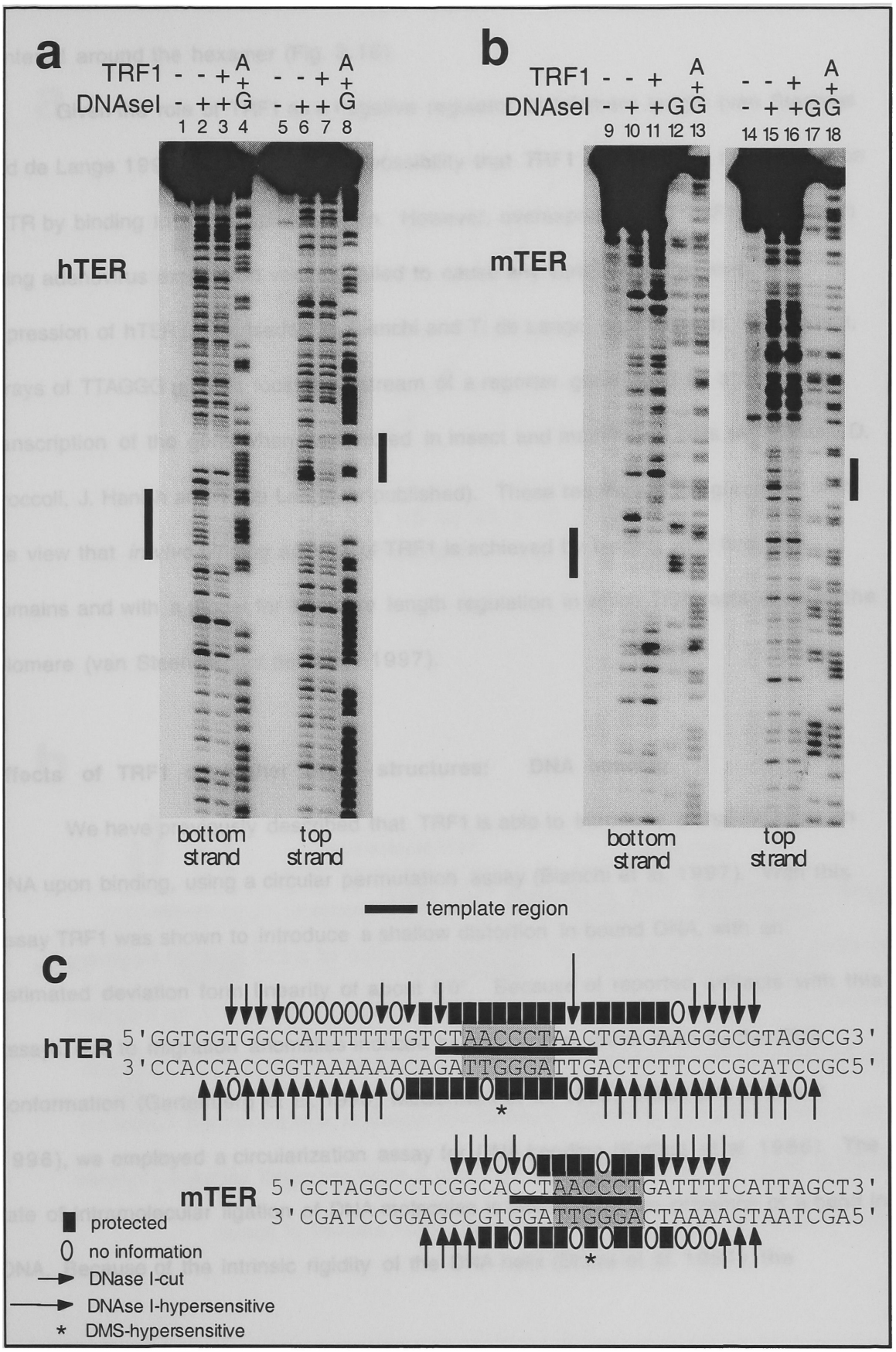

a
TRF1 - - + A +
DNAseI - + + G
1 2 3 4 5 6 7 8
hTER
bottom strand
top strand
b
TRF1 - - + A +
DNAseI - + + G G
9 10 11 12 13 14 15 16 17 18
mTER
bottom strand
top strand
template region
c
hTER
5'GGTGGTGGCCATTTTTTGTCTAACCCTAACTGAGAAGGGCGTAGGCG3'
3'CCACCACCGGTAAAAAACAGATTGGGATTGACTCTTCCCGCATCCGC5'
mTER
5'GCTAGGCCTCGGCACCTAACCCTGATTTTCATTAGCT3'
3'CGATCCGGAGCCGTGGATTGGGACTAAAAGTAATCGA5'
protected
no information
DNAse I-cut
DNAse I-hypersensitive
* DMS-hypersensitive

Fig. 3.16

showed that TRF1 binds specifically to the template region and produces a footprint centered around the hexamer (Fig. 3.16).

Given the role of TRF1 as a negative regulator of telomere length (van Steensel and de Lange 1997), we explored the possibility that TRF1 may regulate the expression of TR by binding to the template region. However, overexpression of TRF1 (and TRF2) using adenovirus expression vectors failed to cause any variation in the level of expression of hTER (J. Karlseder, A. Bianchi and T. de Lange, unpublished). In addition, arrays of TTAGGG repeats located upstream of a reporter gene failed to affect transcription of the gene when transfected in insect and mammalian cells (A. Bianchi, D. Broccoli, J. Hanish and T. de Lange, unpublished). These results are in agreement with the view that *in vivo* binding activity of TRF1 is achieved by binding with two Myb domains and with a model for telomere length regulation in which TRF1 acts *in cis* at the telomere (van Steensel and de Lange 1997).

Effects of TRF1 on higher order structures: DNA bending

We have previously described that TRF1 is able to introduce a shallow bend on DNA upon binding, using a circular permutation assay (Bianchi et al. 1997). With this assay TRF1 was shown to introduce a shallow distortion in bound DNA, with an estimated deviation form linearity of about 60°. Because of reported artifacts with this assays, due to migration anomalies induced by the protein rather than by the DNA conformation (Gartenberg et al. 1990; McCormick et al. 1996; Sitlani and Crothers 1996), we employed a circularization assay for DNA bending (Kotlarz et al. 1986). The rate of intramolecular ligation of DNA molecules is affected by the presence of a bend in DNA. Because of the intrinsic rigidity of the DNA helix (Shore et al. 1981), the

Fig. 3.17 **Enhanced DNA cyclization by TRF1**

(a) effect of increasing amounts of TRF1 on cyclization of a 217 bp DNA fragment containing 27 TTAGGG repeats. (b) The rate of cyclizationof the 217 bp fragment was measured in the presence of either heat-inactivated or active TRF1. Ligation time is indicated on top of lanes. Exonuclease digestion was performed I order to eliminate linear ligation products. Lanes 1 and 2 show unligated samples. In lanes 3 and 4 samples were ligated with a 20-fold higher amount of ligase.

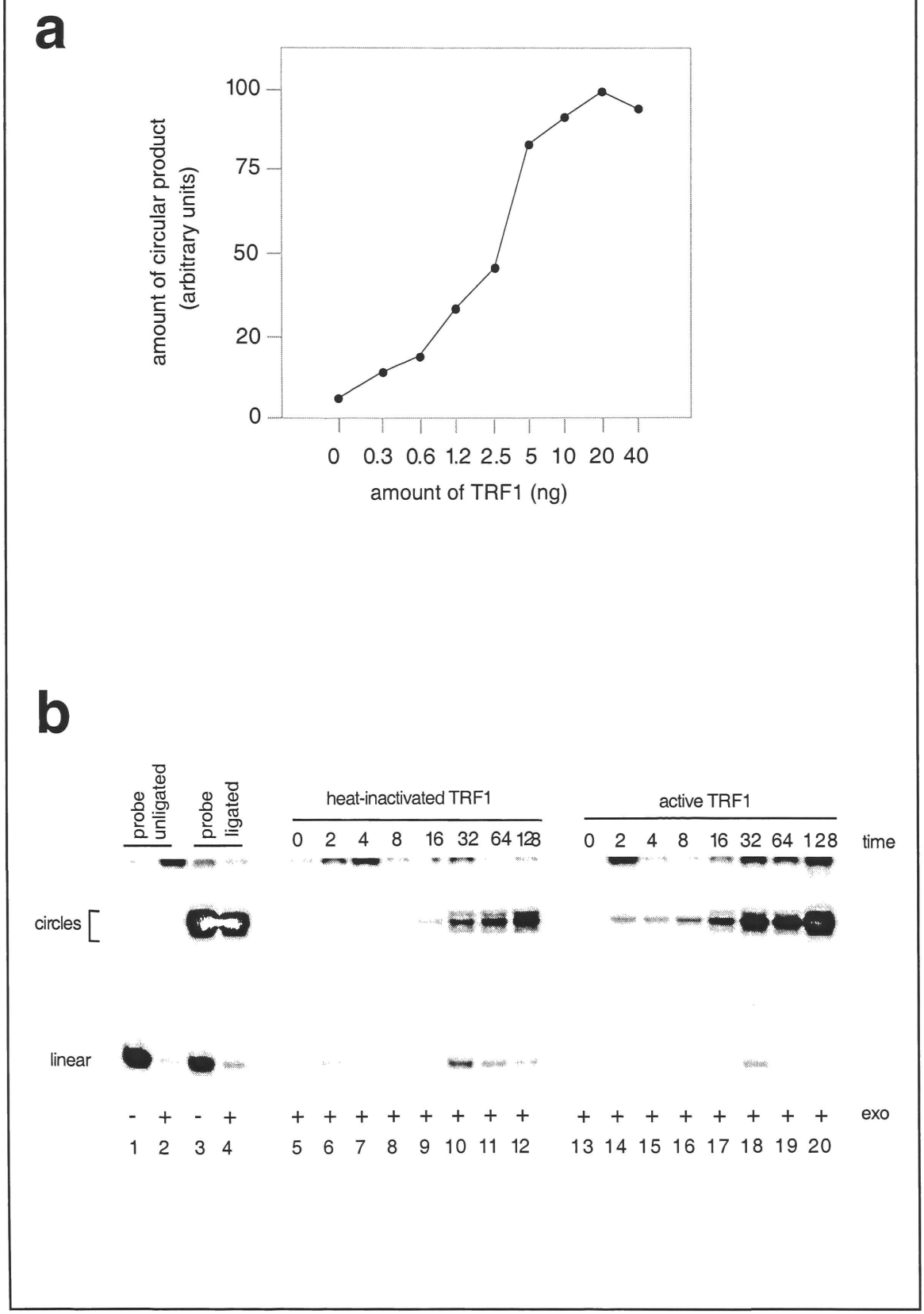

a
amount of circular product
(arbitrary units)
100
75
50
20
0
0 0.3 0.6 1.2 2.5 5 10 20 40
amount of TRF1 (ng)
b
probe unligated
probe ligated
heat-inactivated TRF1
active TRF1
0 2 4 8 16 32 64 128
0 2 4 8 16 32 64 128
time
circles
linear
- + - +
+ + + + + + + +
+ + + + + + + +
exo
1 2 3 4
5 6 7 8 9 10 11 12
13 14 15 16 17 18 19 20

Fig. 3.17

introduction of a bend in the DNA facilitates the interaction of the two DNA ends, which can be measured by changes in the rates of intramolecular ligation. We therefore determined the effect of TRF1 binding to a 217 bp DNA fragment containing 27 tandem TTAGGG repeats. The reaction was monitored by gel electrophoresis of deproteinized DNA samples that were treated with T7 gene 6 exonuclease to visualize only the products of intramolecular (and not intermolecular, see below) ligation. In three independent experiments, an increase in the rate of appearance of the circular ligation product was observed that was dependent on the presence of active TRF1 (Fig. 3.17b, lanes 13-20). No enhancement was observed with heat inactivated TRF1 (Fig. 3.17b, lanes 5-12) or on a 192 bp fragment that did not contain TTAGGG repeats (data not shown). In addition, the amount of circular product obtained was dependent on the amount of TRF1 present (Fig. 3.17a). These results are consistent with the existence of a TRF1-induced distortion in the DNA.

TRF1-induced telomere pairing

EM analysis of the TRF1 complex with DNA conducted in collaboration with Dr. Jack Griffith confirmed the highly specific and non-cooperative binding of the protein (Griffith et al. 1998). In addition these experiments revealed that at higher protein concentrations TRF1 is able to form filaments on arrays of telomeric repeats ($\geq$ 27 repeats), thus completely coating the telomeric sequences. Unexpectedly, TRF1 was shown to be able to promote the pairing of TRF1-coated telomeric DNA with high efficiency, with a preference for a parallel arrangment of the two telomeric tracts. The thickness of the paired filaments suggested that pairing is mediated by protein-protein interactions between the TRF1 molecules bound on each DNA.

Fig. 3.18 **TRF1-induced association of telomeric tracts**

A 217 bp Asp718 fragment containing 27 TTAGGG repeats in tandem arrays was incubated with ligase in the presence of decreasing amounts of TRF1 (from 2000 to 15 ng/ml in 2-fold dilution steps). The position of the dimeric, trimeric, tetrameric and circular ligation products are indicated.

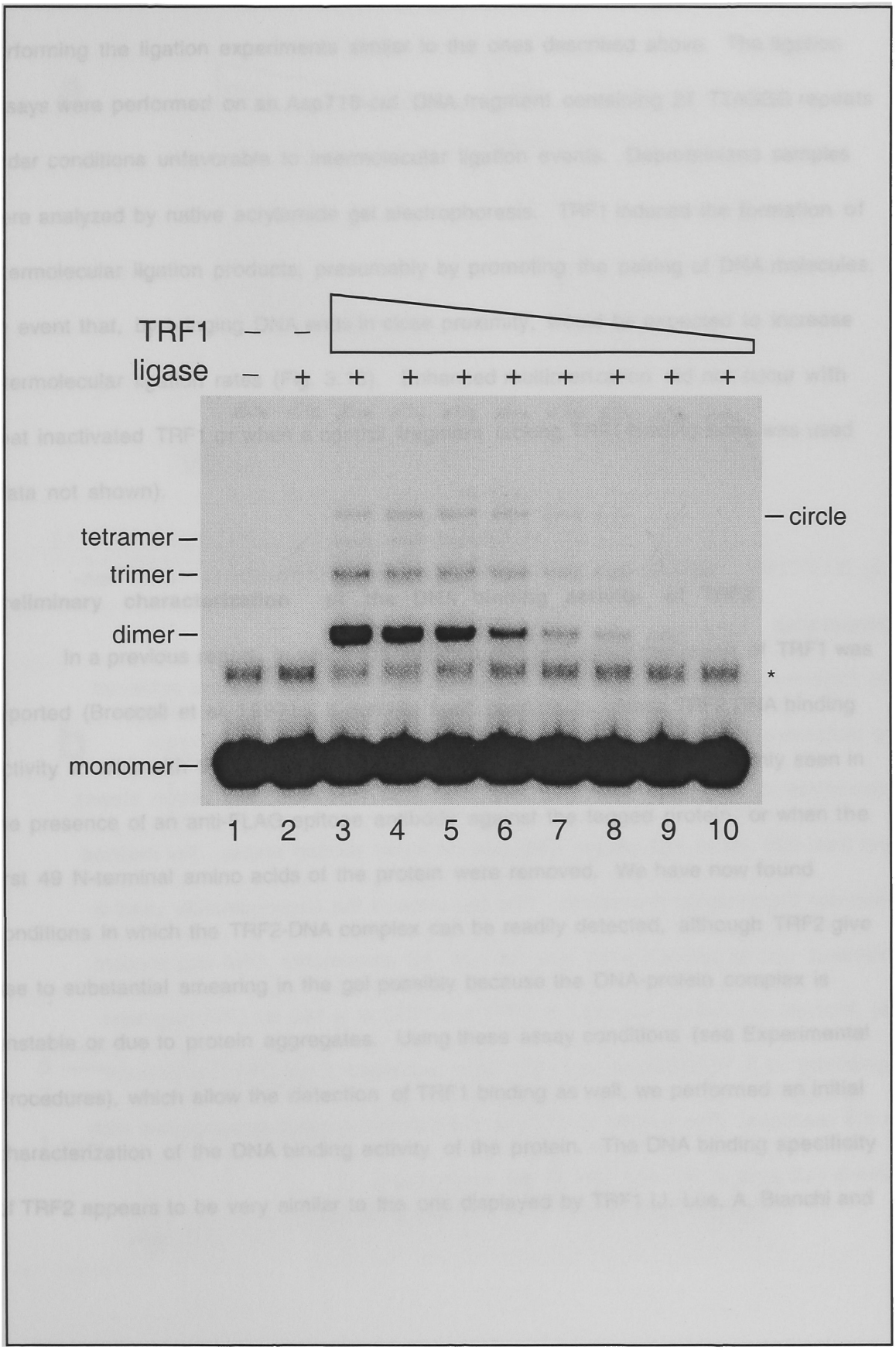

TRF1
ligase
– – + + + + + + + +
tetramer
trimer
dimer
monomer
circle
*
1 2 3 4 5 6 7 8 9 10

Fig. 3.18

We obtained independent evidence for TRF1-induced telomere pairing by performing the ligation experiments similar to the ones described above. The ligation assays were performed on an Asp718-cut DNA fragment containing 27 TTAGGG repeats under conditions unfavorable to intermolecular ligation events. Deproteinized samples were analyzed by native acrylamide gel electrophoresis. TRF1 induced the formation of intermolecular ligation products, presumably by promoting the pairing of DNA molecules, an event that, by bringing DNA ends in close proximity, would be expected to increase intermolecular ligation rates (Fig. 3.18). Enhanced multimerization did not occur with heat inactivated TRF1 or when a control fragment lacking TRF1 binding sites was used (data not shown).

Preliminary characterization of the DNA binding activity of TRF2

In a previous report, in which the isolation and first characterization of TRF1 was reported (Broccoli et al. 1997b), it had not been possible to detect TRF2 DNA binding activity *in vitro* with the full-length protein. A TRF2 complex with DNA was only seen in the presence of an anti-FLAG epitope antibody against the tagged protein, or when the first 49 N-terminal amino acids of the protein were removed. We have now found conditions in which the TRF2-DNA complex can be readily detected, although TRF2 give rise to substantial smearing in the gel possibly because the DNA-protein complex is unstable or due to protein aggregates. Using these assay conditions (see Experimental Procedures), which allow the detection of TRF1 binding as well, we performed an initial characterization of the DNA binding activity of the protein. The DNA binding specificity of TRF2 appears to be very similar to the one displayed by TRF1 (J. Lue, A. Bianchi and

Fig. 3.19 **TRF1 and TRF2 have similar dissociation constants and non-cooperative behavior**

(a) Determination of the dissociation constant for TRF1 and TRF2 was achieved by performing band-shift assays under TRF2 conditions with active protein concentrations ranging from 1.3 μM to 200 pM for TRF1 (in 3-fold dilution steps), and from 800 nM to 122 pM for TRF2 (also in 3-fold dilution steps). For method used see Experimental Procedures. The sequence of the oligonucleotide used is indicated, and its concentration was 100 pM. No competoitor DNA was present.

(b) Analysis of binding beahviour of TRF1 and TRF2 to a 142 bp DNA fragment containing 12.5 TTAGGG repeats. 500 ng of sonicated *E.coli* DNA was present in the reactions. The number of TRF1 or TRF2 dimers present in complexes with DNA is indicated at left and right of gel, respectively.

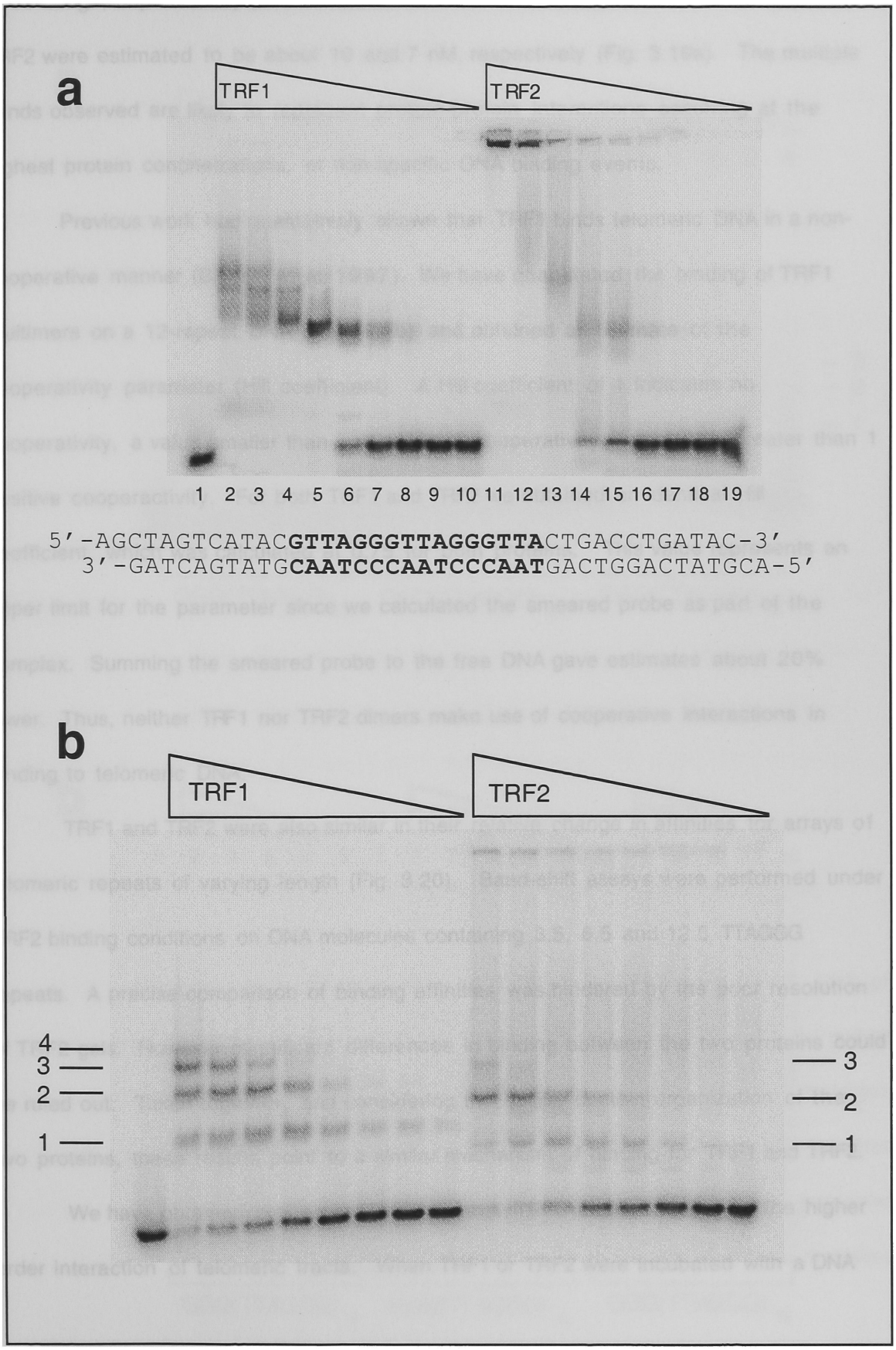

a
TRF1
TRF2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
5′-AGCTAGTCATACGTTAGGGTTAGGGTTACTGACCTGATAC-3′
3′-GATCAGTATGCAATCCCAATCCCAATGACTGGACTATGCA-5′
b
TRF1
TRF2
4
3
2
1
3
2
1

Fig. 3.19

T. de Lange, unpublished) (Broccoli et al. 1997b). Dissociation constants for TRF1 and TRF2 were estimated to be about 10 and 7 nM, respectively (Fig. 3.19a). The multiple bands observed are likely to represent protein-protein interactions occurring at the highest protein concnetrations, or non-specific DNA binding events.

Previous work had qualitatively shown that TRF1 binds telomeric DNA in a non-cooperative manner (Bianchi et al. 1997). We have quantitated the binding of TRF1 multimers on a 12-repeat DNA (Fig. 3.19b) and obtained an estimate of the cooperativity parameter (Hill coefficient). A Hill coefficient of 1 indicates no cooperativity, a value smaller than one negative cooperativity and a value greater than 1 positive cooperactivity. For both TRF1 and TRF2 we obtained an identical Hill coefficient, which was calculated at 0.75 for both proteins. This value represents an upper limit for the parameter since we calculated the smeared probe as part of the complex. Summing the smeared probe to the free DNA gave estimates about 20% lower. Thus, neither TRF1 nor TRF2 dimers make use of cooperative interactions in binding to telomeric DNA.

TRF1 and TRF2 were also similar in their relative change in affinities for arrays of telomeric repeats of varying length (Fig. 3.20). Band-shift assays were performed under TRF2 binding conditions on DNA molecules containing 3.5, 6.5 and 12.5 TTAGGG repeats. A precise comparison of binding affinities was hindered by the poor resolution of TRF2 gels. However, significant differences in binding between the two proteins could be ruled out. Taken together, and considering the similar domain organization of the two proteins, these results point to a similar mechanism of binding for TRF1 and TRF2.

We have obtained preliminary evidence that TRF2 is able to promote the higher order interaction of telomeric tracts. When TRF1 or TRF2 were incubated with a DNA

Fig. 3.20 **Binding of TRF1 and TRF2 to arrays of telomeric repeats** Binding assays were performed under standard TRF2 conditions, with 500 ng sonicated *E.coli* DNA. DNA probes with 3.5, 6.5 and 12.5 repeats were restriction fragments 88, 106 and 142 bp long, respectively. Protein concentrations varied in three-fold steps. The inferred number of dimers present in each DNA complex is indicated at right of gel.

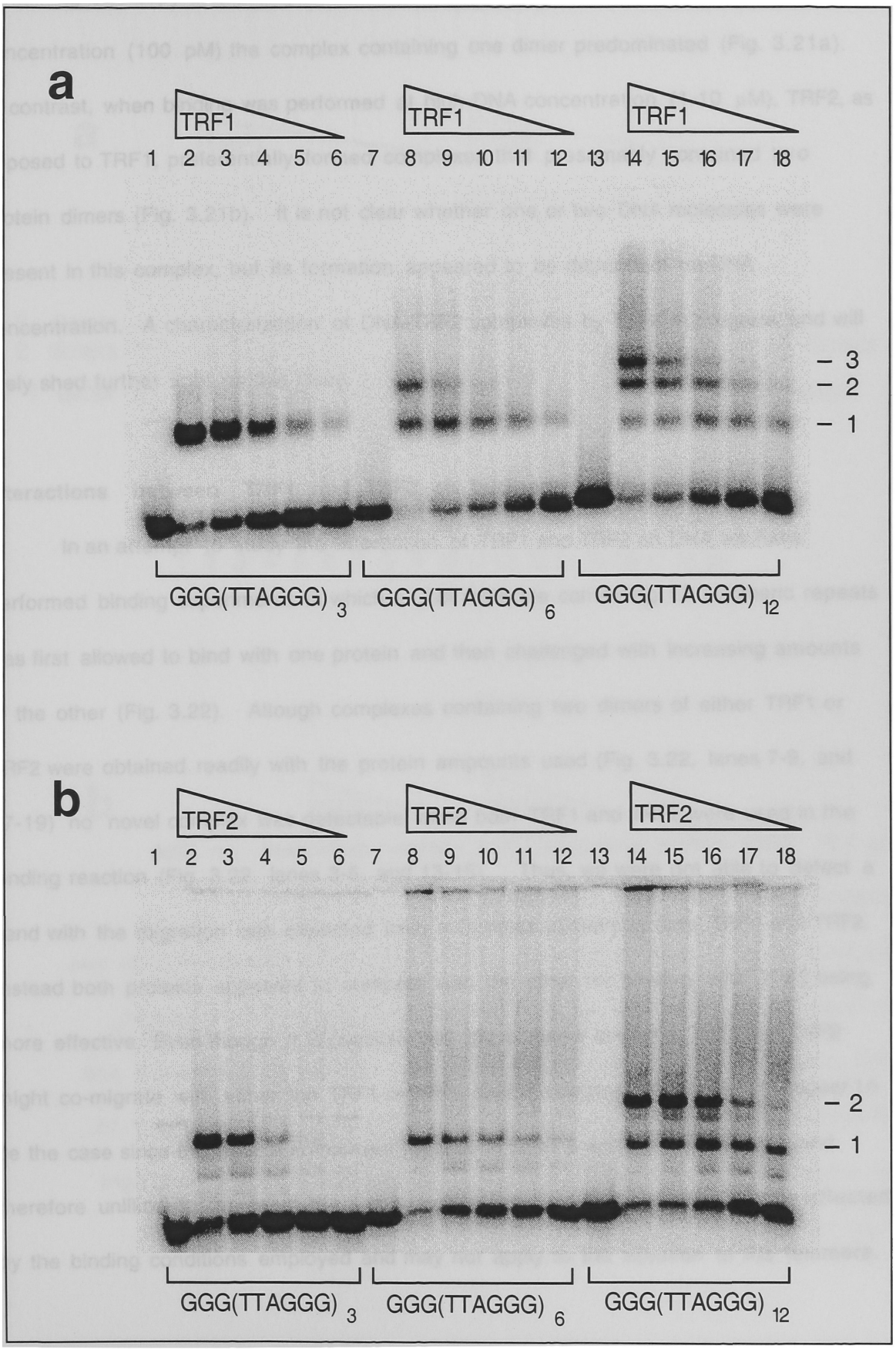

a
TRF1
TRF1
TRF1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
– 3
– 2
– 1
GGG(TTAGGG) 3
GGG(TTAGGG) 6
GGG(TTAGGG) 12
b
TRF2
TRF2
TRF2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
– 2
– 1
GGG(TTAGGG) 3
GGG(TTAGGG) 6
GGG(TTAGGG) 12

Fig. 3.20

containing one copy of the sequence 5'-TTAGGGTTAGGGTTA-3' at low DNA concentration (100 pM) the complex containing one dimer predominated (Fig. 3.21a). In contrast, when binding was performed at high DNA concentration (1-10 μM), TRF2, as opposed to TRF1, preferentially formed complexes that presumably contained two protein dimers (Fig. 3.21b). It is not clear whether one or two DNA molecules were present in this complex, but its formation appeared to be dependent on DNA concentration. A characterization of DNA-TRF2 complexes by EM is in progress and will likely shed further light on this issue.

Interactions between TRF1 and TRF2 on telomeric DNA *in vitro*

In an attempt to study the interaction of TRF1 and TRF2 on DNA we have performed binding experiments in which a DNA molecule containing 12 telomeric repeats was first allowed to bind with one protein and then challenged with increasing amounts of the other (Fig. 3.22). Altough complexes containing two dimers of either TRF1 or TRF2 were obtained readily with the protein ampounts used (Fig. 3.22, lanes 7-9, and 17-19) no novel complex was detectable when both TRF1 and TRF2 were used in the binding reaction (Fig. 3.22, lanes 3-5, and 13-15). Thus, we were not able to detect a band with the migration rate expected from a complex containing both TRF1 and TRF2. Instead both proteins appeared to compete with the other for binding, with TRF2 being more effective. Even though it is possible that the complex containg TRF1 and TRF2 might co-migrate with either the TRF1 or TRF2 double complex, this does not appear to be the case since the double complexes present in lanes 5 and 15 are different and therefore unlikely to represent the same ternary complex. These results may be affected by the binding conditions employed and may not apply to the situation at the telomere.

Fig. 3.21 **TRF2 dimers associate in a manner dependent on DNA concentration**

Gel shown in (a) is the same as the one shown in Fig. 3.19a. Gel shown in (b) was run under identical conditions except that DNA concentration was 10 microM. TRF1 protein concentration in (a) lanes 2-6 are the same as in (b) lanes 2-6; and TRF2 protein concentration in (a) lanes 11-15 are the same as in (b) lanes 7-11. The number of TRF1 and TRF2 dimers in complexes is indicated at left and right of gel, respectively.

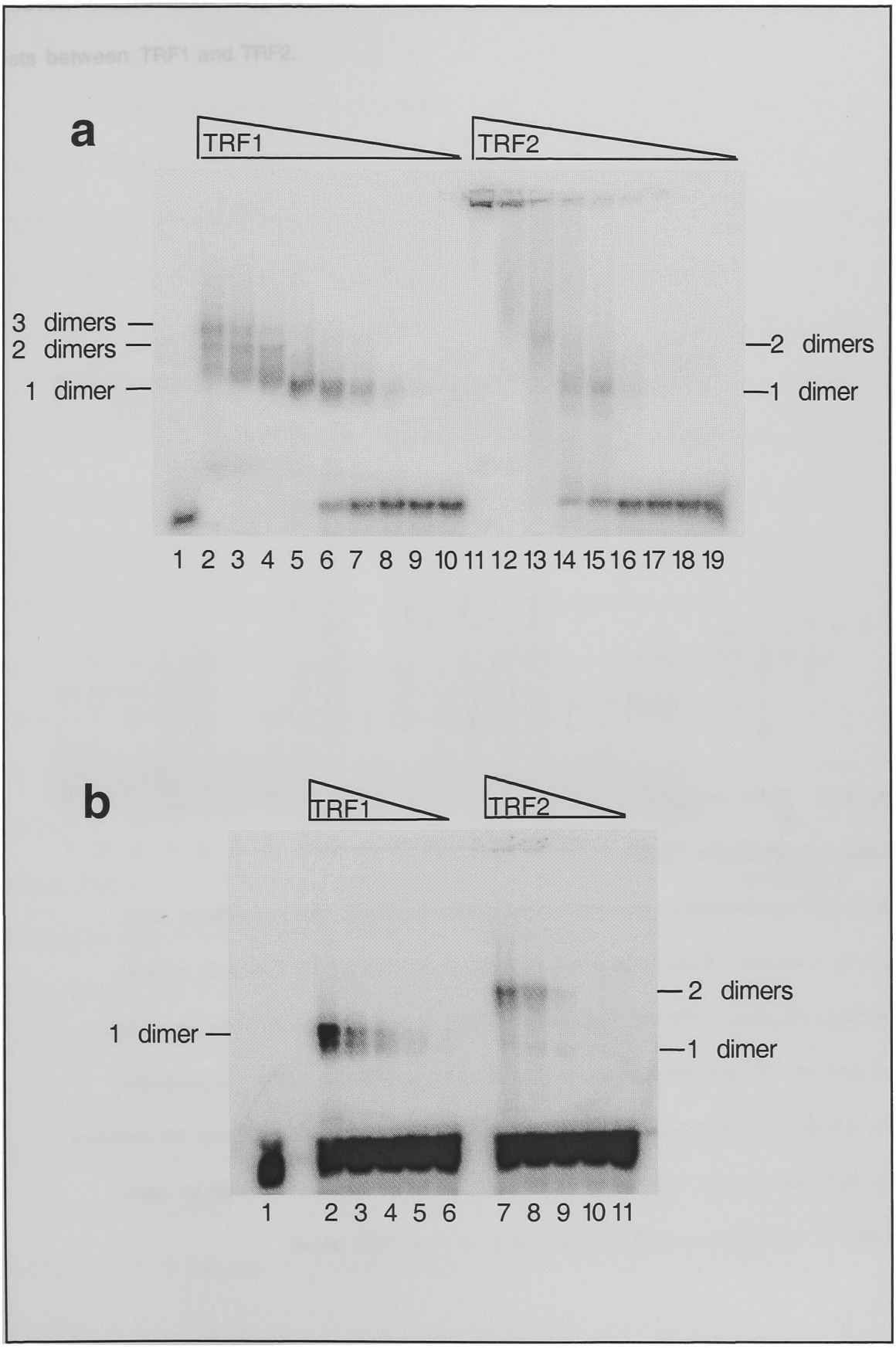

a
TRF1
TRF2
3 dimers
2 dimers
1 dimer
2 dimers
1 dimer
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
b
TRF1
TRF2
1 dimer
2 dimers
1 dimer
1 2 3 4 5 6 7 8 9 10 11

Fig. 3.21

However, at a minimum, they seem to indicate that no cooperativity for DNA binding exists between TRF1 and TRF2.

Fig. 3.22 **TRF1 and TRF2 do not interact cooperatively with each other on telomeric DNA**

Band-shift experiments were performed under standard TRF2 conditions with 500 ng sonicated *E.coli* DNA and a DNA probe containing 12 TTAGGG repeats. Binding with either TRF1 or TRF2 was allowed to proceed for 30 minutes (lanes 1-5 and 11-15) and then an increasing amount of TRF2 or TRF1, respectively, was added and incubations were allowed to continued for an additional 30 minutes. The deduced protein composition of complexes with DNA is indicated at right, where "1" indicates one TRF1 dimer, and "2" one TRF2 dimer.

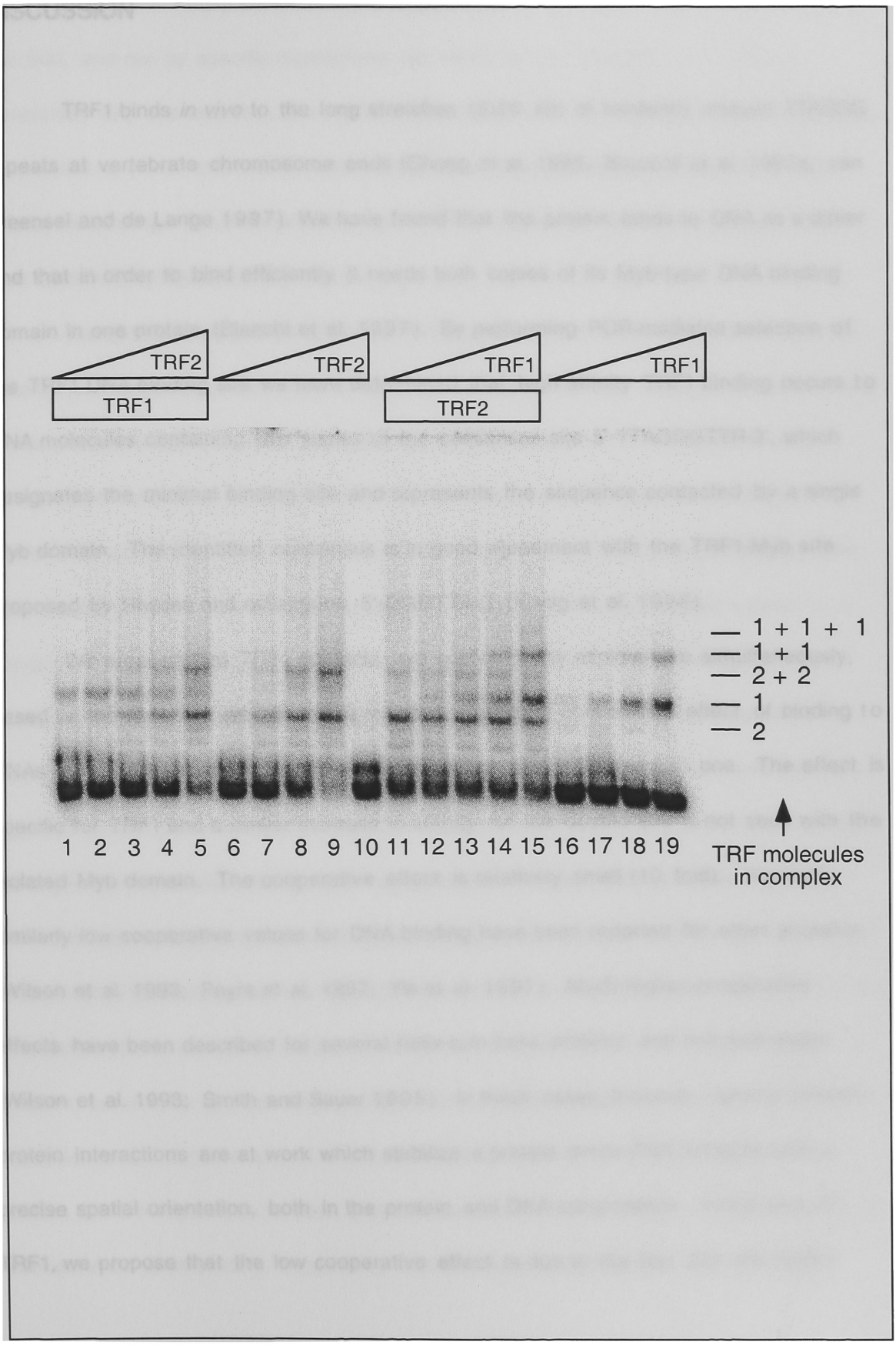

TRF2
TRF1
TRF2
TRF1
TRF2
TRF1
1 + 1 + 1
1 + 1
2 + 2
1
2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
TRF molecules in complex

Fig. 3.22

DISCUSSION

TRF1 binds *in vivo* to the long stretches (2-20 kb) of tandemly arrayed TTAGGG repeats at vertebrate chromosome ends (Chong et al. 1995; Broccoli et al. 1997a; van Steensel and de Lange 1997). We have found that the protein binds to DNA as a dimer and that in order to bind efficiently it needs both copies of its Myb-type DNA binding domain in one protein (Bianchi et al. 1997). By performing PCR-mediated selection of the TRF1 DNA binding site we have determined that high affinity TRF1 binding occurs to DNA molecules containing two copies of the consensus site 5'-YTAGGGTTR-3', which designates the minimal binding site and represents the sequence contacted by a single Myb domain. The identified consensus is in good agreement with the TRF1-Myb site proposed by Rhodes and colleagues, 5'-GGGTTA-3' (Konig et al. 1998).

We propose that TRF1 contacts two copies of the minimal site simultaneously, based on the fact that we observe a mild but significant cooperative effect of binding to DNAs containing two copies of the minimal binding site as opposed to one. The effect is specific for TRF1 and a similar increase in affinity for the double site is not seen with the isolated Myb domain. The cooperative effect is relatively small (10 fold). However, similarly low cooperative values for DNA binding have been reported for other proteins (Wilson et al. 1993; Payre et al. 1997; Yie et al. 1997). Much higher cooperative effects have been described for several helix-turn-helix proteins and homeodomains (Wilson et al. 1993; Smith and Sauer 1995). In these cases, however, specific protein-protein interactions are at work which stabilize a protein dimer-DNA complex with a precise spatial orientation, both in the protein and DNA components. In the case of TRF1, we propose that the low cooperative effect is due to the fact that the higher

affinity is created simply by an increase of proximity of the second Myb upon binding of the first, and not by specific interactions between the two domains. This view is consistent with the fact that the two half sites contacted by the dimer can be arranged in inverted or direct orientation, and with varying spacing, without strongly affecting affinity of binding. Thus we propose that TRF1 contacts the two half sites with each Myb domain in an independent manner.

The spatial arrangement of the two TRF1 monomers in the dimer remains to be determined. In other dimeric DNA binding factors symmetrically arranged monomers are more commonly encountered than asymmetric configurations. If TRF1 dimers also interact symmetrically, one of the two Myb domains would be required to be free to swivel 180° to contact the second (tandemly arranged) half site at the telomere. The poorly conserved region of TRF1 that spans amino acids 265-376 seems a good candidate for a structural hinge between the dimerization and DNA-binding domains. Assuming a length of about 3 Å for an unstructured amino acid, this domain could theoretically span a length of about 33 nm, at least in principle more than enough to span the length of the more distantly spaced sites tested (56 bp). In the case of TRF1 binding to inverted sites, the protein would remain oriented in a symmetrical manner, while the built-in flexibility used to recognize symmetric sites would allow to the protein to have little constraints for the spacing between the sites. This model accounts for the poor sequence conservation between human and mouse TRF1 in the "hinge" region. In addition, proteolysis analysis of the protein has revealed that two trypsin-sensitive sites exist, one at the N-terminal end of the dimerization domain and one in the "hinge" domain (M. O'Reilly and D. Rhodes, personal communication). The protease sensitivity is retained upon DNA binding, suggesting that the region may remain unstructured in

bound TRF1, a result that is in agreement with circular dichroism analysis (L. Fairall and D. Rhodes, personal communication).

Several examples have been reported of proteins capable of changing the relative orientations of DNA binding domain due to the presence of structurally flexible linker regions. A flexibility in the way p53 contacts half sites (consecutive or separated by one or two turns) has been observed and ascribed to the presence of a flexible short linker between its dimerization and DNA binding domains (Arrowsmith and Morin 1996). The presence of a flexible linker (23 amino acids long) has been shown to allow the relative rotation and orientation of the intramolecularly linked POU domains of Oct-1, with effects on DNA specificity (van Leeuwen et al. 1997). In one POU domain family member (the *C.elegans Ceh-18* gene product), the linker is as long as 57 amino acids (Greenstein et al. 1994). Other examples of DNA binding proteins with at least some degree of flexibility in DNA binding are the *E. coli* CytR regulator, the maize Activator transposase, and the yeast α2 repressor (Sauer et al. 1988; Becker and Kunze 1997; Jorgensen et al. 1998). The yeast protein HAP1 dimerizes symmetrically but contacts two tandemly repeated sites, presumably by rotation around a flexible linker (Zhang and Guarente 1996). In the case of HAP1, as for some of the nuclear receptors (that also contact tandem repeats), the rotational flexibilty is coupled to asymmetric protein-protein interactions that take place in the dimerization domain (Kurokawa et al. 1993; Rastinejad et al. 1995).

Thus, the structural flexibility proposed for TRF1 is not unprecedented. Yet, the model remains tentative in the absence of more direct evidence. In the course of EM studies on TRF1-DNA complexes we have failed to observe looping or shortening of the DNA (Griffith et al. 1998). However, the experiments were always conducted on

continuous arrays of TTAGGG repeats, and it is possible that under these conditions the two Myb domains tend to occupy proximal, or close, sites in order to maximize occupancy of the DNA. Therefore, DNA substrates need to be tested that bear only two half sites and at an appropriate distance to visualize looping. Similarly, band-shift experiments might be performed at high DNA concentration with DNA probes harboring one copy of the half site, in an effort to detect ternary complexes containing one TRF1 dimer bound to two DNA molecules.

An alternative model for TRF1 binding in which the cooperative effect is a result of the close proximity of the second, but unengaged DNA binding domain, cannot at present be ruled out. However, this model does not explain why a preferential TRF1-specific increase in binding affinity is not seen as the number of minimal sites intramolecularly linked increases above 2.

If the two Myb domains in the dimer are engaged in binding at the same time, a change in the stability of the complex should be expected. We have not been able to observe differences in the half-lives of the TRF1 and Myb complexes with DNA, because they are too short too be measured be measured by competition experiments in band-shift assays. It remains possible that a significant difference in off-rates exists but that it still too low and remains unmeasurable by electrophoretic approaches. In support of this hypothesis we note that, while TRF1 complexes are readily observed when gel electrophoresis is carried out at room temperature, Myb domain complexes appear very unstable under these conditions and are only efficiently detected at 4°C (A. Bianchi and T. de Lange, unpublished). Possibly, this is a reflection of differences in the off-rates of the two proteins.

Both DNA-ligation assays and EM analysis indicate that TRF1 has the ability to promote the association of telomeric DNA (Griffith et al. 1998). It remains to be demonstrated whether TRF1-promoted telomere associations occur *in vivo*. However, in support of this idea, telomeres visualized *in vivo* by immunogold electron microscopy appear to be very compact structures (Ludérus et al. 1996). In addition, natural telomere sequences are much longer than the ones that have been used in the *in vitro* assays, and therefore could be even more amenable to form associations. Intermolecular telomere associations might have important consequences on the nuclear localization of telomeres, in particular with regard to the clustering of telomeres observed in the early stages of meiotic division (Dernburg et al. 1995; de Lange 1998a). Intramolecular associations at the telomere, which we did not observe *in vitro* but that could conceivably take place on longer TTAGGG tracts, would result in a looping back of the DNA on itself with consequent compacting of the structure. This altered chromatin structure may be inaccessible by elongation by telomerase and thus represent the mechanism by which TRF1 negatively regulates telomere length (van Steensel and de Lange 1997) (see Chapter 4).

Telomeric DNA in human cells represents typically about 0.2-0.4% of the total genomic DNA. We have shown that each Myb domain binds to DNA very specifically with absolute sequence requirements within a hexameric sequence and additional strong base preferences (two of which degenerate) in three adjacent positions. In addition we have shown that for effective DNA binding both *in vitro* and *in vivo* TRF1 needs to contact two of these minimal binding sites. Thus the full TRF1 binding site is 15-18 bp long, a fact that should ensure great specificity of binding *in vivo*. Another factor that seems likely to ensure an effective sequestering of TRF1 at the telomere is the high local

concentration of sites, which, as we have documented in Surface Plasmon Resonance experiments, might create an ideal local environment for rapid rebinding events.

The available evidence and the similar domain organization suggest that TRF2 may bind DNA in a manner similar to TRF1. However, TRF2 is, at least partly, functionally distinct from TRF1. In particular, the protein has a protective function on G-overhangs (van Steensel et al. 1998). Presently, we have no evidence that this function may be carried out by means of a single-stranded DNA binding activity of TRF2, as none was detected by band-shift assays under a variety of conditions and by chemical or enzymatic probing (A. Bianchi, D. Rhodes and T. de Lange, unpublished). It will be of interest to determine if a differential localization of TRF1 and TRF2 exists at different telomeric "subdomains". Our *in vitro* experiments suggests that there is no positive cooperativity in binding between the proteins, and that negative cooperactivity is a possibility. Thus, an alternating arrangement of TRF1 and TRF2 dimers at the telomere seems an unlikely possibility, whereas it is possible that binding occurs randomly or in clusters.

Chapter 4

Interaction of telomerase with *in vitro* substrates containing telomeric proteins

SUMMARY

Telomere length maintenance in mammalian cells is primarily dependent on telomerase. Telomerase catalyzes the addition of telomeric repeats on a single-stranded DNA substrate in an RNA-templated manner. Several DNA binding activities found at telomeres in mammals, budding and fission yeast act as negative regulators of telomere length, possibly by modulating directly or indirectly the action of telomerase at individual telomeres. These activities include TRF1 and TRF2 in human cells. We have tested by an *in vitro* assay whether TRF1 or TRF2 can inhibit telomerase at the DNA terminus. Our results show that this is not the case, suggesting that these proteins may modulate telomerase through accessory factors or formation of a higher order chromatin structure at the telomere.

INTRODUCTION

Although pathways for telomere length control based on recombination exist both in yeast and in mammalian cells (Pluta and Zakian 1989; Lundblad and Blackburn 1993; Bryan et al. 1995; Li and Lustig 1996; McEachern and Blackburn 1996), telomere maintenance is primarily dependent on the activity of telomerase, a ribonucleoprotein with template-dependent DNA polymerase activity (see (Greider 1995; Greider 1996) for review). The RNA component of telomerase contains a short sequence (8-29 nucleotides) complementary to the telomeric G-strand, that serves as the template for DNA synthesis (Greider and Blackburn 1989; Shippen-Lentz and Blackburn 1990; Singer and Gottschling 1994; Blasco et al. 1995; Feng et al. 1995). Although a few telomerase-associated proteins have been described and possibly more remain to be discovered (Collins et al. 1995; Lingner and Cech 1996; Bednenko et al. 1997; Harrington et al. 1997a; Meyerson et al. 1997; Nakayama et al. 1997), one single peptide bearing homology to the catalytic domain of reverse transcriptases (Counter et al. 1997; Harrington et al. 1997b; Lingner et al. 1997b; Meyerson et al. 1997; Nakamura et al. 1997; Nakamura and Cech 1998), together with the RNA may be sufficient for reconstitution of telomerase activity *in vitro* (Weinrich et al. 1997; Beattie et al. 1998).

Even though telomerase was originally identified based on its ability to add telomeric repeats to a single-stranded telomeric primer *in vitro* (Greider and Blackburn 1985), it was later shown that non-telomeric primers also function in this assay (Harrington and Greider 1991; Morin 1991; Melek et al. 1996; Wang and Blackburn 1997). In fact, the sequence requirements for primer elongation are not very strict and not fully understood. Experiments conducted with permutated and antisense primers

and with mutated template RNAs have established that only one segment of the template RNA sequence (at the 5' end) is used as for templating activity while the rest is used for primer alignment (Greider and Blackburn 1989; Shippen-Lentz and Blackburn 1990; Autexier and Greider 1994).

In vitro, telomerase is able to synthesize many repeats on a single primer molecule in a highly processive manner (Greider 1991). This is achieved by a translocation mechanism in which the elongated DNA, having reached the end of the template region, dissociates from the RNA and, in a translocation step, reassociates at the beginning of the template sequence for a subsequent round of synthesis. Telomerase processivity *in vitro* is dependent on reaction conditions, one critical parameter being primer length. Long primers (about two repeats-long or more) result in processive elongation, whereas accumulation of short products is observed with short primers (Morin 1989; Collins and Greider 1993; Lee and Blackburn 1993). Since a short primer is rapidly converted by telomerase into a long one, it could as a consequence be expected to give rise to processive elongation. To explain these observations, it has been proposed that telomerase possesses a second ("anchor") binding site that contacts the 5' region of the primer stabilizing the interaction of the DNA with the enzyme and thus allowing several rounds of synthesis to take place before dissociation occurs. This model postulates that the newly synthesized DNA is looped out, between the two binding sites, and indirect evidence for such a loop has been obtained by crosslinking studies (Hammond et al. 1997).

DNA conformation and additional protein factors may contribute to regulate telomerase activity *in vivo*. For example, experiments in which a mutated RNA was introduced in *Tetrahymena* cells that conserved also the wild type copy, showed that

mutated and normal telomeric repeats interspersed randomly and at high frequency at telomeres (Yu et al. 1990). This suggests that telomerase may act non-processively *in vivo*, although an alternative explanation based on template switching due to the fact that telomerase appears to be a dimer is also possible (Prescott and Blackburn 1997). If true, discrepancies in processivity between the *in vitro* and *in vivo* situation could possibly reflect the presence of modulating factors or perhaps a difference in the DNA substrate with respect to conformation and accessibility. In all cases examined telomerases appear to be unable to elongate blunt-ended double-stranded DNA but variations are observed with respect to the minimal required length of the overhang, which ranges from 4-6 and 13-20 nucleotides for *Euplotes* and *Tetrahymena* telomerase, respectively, depending on sequence (Lee et al. 1993; Lingner and Cech 1996; Wang and Blackburn 1997). At present it is unclear whether only single-stranded DNA can be bound at the anchor site, or double-stranded DNA can bind as well. The situation may be different for telomerases from different species. Whereas *Euplotes aediculatus* telomerase can be crosslinked to double-stranded DNA in the anchor site (Hammond et al. 1997), no increase of processivity with partially double-stranded substrates has been observed for *Tetrahymena* telomerase (Lee et al. 1993) and actually double-stranded DNA may decrease the interaction of this telomerase with the primer (Wang and Blackburn 1997). No studies on partially double-stranded DNA substrates have been reported for human telomerase.

Telomere regeneration experiments in human cells suggest that the length of individual telomeres is determined *in cis*, presumably by telomere binding factors (reviewed in (Smith and de Lange 1997), see (van Steensel and de Lange 1997)). More directly, telomere directed Rap1p-tethering studies performed in budding yeast have

shown that Rap1p affects telomere length by means of a protein-counting mechanism (Marcand et al. 1997). How the number of telomeric proteins present at the telomere generates a signal that inhibits telomerase remains to be established. In mammalian cells the telomeric protein TRF1 and TRF2 (Chong et al. 1995; Bilaud et al. 1996, 1997; Broccoli et al. 1997b), act as negative regulators of telomere length ((van Steensel and de Lange 1997), A. Smogorzewska, B. van Steensel and T. de Lange, unpublished results). The effect of TRF1 and TRF2 on telomerase may be indirect and achieved through the recruitment of additional factors at the telomere. We have tested the alternative possibility that the effect may be achieved by direct inhibition of telomerase *in cis*. To this end we performed standard telomerase assays *in vitro* on DNA substrates bound by TRF1 or TRF2. Our results suggest that the effect of TRF1 and TRF2 on telomere length is either indirect or, if direct, mediated by some higher order chromatin structure.

RESULTS

TRF1 and TRF2 do not affect telomerase activity *in vitro*

In order to test the possibility that TRF1 and TRF2 could potentially directly affect telomerase activity *in cis* at a telomere terminus, we constructed an *in vitro* system with semi-purified components. N-terminally 6-Histidine-tagged TRF1 and TRF2 were expressed in baculoviruses and purified on Ni-column to apparent homogeneity as judged from coomassie staining of SDS-PAGE gels. The full activity of both TRF1 and TRF2 in protein preparations was confirmed by performing binding assays at high (1-10 μM) DNA concentration. Telomerase was partially purified from HeLa cells and provided by Dr. Gisela Schnapp (Boehringer Ingleheim, Germany). The increase of specific activity in telomerase preparations over the starting material in the purification was about 100 fold and total protein concentration was about 1 mg/ml.

Double-stranded DNA molecules were generated that have a TRF1 site (presumably the same as a TRF2 site) at one end, with the same strand polarity that is found at a natural telomere (Fig. 4.2). A short (4 nucleotides) telomeric overhang was present in the G-strand just adjacent to the TRF1-2 binding site. A snap-back type DNA molecule with the ability to self-anneal was used, to ensure both correct strand annealing and that only one 3' end would be available as a substrate for telomerase activity.

Conditions were established under which the amount of synthesis of telomeric repeats was dependent on the concentration of primer in an approximately linear manner (Fig. 4.1a). Having thus chosen the primer concentration for the telomerase assay (10

Fig. 4.1a **Effect of DNA substrate concentration on telomerase**

Standard telomerase assay on primer SB2 (see Fig. 4.2) at the indicated concentrations. Telomerase reactions are as described in Experimental Procedures. Gel at left contained formamide (see Experimental Procedures) whereas gel at right did not.

Fig. 4.1b **DNA binding assay for TRF1 and TRF2 under conditions of telomerase assay**

Band shift assay were in 0.6% agarose, 0.1xTBE. Binding reactions were as for telomerase assays except that dNTPs were omitted. TRF1 and TRF2 concentrations are as indicated on top of gel.

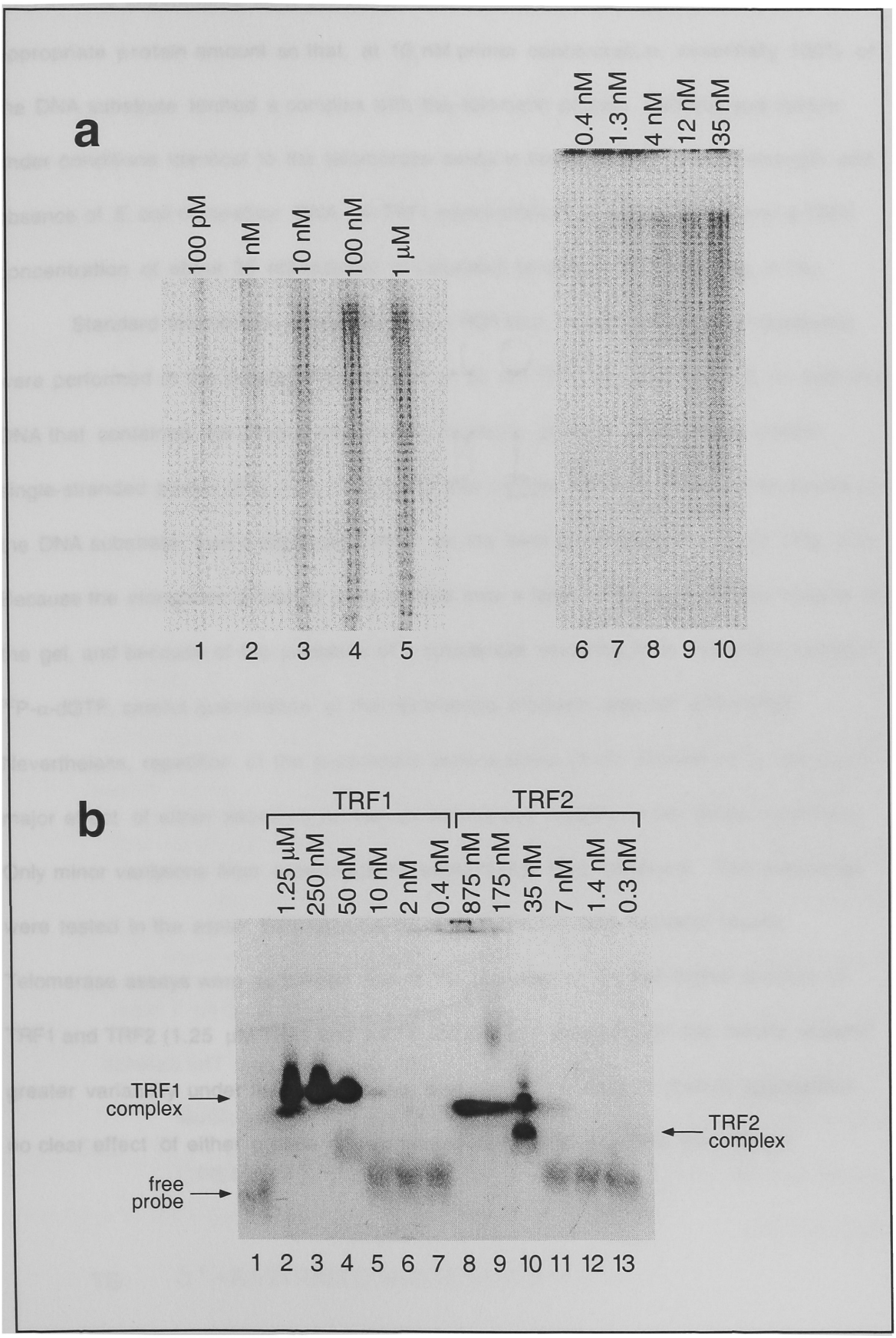

a
100 pM
1 nM
10 nM
100 nM
1 µM
0.4 nM
1.3 nM
4 nM
12 nM
35 nM
1 2 3 4 5 6 7 8 9 10
b
TRF1
TRF2
1.25 µM
250 nM
50 nM
10 nM
2 nM
0.4 nM
875 nM
175 nM
35 nM
7 nM
1.4 nM
0.3 nM
TRF1 complex
free probe
TRF2 complex
1 2 3 4 5 6 7 8 9 10 11 12 13

Fig. 4.1

nM), binding experiments were performed with both TRF1 and TRF2 to determine the appropriate protein amount so that, at 10 nM primer concentration, essentially 100% of the DNA substrate formed a complex with the telomeric protein. Binding was tested under conditions identical to the telomerase assay in terms of buffer, ionic strength and absence of *E. coli* competitor DNA. A TRF1 concentration of about 50 nM and a TRF2 concentration of about 35 nM resulted in saturated binding to the DNA (Fig. 4.1b).

Standard telomerase assays (lacking a PCR step, see Experimental Procedures) were performed in the presence or absence of 50 nM TRF1 and 35 nM TRF2, on both the DNA that contained the binding site for the telomeric proteins (SB2) and a control single-stranded primer (Fig. 4.2). We found that neither TRF1 nor TRF2, when bound to the DNA substrate, had a significant effect on the level of telomerase activity (Fig. 4.2). Because the elongation products were spread over a large range of molecular weights on the gel, and because of the presence of a substantial smearing from the unincorporated ^{32}P-α-dGTP, careful quantitation of the telomerase products was not attempted. Nevertheless, repetition of the experiment several times (n=7) allowed us to rule out a major effect of either telomeric protein on telomerase activity under these conditions. Only minor variations from experiment to experiments were observed. Two overhangs were tested in the assay, 5'GTTAGGGTT3' and 5'GGTT3', with identical results. Telomerase assays were performed also in the presence of 25 fold higher amount of TRF1 and TRF2 (1.25 μM TRF1 and 0.877 μM TRF2). Even though the results showed greater variability under these conditions, presumably because of protein aggregation, no clear effect of either protein on telomerase was observed (data not shown).

Fig. 4.2 **Effect of TRF1 and TRF2 on telomerase activity *in vitro*** Standard telomerase assays with unlabeled primer and P^{32}dGTP were performed on primers SB2 and TS in the presence or absence of telomeric proteins. Primer concentration was 10 nM and protein concentration was about 50 nM for TRF1 and about 35 nM for TRF2. TRF1 and TRF2 were incubated with the oligonucleotide in telomerase buffer for 30 minutes at room temperature. Then telomerease was added, and samples were incubated at 30°C for 1 hour. Reactions were stopped, deproteinized and run on denaturing gel. The asterisk indicates a band that was dependendent on incubation with baculovirus-purified proteins and run a few nucleotides smaller than the untreated oligonucleotide.

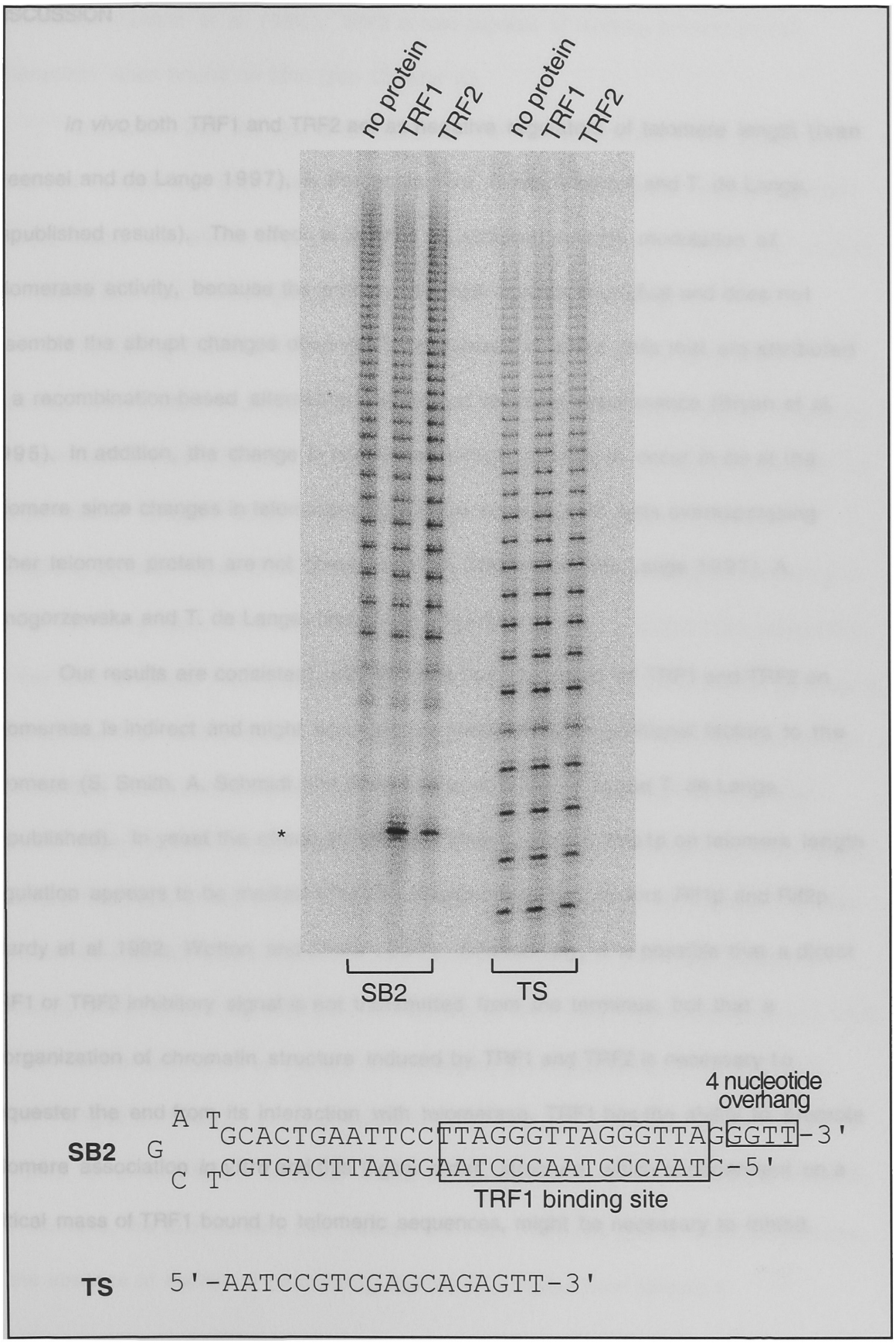

no protein
TRF1
TRF2
no protein
TRF1
TRF2
*
SB2
TS
4 nucleotide overhang
SB2
A T
G
C T
GCACTGAATTCCTTAGGGTTAGGGTTAGGGTT-3'
CGTGACTTAAGGAATCCCAATCCCAATC-5'
TRF1 binding site
TS
5'-AATCCGTCGAGCAGAGTT-3'

Fig. 4.2

DISCUSSION

In vivo both TRF1 and TRF2 act as negative regulators of telomere length ((van Steensel and de Lange 1997), A. Smogorzewska, B van Steensel and T. de Lange, unpublished results). The effect is likely to be achieved through modulation of telomerase activity, because the pattern of length change is gradual and does not resemble the abrupt changes observed in telomerase negative cells that are attributed to a recombination-based alternative pathway of telomere maintenance (Bryan et al. 1995). In addition, the change in telomerase activity is likely to occur *in cis* at the telomere since changes in telomerase activity in extracts from cells overexpressing either telomere protein are not observed ((van Steensel and de Lange 1997), A. Smogorzewska and T. de Lange, unpublished results).

Our results are consistent with the idea that the effect of TRF1 and TRF2 on telomerase is indirect and might occur via the recruitment of additional factors to the telomere (S. Smith, A. Schmidt and T. de Lange, in press; B. Li and T. de Lange, unpublished). In yeast the effect of telomere binding protein Rap1p on telomere length regulation appears to be mediated by two Rap1p-interacting factors Rif1p and Rif2p (Hardy et al. 1992; Wotton and Shore 1997). Alternatively, it is possible that a direct TRF1 or TRF2 inhibitory signal is not transmitted from the terminus, but that a reorganization of chromatin structure induced by TRF1 and TRF2 is necessary to sequester the end from its interaction with telomerase. TRF1 has the ability to promote telomere association *in vitro* and this higher order structure, which is dependent on a critical mass of TRF1 bound to telomeric sequences, might be necessary to inhibit

telomerase (Griffith et al. 1998). TRF2 is also capable of forming protein-protein interaction when bound on DNA (see Chapter 3).

Consistent with the lack of an effect of TRF1 on telomerase in our assay, TRF1 does not co-sediment with telomerase activity on glycerol gradients. Similarly, a TRF1 affinity matrix was unable to preferentially retain telomerase RNA or telomerase activity (Dr. Gisela Schnapp, Boehringer Ingelhelm, personal communication).

The finding that TRF1 and TRF2 molecules positioned so closely to the 3' terminus were unable to prevent telomerase elongation was, to some extent, surprising. The telomere protein α/β from *Oxytricha nova* has been recently reported (contrarily to earlier results (Shippen et al. 1994)) to be able to efficiently prevent terminus elongation by telomerase (Froelich-Ammon et al. 1998). In the α/β instance binding is to the overhang itself, rather than to the duplex, and this may prevent DNA recognition by the RNA template. Alternatively it is possible that the DNA in the α/β-DNA complex is exposed but that telomerase cannot contact it because of steric hindrance. We cannot rule out that steric hindrance between TRF1 or TRF2 and telomerase occurs in our system. Given the low abundance of telomerase protein in our fractions we were unable to assess the state of the TRF-DNA complex after incubation with telomerase fractions because we estimate only a minor fraction of the DNA present to be elongated and bound by the enzyme. We note that the half-life of TRF1 and TRF2 complexes with DNA is much shorter compared to the α/β-DNA complex (see Chapter 3, (Price and Cech 1989; Fang et al. 1993)), and it is therefore possible that human telomerase is able to compete for binding to the terminus under our assay conditions. Nevertheless, even in this scenario, the fact remains that neither TRF1 nor TRF2 are able to block telomerase in the absence of additional (stabilizing) factors or a higher level structure.

The single-stranded overhangs found at mammalian telomeres (Makarov et al. 1997;. McElligot and Wellinger 1997; Wright et al. 1997) are much longer than the 4-nucleotide overhang present in the substrate we tested. We chose a short overhang because we believe it should represent a more stringent test for a direct effect of TRF1 or TRF2 on telomerease activity at the telomere terminus. Since neither TRF1 nor TRF2 appear to posses DNA binding activity, it seems unlikely that they may influence telomerase action on DNA substrates with longer overhangs in the absence of accessory factors.

Collins and Greider have noted that small primers bind *Tetrahymena* telomerase with much lower affinity compared to long ones: more than 10 μM of TTGGGG is needed to compete for 25 nM $(TTGGGG)_3$ (Collins and Greider 1993). This result suggests that binding at the anchor site contributes to overall affinity of telomerase for the DNA. We have observed that fully single-stranded primers are elongated more efficiently by human telomerase compared to partially double-stranded ones (A. Bianchi and T. de Lange, unpublished). This suggests that the anchor site of human telomerase has a higher affinity for single-stranded over double-stranded DNA. If a small change in affinity of telomerase for our primer occurs when the anchor site is occupied by TRF1 or TRF2, then a comparatively small change in telomerase activity might take place. A small change in telomerase affinity would go undetected in our assay but could have more profound implications *in vivo* where TRF1-induced telomere shortening is slow ((van Steensel and de Lange 1997), A. Smogorzewska, B. van Steensel and T. de Lange, unpublished results) and where, in general, regulatory mechanisms are more finely tuned.

Conclusion and perspectives

Single-stranded DNA binding activities: a capping function at vertebrate telomeres?

Ciliate proteins have offered, since their discovery in 1986 (Gottschling and Zakian 1986), the paradigm for end binding activities. The tenacious binding mode of these factors makes them seem ideally suited for a protective function. However, they bind to telomeres that are perhaps unique in their being extremely short and of homogeneous configuration. As in the last few years we have learned more about the vertebrate telomeric complex, the possibility that protein factors may carry out a protective role at vertebrate telomeres in a manner analogous to ciliates seems to be more remote. The existence of long single-stranded tails at mammalian telomeres (Makarov et al. 1997; McElligott and Wellinger 1997; Wright et al. 1997) raises the possibility that a multimeric protein complex could be bound at the telomere terminus. A vertebrate capping factor with 3'-end binding requirement is not precluded by these considerations. However, if this factor occurs one needs to invoke either the presence of additional factors or the ability of the end factor to bind internally on the overhang, possibly in a manner facilitated by cooperative interactions.

ATEF remains the best candidate for such a terminal protein at vertebrate telomeres. Its preference for an adjacent duplex DNA is not very stringent, thus suggesting that ATEF may bind at the 3' end of a long G-strand overhang ((Cardenas et al. 1993) and A. Bianchi and T. de Lange, unpublished). Our work, with the development of an affinity column step, coupled with more recent advances in nanoelectrospray protein microsequencing techniques, may render the future biochemical purification of this activity feasible (Wilm et al. 1996). However, the identification of some of the components of the mammalian telomere has opened an alternative route to the isolation

of telomeric proteins, by taking advantage of isolation techniques based on protein-protein interactions or of telomere-specific immunoreagents. The use of these methods appears today to be the most powerful and promising point of entry to the biochemical characterization of the vertebrate telomere terminus.

It is interesting that recent progress towards the identification of telomere end factors has identified a heterogeneous nuclear ribonucleoprotein (hnRNP A1) in mammals (LaBranche et al. 1998) and a repair protein with non-specific DNA binding activity (Yku70-80p) in yeast (Gravel et al. 1998; Nugent et al. 1998; Polotnianka et al. 1998). These findings suggest that the overhang binding activity expected to be required at telomeres in higher eukaryotes may be taken up by factors that are not exclusive for the telomere. In this view, telomeric specificity would be imparted by exclusively-telomeric proteins, such as the TRF proteins and, possibly, ATEF, but the formation of a functional telomeric complex would require the additional presence of non-exclusively-telomeric factors. The situation is reminiscent of the yeast telomeric complex where Sir proteins, although essential for the establishment of telomeric heterochromatin, are also involved in mating-loci silencing, DNA repair, and nucleolar-mediated senescence ((Shore 1995) and (Sinclair et al. 1998), for review).

Our results indicate that, while ATEF could be potentially bound to the 3' end of the overhang, the human Ku70/80 complex, on the other hand, is biochemically capable of recognizing efficiently the overhang-duplex transition. Genetic experiments with Ku-deficient rodent cell lines will be required to address whether Ku is indeed part of the telomeric complex in mammals as it is in yeast. However, even the interpretation of these results will not be necessarily straightforward, as redundancy of function may be a feature of mammalian telomeres. In yeast, for example, Cdc13p and Yku70/80p appear

to have partially redundant protective function (Gravel et al. 1998; Nugent et al. 1998; Polotnianka et al. 1998).

The possibility that Ku is a component of the vertebrate telomeric complex raises a conundrum, resulting from the fact that Ku is normally involved in the processing of double-strand breaks ((Jackson 1996; Lieber et al. 1997) for review) while one of the functions of the telomeric complex is to sequester chromosome ends from this pathway (Sandell and Zakian 1993; van Steensel et al. 1998). Again, telomere–specific activities may then come into play to suppress, at the telomere, the Ku-dependent joining of DNA ends.

Thus, a complex interplay of factors is likely to be necessary to ensure the correct functioning of the telomeric complex in higher eukaryotes. In addition to protecting the end, the terminal telomeric complex in these organisms may have an active role in modulating the action of telomerase at a given telomere. Telomeric DNA in vertebrates has heterogeneous length as a result of shortening and lengthening processes. End factors are likely to play a role in affecting the rate of shortening with their protective function. Such a function may be carried out by hnRNP A1, since cells defective in this protein have shorter telomeres. End factors may conceivably also play a role in determining the rate of lengthening by affecting the accessibility of the DNA terminus to telomerase (Froelich-Ammon et al. 1998).

That protective function and telomere-length regulating activity can be the result of modifications in the activity of individual binding factors is, somewhat paradoxically, illustrated by the function of a duplex binding factor, TRF2. TRF2 is both a negative regulator of telomere length (presumably by acting on telomerase) and a (functionally defined) capping activity required for the maintenance of the G-strand overhang (van

Steensel et al. 1998). How the loss of TRF2 results in erosion of the G-overhang is still a puzzling question. Our efforts to detect a single-stranded DNA binding activity by TRF2 have yielded negative results (A. Bianchi, D. Rhodes and T. de Lange, unpublished), even though the presence of a small basic terminal domain in the protein suggests an attractive similarity to p53 which binds single-stranded DNA with a basic short C-terminal domain. The biochemical characterization of TRF2 binding activity has suffered from a tendency of the factor to aggregate and from the requirement for unusual binding conditions. However, the protein deserves further study, and its unusual and unexpected function in protecting the terminus may require imaginative approaches. In a simpler scenario, it is possible that TRF2 recruits to the overhang an activity capable of binding to the single-stranded DNA. In this case it is conceivable that the single-strand binding factor would display low binding specificity since targeting by TRF2 would ensure highly efficient binding.

These considerations leave open many possibilities for what type of factors may represent the vertebrate capping activity and an open mind should be kept for what biochemical properties are to be expected from it.

Double-stranded DNA binding activities: sensors of a functional telomere?

The characterization of duplex binding activities at the vertebrate telomere has been an area of more rewarding progress, with the identification of TRF1 and TRF2 (Chong et al. 1995; Bilaud et al. 1997; Broccoli et al. 1997b). A search of other TRF homologues in cDNA libraries and EST databases, suggests that no additional components of the TRF family remain to be discovered (S. Oelmann, D. Broccoli and T.

de Lange, unpublished). It is thus possible that we are already in possession of all the elements that are responsible for duplex TTAGGG repeat binding activity in higher eukaryotes. This is an exciting situation, as it is possible to start addressing questions regarding how these protein organize telomeric chromatin in view of their binding characteristics and of their ability to tether to the telomere additional proteins devoid of DNA binding activity (S. Smith, A. Schmidt and T. de Lange; B. Li and T. de Lange; both unpublished).

We have begun to elucidate the organization of the telomeric complex by studying the DNA binding characteristics of TRF1. We have proposed a relatively unusual binding mode for the protein which hypothesizes a high degree of structural flexibility in DNA binding. Although the protein might have the ability to contact two 5'-AGGGTT-3' sites simultaneously at relatively large distance, *in vitro* it appears not to do so, and, given a choice, will bind to two closely arranged half sites. Whether the situation may be different *in vivo* in the possibly densely packed telomeric DNA, wrapped in nucleosomes, remains to be seen. The high concentration of TRF1 sites at the telomere, coupled with the high off-rate of the protein, may efficiently serve as a "sink" for TRF1 at that nuclear location, leaving possibly little TRF1 unbound or at chromatin elsewhere in the nucleus. Consistent with this notion is the fact that no TRF1 can be extracted from the nucleus at low salt (Zhong et al. 1992; Ludérus et al. 1996). At the same time, telomeric DNA may be kept in an open state, with TRF1 coming on and off, that could render it accessible to additional proteins with regulatory function. Perhaps for this reason TRF1 and TRF2 are unable to inhibit telomerase when bound at a site as close as four nucleotides from the DNA terminus.

EM studies of TRF1 complexes with telomeric DNA have revealed that the protein is capable of imparting an ultrastructural organization to telomeric DNA, by promoting telomere-telomere interactions that may be intra- or inter-molecular (Griffith et al. 1998). The implications of this phenomenon are largely unexplored and may vary from an involvement in telomere length regulation (van Steensel and de Lange 1997) to a role in the meiotic synapses of chromosome ends (de Lange 1998a). Clearly more needs to be done in this regard and the next stage will be to determine how TRF1 affects the organization of nucleosomal telomeric DNA.

A striking parallel that seems to run across organisms is the shared role of duplex binding telomeric factors is in their ability to negatively regulate telomere length *in cis* (Cooper et al. 1997; Marcand et al. 1997; van Steensel and de Lange 1997). From a mechanistic point of view this is still a very puzzling phenomenon. It appears that it is the total number of molecules of telomeric protein per telomere that is counted, and not simply the presence of telomeric factors near the chromosome end. Models need to be generated to explain how the amount of protein bound can transmit an "on" or "off" signal to telomerase. Probably, in the future some of the most valuable information on these and analogous issues will be provided by experiments that will follow telomere dynamics at individual telomeres. The ability to target telomeric proteins at these telomeres (Marcand et al. 1997) will be a powerful tool in this regard. If our model for TRF1 binding is correct, it may be possible to generate artificial telomeres in which the spacing between TRF1 half sites is changed, without affecting telomere function. This could constitute a first step in manipulating telomeric DNA in an attempt to monitor telomere dynamics.

Exciting times ahead for the telomeric complex

Considerable excitement in telomere biology is provided by recent studies concerning the effects of altering the expression of telomerase on cellular life-span. The enzyme has potentially profound implications in medical science, specifically cancer research. Clearly, an understanding of its regulation at the level of expression is of primary importance, and so is the characterization of its structure and biochemical activity. However, potentially beneficial applications may come also from a deeper understanding of the role of the telomeric complex and its components in regulating telomere length.

On less applied grounds, a source of great interest in the next few years will be to see how in different organisms different answers have been given to the common problem of assembling a "telomere", in the operationally defined sense of Muller, at the tip of the chromosome. In *Drosophila* perhaps the most unusual path of all has been taken, by making use of transposable elements (Levis et al. 1993). What most other eukaryotes appear to have in common is the presence of short telomeric repeats and their replication by telomerase. From here, though, the assembly of the telomeric complex seems to have taken different paths. The presence of Myb-repeats in telomeric double-stranded factors is emerging as a common theme, being present in the yeast telomeric proteins Rap1p and Taz1p, in addition to the TRFs (Konig and Rhodes 1997). However, the TRF proteins are otherwise unrelated to Rap1p. Thus a modular organization of telomeric proteins may be expected, in which a shared feature may be the ultimate recruitment of similar activities (such as telomerase) to the telomere. In general, the approach of identifying telomeric proteins by homology searches has proven unsuccessful and the rate of divergence for TRF1 and TRF2 is high even between mouse

and human (Broccoli et al. 1997a). This poses additional challenges in the identification of new components but identified parallelisms will be highly informative on function. Seeing how many of the same problems posed by the chromosome tip have gotten either similar or totally different solutions across phyla is likely to be one of the most fascinating aspects of the biology of the telomeric complex in the next several years.

Experimental Procedures

ATEF DNA binding assays

Binding assays were performed as described in (Cardenas et al. 1993). Briefly, for reactions involving partially double-stranded DNAs, 0.05 pmoles of radiolabeled DNA probe, 2 μg of sonicated *E.coli* DNA and the appropriate protein preparation were incubated for 20 min at room temperature in 20 mM Hepes pH 7.75, 100 mM KCl, 5 mM MgCl2, 0.02% NP40, 0.5 mM DTT and 10% glycerol. For DNA binding reactions with single-stranded DNA and RNA probes, 3-5 μg of cytoplasmic RNA from HeLa cells and 20 units RNasin (Promega) were included; $MgCl_2$ was omitted and 1 mM EDTA and I mM EGTA were included. After incubation the mixture was loaded on 6% polyacrilamide gels (29:1) in TBE. Gels were electrophoresed at 150V for 2.5 hours

Preparation of nuclear extracts from chicken erythrocytes

Nuclear extracts were made in batches from 500 ml of chicken blood at a time. The blood was diluted to a final volume of 4 l with buffer A' (50 mM KCl, 50 mM Hepes, 1 mM $MgCl_2$, 0.5 mM DTT, 0.3 mM PMSF – DTT and PMSF were included only after cell lysis) and centrifuged in 1 l bottles at 3500 rpm at 4°C in a Sorvall centrifuge for 30 min. Pellets were resuspended in 2 l of buffer A' and spun again. Pellets were again resuspended, in 800 ml of buffer A' and filtered through cheesecloth to eliminate clots. Cells were lysed by the addition of 40 ml of 10% NP40, to a final concentration of 0.5%. Lysis was immediately checked by phase contrast microscopy and nuclei were quickly diluted to 4 l with buffer A'. Nuclei were then centrifuged at 4000 rpm for 30 min in a Beckman JA10 rotor at 4°C. Pellets were resuspended in 3 l buffer A' and centrifuged again at 3000 rpm for 20 min in a Beckman JA10 rotor at 4°C. Pellets appeared light pink at this stage. Final volume of combined pellets was about 200 ml. One volume of

Dignam buffer (300 mM KCl, 33% glycerol, 10 mM Hepes pH 7.9, 0.2 mM EDTA, 7 mM $MgCl_2$ with protease inhibitors pepstatin, leupeptin, apoproteinA) was added and nuclei were incubated at 4°C for 30 min with stirring. Nuclei were finally spun in rotor JA10 at 4000 rpm for 30 min and and the supernatant was dialyzed for 2-4 hours in 20 mM Hepes pH 7.75, 20% glycerol, 100 mM KCl, 1 mM EDTA, 1 mM DTT).

Partial purification of ATEF

The following materials were used: Heparin-Sepharose (Pharmacia), Phosphocellulose (Whatman), TMAE-fractogel (EM Separations), CM-Sepharose (Pharmacia), Sephacryl-300 (Pharmacia), Single-Stranded DNA (calf thymus)-Cellulose (SIGMA).

In all cases the protein sample was applied to the column in the dialysis buffer (20 mM Hepes pH 7.75, 20% glycerol, 100 mM KCl, 1 mM EDTA, 1 mM DTT), sometimes glycerol was omitted. Salt gradients were from 0.1 to 1 M KCl and were obtained by gradual mixing of equal volumes of the two extreme salt concentrations.

Two types of affinity columns were used, with identical results. Both were based on the binding of biotinylated oligonucleotides coupled to a streptavidin-agarose resin (Pierce). The DNA used in the first column was the annealing product of oligonucleotides 5'AAACGTCTACAGCATCTTAGGGTTAGGG3' and 5'GATGCTGTAGACGTTTX3'. The oligonucleotide used for the second column was 5'XXXAAACGTCTACAGCATCTTAGGGTTAGGG3'. In both cases X represents a biotinylated residue. Protein samples were applied to the affinity column in buffer containing either 0.1 or 1 M KCl. Column was washed with dialysis buffer + 2 M KCl and eluted with 100 mM carbonate buffer pH 10.7, 1 mM DTT, 500 mM KCl.

Purification of human Ku from HeLa nuclear extracts

Ku heterodimer was purified to apparent homogeneity (Fig. 2.1b) from HeLa nuclear extracts. Four chromatographic steps were used for the purification (de Vries et al. 1989): DEAE-fractogel 650 (EM Separations, Ku eluting at 0.2-0.3 M KCl), phosphocellulose (Whatman, Ku eluting at 0.6 M KCl), double-stranded DNA cellulose (SIGMA, Ku eluting at 0.4-0.5 M KCl) and single-stranded DNA cellulose (SIGMA, Ku eluting at 0.5 M KCl). About 10 μg of purified Ku were obtained from about 4 g of total nuclear protein.

Preparation of oligonucleotides and probe labelling for Ku binding assays and competitions

Oligonucleotides were synthesized on an Applied Biosystem DNA Synthesizer and gel purified before use. Labelling of 5' ends was carried out by T4 polynucleotide kinase and γ^{32}P-ATP (3000 Ci/mmol). Annealing of double-stranded substrates was achieved by incubation for 2 hours at room temperature in TE with or without added salts at a DNA concentration of 4 pm/μl and was carried out after labelling and folding of overhangs when appropriate. Folding of overhangs was induced by incubating at 90°C for 3 min and at room temperature for 20 min in 10 mM Tris pH 8.0, 0.1 mM EDTA with or without 50 mM KCl or LiCl, at DNA concentrations ranging from 0.2 to 40 pm/μl (Williamson et al. 1989). DMS protection assays were as described in (Williamson et al. 1989).

Ku DNA binding assays and gel electrophoresis

Binding reactions were carried out at room temperature for 15 min. Reactions described in Fig. 2.2 were carried out in 20 mM Hepes pH 7.75, 2 mM $MgCl_2$, 0.1 mM EDTA, 0.25 mM DTT, 200 mM KCl, and 7% glycerol. All other reactions were performed in 10 mM Tris pH 8.0, 0.1 mM EDTA, with or without added salts, as indicated in the individual experiments. Electrophoresis was performed in 5% acrylamide (29:1), 20 cm long gels in 1xTBE, at 130V, for 2 hours at room temperature. Experiments were quantitated on a PhosphorImager using ImageQuant software.

Partial purification of human telomerase (performed by Dr. Gisela Schnapp, Boehringer Ingelheim, Biberach, Germany)

Nuclear extracts from HeLa cells were prepared according to (Dignam et al. 1983). The extracts were dialyzed against buffer A-100 (20 mM Hepes, pH7.9, 1 mM EDTA, 1mM EGTA, 10 % glycerol, 100 mM KCl, 0.5 mM PMSF, 0.5 mM DTT), and applied to a Spermine-Agarose column (Sigma). The column was washed with 250 mM KCl, and 500 mM KCl in buffer A. Telomerase activity was eluted with buffer A containing 1 M KCl. The active fractions were pooled, dialyzed against buffer A-100, and loaded onto a Heparin Cl-6B column (Pharmacia) equilibrated in buffer A-100. The column was washed with buffer A containing 220 mM KCl, and active telomerase was step eluted with 500 mM KCl. Active fractions were pooled, dialyzed as described above, and applied to a MonoQ-FPLC column (Pharmacia), equilibrated with buffer A-100. The column was washed with buffer A containing 300 mM KCl, and telomerase was eluted with a linear gradient from 0.3 to 1 M salt in buffer A. The peak of active telomerase eluted at 500 mM KCl. Active fractions were pooled (1 mg/ml total protein) and used for direct

telomerase assays. These fractions were up to 100-fold purified with respect to specific telomerase activity determined in the conventional telomerase assay (Morin 1989).

Telomerase reactions

Telomerase reactions were performed in a volume of 20 µl using typically 1-2 µl of telomerase fractions and 25 mM Tris Acetate pH 8.2, 50 mM Na Acetate pH 8.3, 1 mM $MgCl_2$, 1 mM EGTA, 5 mM dATP, 5 mM dTTP, 4.5 µM dGTP, 0.1 µl RNAsin and 1 mM spermidine. Typically 1-5 µl of labelled dGTP (NEN, 3000 Ci/mmol, 10 mCi/ml – final concentration 0.165-0.825 µM) were used per reaction. Concentration of primer was 10 nM. Reactions performed in the presence of TRF1 or TRF2 were done both in the presence and absence of spermidine. Baculovirus-expressed TRF1 or TRF2 protein (see below), or BSA as a control were incubated with the DNA in the telomerase reaction buffer for 30 minutes at room temperature. Subsequently, telomerase fractions were added and the reactions were allowed to proceed for 1 hr at 30°C, and stopped by adding 20 µl H_20 and 50 µl of 10 mM Tris pH 8.0, 5 mM EDTA, 0.1 mg RNAse/ml, and incubating at 45°C for 10 min. Next, 50 µl 10 mM Tris pH 8.0, 1% SDS, 0.3 mg protK/ml and the samlpes were incubated at 45°C for 10 min. Samples were extracted with 200 µl of phenol/chloroform/isoamyl alcohol and precipitated by adding 50 µl 3 M Na Acetate pH 5.2, 2.5 µg tRNA and 500 µl ethanol and incubation on ice for 15 min. Samples were centrifuged at room-temperature for 10 min and the ethanol was removed. Pellets were dissolved in 4 µl of formamide loading dye and all or half the sample was loaded on an gel (see below).

Gel electrophoresis of telomerase products

Electorphoresis was carried out in 6% polyacrylamide (19:1) gels containing 7 M urea and 20 % (v/v) formamide in 1xTBE. Gels were run at 40-55 W, so that the external temperature of the gel was about 50-60°C.

Oligonucleotides for telomerase assays

Oligonucleotides were gel purified on PAA-7 M urea gels and annealed in 10 mM Tris pH7.5, 1 mM EDTA, 100 mM NaCl by boiling for 5 min and slow cooling (over a few hours) in a water-bath from 80°C to room temperature.

Band-shift analysis for telomerase assays

Gels shift analysis to monitor TRF1 and TRF2 binding was performed in 0.6% agarose, 0.1xTBE at 130V for about 45 min. Binding conditions were exactly as for telomerase assays except that dNTPs were omitted.

***In vitro* translation reactions**

TRF1 deletion mutants used for the *in vitro* coupled transcription-translation experiments were cloned in the vector pET28(a) (PROMEGA) in the NcoI and EcoRI sites using PCR-generated fragments. The GFP-TRF1 fusion product was cloned in pBluescriptKS+. PCR-directed mutagenesis was used to eliminate from this construct the start codon of the TRF1 gene by mutating it from ATG to ATT in order to suppress the occurrence of internal translation at this position. The GFP sequence was obtained from pS65T-C1 (Clontech). Expression of TRF1 derivatives was achieved using a rabbit reticulocyte lysate system (Promega) using reaction conditions essentially as described

by the supplier. Briefly, between 0.2 and 1 μg of plasmid DNA (prepared by the alkalyne lysis procedure and RNase and protK treated), was used per 20 μl reaction containing T7 RNA polymerase in the presence of [^{35}S]methionine (to visualize products on SDS-page) or without labeled amino acids (for gel-shift assays). After the transcription-translation reaction, samples were diluted 1:5 with the addition of 80 μl of buffer D (Chong et al. 1995) and stored at –70°C. Of this mixture, 0.5-5 μl was used in gel-shift reactions. For SDS-gels, typically 2-5 μl of samples were loaded. Gels were fixed and enhanced (Amersham). Exposure was overnight.

Band-shift assays for TRF1

Gel shift assays were performed as described previously (Zhong et al. 1992) using either labeled restriction fragments or oligonucleotides as probes. Binding incubations were in 5% glycerol, 4% Ficoll, 20 mM Hepes pH 7.9, 150 mM KCl, 1 mM $MgCl_2$, 0.1 mM EDTA, 5 ng/μl β-casein, and 0.5 mM DTT. In some experiments 0.1% NP40 was added, as it was found to improve binding. Samples were incubated at room temperature for 30 min and then loaded on a 5% polyacrylamide gel (29:1) in 1xTBE run at room temperature, or 0.5xTBE runa t 4°C. Gels were run at 130-140 V (about 60 mA) for about 2 hours and then dried and either exposed to X-ray films or quantitated by PhosphorImager.

Band-shift assays for the TRF1Myb domain

E.coli-expressed, purified TRF1 Myb domain was provided by P. Konig and D. Rhodes (Konig et al. 1998). Binding reactions were as for TRF1 (0.1% NP40 was present). Gels were as for TRF1 assays, in 0.5xTBE, run at 4°C.

Preparation of DNA fragments for binding assays

In order to ensure correct annealing, probes containing telomeric repeats were preferably prepared from restriction fragments. Occasionally, when large amounts of DNA were needed (as in the case of substrate KdGC), oligonucleotides were employed. Cloning was achieved by standard procedures. DNA fragments were isolated from low-melt point agarose gel (SeaPlaque, FMC, typically 0.8% in 0.5xTBE). Briefly, gel slices containing the band of interest were cut out and one volume of TE buffer was added in an Eppendorf tube. The slice was then incubated at 65°C for 15 min and thoroughly melted. The agarose was extracted (twice) by adding one volume of Tris-buffered phenol. The aqueous phase from the second extraction was transferred to a new tube and 0.2 volumes of 2 M LiCl were added, mixing immediately and incubating on ice for 2 min. The sample was then centrifuged in a microcentrifuge at top speed for 3 min and the supernatant was ethanol-precipitated. The concentration of the purified DNA fragments was determined by measuring absorbance at 260 nm and by agarose electrophoresis against markers of known concentration.

Labelling of probes for binding assays

Probes were typically labeled using the Klenow fragment of DNA polymerase I from *E. coli*. Typically about 0.5-1 pM of DNA were labeled in a 20 µl reaction. Reactions were allowed to proceed for 1 hour at room temperature in the presence of either labeled dCTP or dGTP (NEN, 3000 Ci/mmol, 10 mCi/ml – final concentration 0.330-0.825 µM) and the corresponding remaining three cold nucleotides (0.5 mM). Reactions were stopped by adding 80 µl of 10 mM Tris pH 7.5, 1 mM EDTA, and 0.1%

SDS(TES) and applied on a Sephadex G50 column (Pharmacia) packed in a 3 ml syringe. The labeled fragment was eluted in a volume of 800 µl (after 1 ml of void volume) of TES, extracted with phenol/cloroform/isoamyl alcohol and ethanol-precipitated. Probes were resuspended in 10 mM Tris pH7.5, 1 mM EDTA, 100 mM NaCl.

Expression of TRF1 and TRF2 proteins in insect cells using baculovirus expression vectors

N-teminally $[His]_6$-tagged versions of TRF1 and TRF2 were cloned in the baculovirus expression vectors pBacPak8 and pBacPak9 (Clontech), respectively. Either of these plsmids was used to co-transfect insect Sf21 cells together with linearized baculovirus BacPak6 (Clontech). Recombinant viruses were plaque-purified, screened for TRF1 and TRF2 expression and amplified. For protein production, 500 ml of SF-21 cells in a 1 l spinner flask were grown to a density of approximately $5x10^5$ cells/ml in a 27°C incubator. Cells were infected with baculovirus expressing TRF1 or TRF2 at an m.o.i. of 4-5 and harvested after approximately 40 hours. Cells were centrifuged in a swing-out rotor at 230 g (JS 4.3 rotor in a Beckman JS 6B at 1K) at 4° C for 15 min. The supernatant was removed and the pellets washed with 40 ml cold PBS. The cells were centrifuged at 230 g, 4°C for 10 min (Sorvall RT6000B at 1 K). The pellet was then resuspended in 12 ml of 5 mM imidazole, 500 mM NaCl, and 20 mM Tris-HCl pH 7.9 at 4° C. Cell lysis was achieved by sonicating on ice for 4 min, pulsing with a 0.5 seconds on/off rate. Lysate was centrifuged in an ultra-centrifuge at 45,000 g (SW41 rotor at 19.1 K) at 4° C for 30 min. Supernatant was added to 0.3 ml of Ni^+-activated Sepharose resin (Chelating Sepharose Fast Flow, Pharmacia) and incubated, rotating, at 4°C for 1 hour. Resin was centrifuged in a swing-out rotor at 1000g (Sorvall RT6000B,

1K) for 10 min, at 4° C, and washed batch wise 2 times with 5 volumes of the same buffer, and then 2 times with 3 volumes of 60 mM imidazole. Proteins were eluted by washing the resin 2 times with 2 volumes of 1 M imidazole, 0.5 M NaCl, and 20 mM Tris-HCl pH 7, and finally dialyzed in 20 mM Hepes pH 7.9, 20% glycerol, 500 mM KCl.

Resin was activated by washing it 3 times with 1 volume of ddH_2O and 5 times with 1 volume of 50 mM $NiSO_4$. Finally the resin was equilibrated by washing it 3 times with one volume of 5 mM imidazole, 500 mM NaCl, and 20 mM Tris-HCl pH 7.9.

Sf21 cells were grown in Grace's Insect Cell media (Gibco) supplemented with penicillin and streptomycin, and 10% Fetal Bovine Serum (SIGMA).

Protein yield was typically 2-5 mg/500 ml of cells. Amount of active protein was determined by band-shift assays at high DNA concentration or by comparison to preparations which specific activity had been determined.

SELEX of TRF1 binding sites

Selex (Systematic Evolution of Ligands by Exponential enrichment) was performed on an oligonucleotide library containing a central degenerate sequence of 38 nucleotides. To this purpose the 84 nucleotide long single-stranded oligonucleotide 5'ATCGGATCCTTGATCAAGCTGCA$(N)_{38}$CGACATGTATCGATGAATTCGAC3' 1Sel38 was chemically synthesized (Genelink). Double stranded molecules were obtained by annealing the randomized library with the 20 mer 5'GTCGAATTCATCGATACATG3' (1Eco) and by extending the annealed primer with Klenow enzyme (New England Biolabs). Annealing was obtained by mixing approximately 380 pmoles of 1Sel38 and 650 pmoles of 1Eco, in TE plus 100 mM NaCl, in a volume of 20 µl. Incubation was at 94°C for 5 min followed by incubation at 44°C for 30 min. Primer extension was

performed by increasing the volume to 80 µl to a final concentration of 50 mM Tris pH 7.5, 10 mM $MgCl_2$, 1 mM DTT, 50 µg/ml BSA, 2.5 mM each dNTP, and 0.25 u/µl Klenow. The primer extension reaction was allowed to proceed at 30°C for 20 hours, and samples were phenol-extracted, ethanol-precipitated , and resuspended in TE with 100 mM NaCl. The products of the polymerization reaction were analyzed on acrylamide gel. The concentration of the double-stranded 84mer oligonucleotide was estimated to be about 400 ng/µl. About 1.2 µg of double-stranded oligonucleotide were used in the first round of TRF1 binding reactions (in two separate experiments). Each round of binding reactions was performed at three different TRF1 concentrations (approximately 1 µM, 200 nM and 40 nM). Binding reactions were done as described above except that sonicated *E.coli* competitor DNA was substituted by ϕX174 HaeIII-digested DNA (GIBCO). This DNA was chosen because it lacks BamHI and EcoRI sites. Reaction mixtures were loaded on acrylamide gels as described above for TRF1 assays, and TRF1-bound DNA was isolated by excision of the gel slice containing the DNA-TRF1 complex. Since the DNA was unlabelled, the position of the complex in the gel was determined by loading, in an adjacent lane, a binding reaction in which a labelled DNA was used which had unrelated non-telomeric sequence compared to the librbary oligonucleotide and which gave a complex of identical mobility. The gel slice was crushed in 400 µl of 500 mM NH_4Cl and 0.1% SDS, briefly heated to inactivate the protein and incubated, rotating, at room temperature overnight. The DNA was then phenol–extrated, precipitated with 4 volumes of ethanol, and resuspended in 30 µl ddH2O. PCR reactions using primers 1Eco and 1Bam (5'ATCGGATCCTTGATCAAGCT3') were performed using 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 pmoles/µl of each primer, and Taq

polymerase (concentration approximately 0.1 units/μl, provided by Dr. Jan Karlseder). Reactions were at 94°C for 30 sec, 44°C for 30 sec and 72°C for 30 sec. A total of 20 cycles were performed and then one volume of fresh reaction mix (excluding DNA) was added and one more cycle performed at 94°C for 2 min, 44°C 1 min and 72°C 10 min. The last cycle was added to maximize the percent of DNA molecules in a double-stranded form. Sample were then phenol-extracted, ethanol-precipitated and resuspended in 5 μl TE plus 100 mM NaCl. The amount of PCR products was verified at each round of selection by loading 0.4 μl on a 2% agarose gel. 2 μl of PCR products were used for the binding reaction in the subsequent round of selection. A total of 7 rounds of selection were performed, at the end of which PCR products were cut with EcoRI and BamHI, and cloned in pBluescriptKS (Stratagene). DNA from individual clones was prepared and analyzed by sequencing.

To determine the base composition in the unselected library, we performed four PCR cycles as above, and sequenced 5 resulting clones, for a total of 190 nucleotides.

Estimates of dissociation constants

The dissociation constant (Kd) for a protein (P) binding to a DNA molecule (D) is given by the equation $Kd = (P_f \times D_f)/PD_c$, where P_f is the concentration of free protein, D_f is the concentration of free DNA, and PD_c is the concentration of the protein-DNA complex. If binding reactions are performed at low DNA concentration, the concentration of total protein (P_t) approximates the concentration of free protein and therefore $Kd = (P_t)(D_f)/PD_c$. At 50% DNA binding $D_f = PD_c$, and therefore $Kd = P_t$ (Fairall et al. 1992). The concentration of active protein in the preparations was determined by performing binding assays with a DNA concentration in the μmolar range. All binding

experiments aimed at measuring dissociation constants were performed at DNA concentrations of 50 pM or lower.

Relationship between microscopic and macroscopic dissociation constants

The measured (macroscopic) dissociation constant (K) for binding of a protein site to a DNA site differs from the microscopic dissociation constant (k) if multiple copies of the binding sites are present linked in either the protein or the DNA. For a protein with one binding site (as is the case for the Myb domain of TRF1 - abbreviated as Myb) binding to a DNA containing one copy of the binding site (5'-YTAGGGTTR-3'), the macroscopic and microscopic dissociation constants are equal. Instead, if binding events are independent of one another, the K for binding of Myb to a DNA containing two copies of the binding site, is $K = k/2$ as can be seen by the following equations, where PD is the protein-DNA complex at the macroscopic level and PD_1 and PD_2 indicate the microscopic complexes in which the first or the second site in the DNA are occupied, respectively.

The macroscopic equilibrium is: $P + D \leftrightarrow PD$

whereas microscopically: $P + D \leftrightarrow PD_1$

and $P + D \leftrightarrow PD_2$

Therefore: $K = (P)(D)/PD$

and: $k = (P)(D)/PD_1 = (P)(D)/PD_2$

Since: $PD = PD_1 + PD_2$

it follows that: $K = k/2$

Similarly, the binding of a protein with two independent binding sites (as we propose is TRF1) to a DNA containing one copy of a binding site, also gives a macroscopic dissociation constant that is half the microscopic dissociation constant. The binding of such a protein to a DNA containing two binding sites instead gives (for the first binding event) a dissociation constant that is 1/4 the microscopic constant (K = k/4), 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.

Analysis of TRF1 kinetics of dissociation from DNA using Plasmon Surface Resonance

The DNAs used in the Biacore experiments were EcoRV-XhoI fragments (for the DNA containing 2 TTAGGG repeats) or EcoRV-HindIII (for the DNAs containing 3, 4, and 5 TTAGGG repeats). Sizes of fragments were 93, 87, 93, and 99 bp, respectively. Fragments were biotinylated by fill-in reactions using Klenow polymerase. The DNAs were immobilized on streptavidine-coated chips. About 49, 50, 73 and 54 RU units were immobilized on the chips for the 4 DNAs, respectively. Experiments on the Biacore were performed both in the Hepes binding buffer (see above) and in the glycine binding buffer (with 50 mM KCl; see below). Protein concentrations were as indicated. Injection time was 120 sec at a flow rate of 20 µl/min at 10°C. The chips were regenerated after binding by washing with 0.05% SDS. Data were analyzed using BiaEvaluation 3.0.

DNA ligation assays

An Asp718-cut kinase end-labeled 217 bp DNA fragment containing an array of 27 TTAGGG repeats was used. The DNA was incubated for 20 min at room temperature

either with active or heat-inactivated (55°C for 30 min) baculovirus-expressed TRF1 in 20 mM Hepes pH 7.9, 200 mM KCl, 10 mM $MgCl_2$, 2 mM DTT. For the TRF1 titration experiments, reactions were carried out with 35 ng/ml of DNA and a TRF1 concentration that varied from 15 to 2000 ng/ml. For rate measurements, the DNA concentration was 40 ng/ml and TRF1 was added to 500 ng/ml. ATP was added to 1 mM and ligase to 10 U/ml (protein titration) or 1000 U/ml (rate measurements). Ligation reactions were performed at 23°C and allowed to proceed for 30 min (protein titration) or from 0 to 128 min (rate measurements). Reactions were stopped by the addition of 1/2 volume of stop buffer (75 mM EDTA, 3 mg of proteinase K/ml, 15% glycerol) and incubation at 55°C for 15 min. Exonuclease-treated samples were phenol-extracted, ethanol-precipitated, resuspended in 20 µl of 40 mM Tris pH 7.5, 20 mM $MgCl_2$, 50 mM NaCl and 50 U of T7gene6 exonuclease (USB) and incubated for 2 hours at 37°C. Digestions were terminated by the addition of 1/2 volume of stop buffer and incubated at 55°C for 15 min. Samples were run on 6% polyacrylamide gels in TBE. Products were quantitated using a PhosphorImager (Molecular Dynamics).

DNAseI footprinting of TRF1 on hTER and mTER

DNA substrates for footprints was obtained by PCR on plasmids containing the hTER and mTER genes. For footprints on hTER, primers hTRfoot (5'GCGGCAGGCCGAGGCTT3') and SP6 were used; for mTER, primers mTRfoot (5'CGCTGCAGGTCTGGACTT3') and T7 were used. The template sequence in hTER and mTER in both PCR products was centered at about 100 bp from each end. In each PCR reaction 10 pmoles of each primer were used, one of which was labeled by T4 polynucleotide kinase. PCR reactions were performed in the presence of 1 ng of

template plasmid, in a volume of 100 µl, for 30 cycles at an annealing temperature of 55°C. TRF1 binding was in standard conditions except that 1 mM $MgCl_2$ was added and the reaction volume was 40 µl. TRF1 concentration was approximately 500 nM. Binding was allowed to proceed for 20 min at room temperature, 2 units DNase I were added, and the reaction was stopped after 10 sec by addition of one volume of 20 mM EDTA. Samples were phenol-extracted, ethanol-precipitated and analyzed on 6% PAA (19:1)-7 M urea gels. Maxam and Gilbert G and A+G marker lanes were produced as described(Ausubel et al. 1994).

Band-shift analysis of TRF2

Binding conditions for TRF2 (and occasionally for TRF1) were 20 mM glycine-KOH pH 9.0, 4% glycerol, 0.5% DTT and variable concentration of KCl (0-50 mM). Affinity was similar under the different salt conditions, but lower salt gave sharper looking bands. Complexes were analyzed in 0.6% agarose gels in 0.1x TBE, which were run at 130 V for about 40 min. Gels were dried on DE81 paper (Whatman) at 80°C and either exposed on X-ray films or analyzed using a PhosphorImager.

References

Allshire, R.C., M. Dempster, and N.D. Hastie. 1989. Human telomeres contain at least three types of G-rich repeats distributed non-randomly. *Nucl. Acids Res.* 17: 4611-4627.

Allshire, R.C., J.R. Gosden, S.H. Cross, G. Cranston, D. Rout, N. Sugawara, J.W. Szostak, P.A. Fantes, and N.D. Hastie. 1988. Telomeric repeat from T. thermophila cross-hybridizes with human telomeres. *Nature* 332: 656-659.

Allsopp, R., H. Vaziri, C. Patterson, S. Goldstein, E. Younglai, B. Futcher, C.W. Greider, and C.B. Harley. 1992. Telomere length predicts replicative capacity of human fibroblasts. *Proc. Natl. Acad. Sci. USA* 89: 10114-10118.

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

Ausubel, F.M., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl. 1994. Current protocols in molecular biology. Wiley Interscience Press.

Autexier, C. and C.W. Greider. 1994. Functional reconstitution of wild-type and mutant Tetrahymena telomerase. *Genes Dev.* 8: 563-575.

Balagurumoorthy, P. and S.K. Brahmachari. 1994. Structure and stability of human telomeric sequence. *J. Biol. Chem.* 269: 21858-21869.

Baranowskij, N., C. Frohberg, S. Prat, and L. Willmitzer. 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.

Barnett, M., V. Buckle, E. Evans, A. Porter, D. Rout, A. Smith, and W.R.A. Brown. 1993. Telomere directed fragmentation of mammalian chromosomes. *Nucl. Acids Res.* 21: 27-36.

Beattie, T.L., W. Zhou, M.O. Robinson, and L. Harrington. 1998. Reconstitution of human telomerase activity in vitro. *Curr Biol* 8: 177-180.

Becker, H.A. and R. Kunze. 1997. Maize Activator transposase has a bipartite DNA binding domain that recognizes subterminal sequences and the terminal inverted repeats. *Mol Gen Genet* 254: 219-230.

Bednenko, J., M. Melek, E.C. Greene, and D.E. Shippen. 1997. Developmentally regulated initiation of DNA synthesis by telomerase: evidence for factor-assisted de novo telomere foundation. *EMBO J.* 16: 2507-2518.

Berman, J., C.Y. Tachibana, and B.-K. Tye. 1986. Identification of a telomere-binding activity from yeast. *Proc. Natl. Acad. Sci. USA* 83: 3713-3717.

Bernstein, H.S. and S.R. Coughlin. 1997. Pombe Cdc5-related protein. A putative human transcription factor implicated in mitogen-activated signaling. *J Biol Chem* 272: 5833-5837.

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

Biedenkapp, H., U. Borgmeyer, A.E. Sippel, and K.H. Klempnauer. 1988. Viral myb oncogene encodes a sequence-specific DNA-binding activity. *Nature* 335: 835-837.

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

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

Blackburn, E. and J. Gall. 1978. A tandemly repeated sequence at the termini of extachromosomal DNA coding for rRNA in Tetrahymena. *J. Mol. Biology* 120: 33-53.

Blackburn, E.H. and S.-S. Chiou. 1981. Non-nucleosomal packaging of a tandemly repeated DNA sequence at termini of extrachromosomal DNA coding for rRNA in Tetrahymena. *Proc. Natl. Acad. Sci. USA* 78: 2263-2267.

Blackburn, E.H. and J.W. Szostak. 1984. The molecular structure of centromeres and telomeres. *Ann. Rev. Biochem.* 53: 163-194.

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

Blasco, M.A., H.-W. Lee, M.P. Hande, E. Samper, P.M. Lansdorp, R.A. DePinho, and C.W. Greider. 1997. Telomere shortening and tumor formation by mouse cells lacking telomerase RNA. *Cell* 91: 25-34.

Blasco, M.A., M. Rizen, C.W. Greider, and D. Hanahan. 1996. Differential regulation of telomerase activity and telomerase RNA during multi-stage tumorigenesis. *Nature Genetics* 12: 200-204.

Blier, P.R., A.J. Griffith, J. Craft, and J.A. Hardin. 1993. Binding of Ku protein to DNA. Measurement of affinity for ends and demonstration of binding to nicks. *J. Biol. Chem.* 268: 7594-7601.

Bodnar, A.G., M. Ouellette, M. Frolkis, S.E. Holt, C.P. Chiu, G.B. Morin, C.B. Harley, J.W. Shay, S. Lichtsteiner, and W.E. Wright. 1998. Extension of life-span by introduction of telomerase into normal human cells. *Science* 279: 349-352.

Boulton, S.J. and S.P. Jackson. 1996a. Identification of a Saccharomyces cerevisiae Ku80 homologue: roles in DNA double strand break rejoining and in telomeric maintenance. *Nucleic Acids Res* 24: 4639-4648.

Boulton, S.J. and S.P. Jackson. 1996b. Saccharomyces cerevisiae Ku70 potentiates illegitimate DNA double-strand break repair and serves as a barrier to error-prone DNA repair pathways. *EMBO J.* 15: 5093-5103.

Boulton, S.J. and S.P. Jackson. 1998. Components of the Ku-dependent non-homologous end-joining pathway are involved in telomeric length maintenance and telomeric silencing. *EMBO J.* 17: 1819-1828.

Brigati, C., S. Kurtz, D. Balderes, G. Vidali, and D. Shore. 1993. An essential yeast gene encoding a TTAGGG repeat-binding protein. *Mol. Cell. Biol.* 13: 1306-1314.

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

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

Broccoli, D., J.W. Young, and T. de Lange. 1995. Telomerase activity in normal and malignant hematopoietic cells. *Proc. Natl. Acad. Sci. USA* 92: 9082-9086.

Brown, L. and R. Baer. 1994. HEN1 encodes a 20-kilodalton phosphoprotein that binds an extended E- box motif as a homodimer. *Mol Cell Biol* 14: 1245-1255.

Brown, W.R.A. 1989. Molecular cloning of human telomeres in yeast. *Nature* 338: 774-776.

Bryan, T.M., A. Englezou, J. Gupta, S. Bacchetti, and R.R. Reddel. 1995. Telomere elongation in immortal human cells without detectable telomerase activity. *EMBO J.* 14: 4240-4248.

Buchman, A.R., W.J. Kimmerly, J. Rine, and R.D. Kornberg. 1988a. Two DNA-binding factors recognize specific sequences at silencers, upstream activating sequences, autonomously replicating sequences, and telomeres in Saccharomyces cerevisiae. *Mol. Cell. Biol.* 8: 210-225.

Buchman, A.R., N.F. Lue, and R.D. Kornberg. 1988b. Connections between transcriptional activators, silencers, and telomeres as revealed by functional analysis of a yeast DNA-binding protein. *Mol. Cell. Biol.* 8: 5086-5099.

Buck, S.W. and D. Shore. 1995. Action of a RAP1 carboxy-terminal silencing domain reveals an underlying competition between HMR and telomeres in yeast. *Genes Dev.* 9: 370-384.

Cantor, C.C. and P.R. Schimmel. 1980. Biophysical Chemistry, Part III: The behaviour of biological macromolecules. Freeman, New York.

Cardenas, M.E., A. Bianchi, and T. de Lange. 1993. A Xenopus egg factor with DNA-binding properties characteristic of terminus-specific telomeric proteins. *Genes Dev.* 7: 883-894.

Carlson, D.L., R. Skopp, and C.M. Price. 1997. DNA-Binding properties of the replication telomere protein. *Biochem.* 36: 15900-15908.

Chan, C.S.M. and B.-K. Tye. 1983. Organization of DNA sequences and replication origins at yeast telomeres. *Cell* 33: 563-573.

Chikashige, Y., D.-Q. Ding, H. Funabiki, T. Haraguchi, S. Mashiko, M. Yanagida, and Y. Hiraoka. 1994. Telomere-led premeiotic chromosome movement in fission yeast. *Science* 264: 270-273.

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

Collins, K. and C.W. Greider. 1993. Tetrahymena telomerase catalyzes nucleolytic cleavage and nonprocessive elongation. *Genes Dev.* 7: 1364-1376.

Collins, K., R. Kobayashi, and C.W. Greider. 1995. Purification of Tetrahymena telomerase and cloning of genes encoding the two protein components of the enzyme. *Cell* 81: 677-686.

Conrad, M.N., J.H. Wright, A.J. Wolf, and V.A. Zakian. 1990. RAP1 protein interacts with yeast telomeres in vivo: Overproduction alters telomere structure and decreases chromosome stability. *Cell* 63: 739-750.

Cooke, H.J. and B.A. Smith. 1986. Variability at the telomeres of the human X/Y pseudoautosomal region. *Cold Spring Harbor Symp. Quant. Biol.* LI: 213-219.

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

Cooper, J.P., Y. Watanabe, and P. Nurse. 1998. Fission yeast Taz1 protein is required for meiotic telomere clustering and recombination. *Nature* 392: 828-831.

Counter, C.M., A.A. Avilion, C.E. LeFeuvre, N.G. Stewart, C.W. Greider, C.B. Harley, and S. Bacchetti. 1992. Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. *EMBO J.* 11: 1921-1929.

Counter, C.M., H.W. Hirte, S. Bacchetti, and C. Harley. 1994. Telomerase activity in human ovarian carcinoma. *Proc. Natl. Acad. Sci. USA* 91: 2900-2904.

Counter, C.M., M. Meyerson, E.N. Eaton, L.W. Ellisen, S.D. Caddle, D.A. Haber, and R.A. Weinberg. 1998. Telomerase activity is restored in human cells by ectopic expression of hTERT (hEST2), the catalytic subunit of telomerase. *Oncogene* 16: 1217-1222.

Counter, C.M., M. Meyerson, E.N. Eaton, and R.A. Weinberg. 1997. The catalytic subunit of yeast telomerase. *Proc. Natl. Acad. Sci. USA* 94: 9202-9207.

Cross, S.H., R.C. Allshire, S.J. McKay, N.I. McGill, and H.J. Cooke. 1989. Cloning of human telomeres by complementation in yeast. *Nature* 338: 771-774.

da Costa e Silva, O., L. Klein, E. Schmelzer, G.F. Trezzini, and K. Hahlbrock. 1993. BPF-1, a pathogen-induced DNA-binding protein involved in the plant defense response. *Plant J.* 4: 125-135.

de Lange, T. 1998a. Ending up with the right partner. *Nature* 392: 753-754.

de Lange, T. 1998b. Telomeres and senescence: ending the debate. *Science* 279: 334-335.

de Lange, T., L. Shiue, R.M. Myers, D.R. Cox, S.L. Naylor, A.M. Killery, and H.E. Varmus. 1990. Structure and variability of human chromosome ends. *Mol. Cell. Biol* 10: 518-527.

de Vries, E., W. van Driel, W.G. Bergsma, A.C. Arnberg, and P.C. van der Vliet. 1989. HeLa nuclear protein recognizing DNA termini and translocating on DNA forming a regular DNA-multimeric protein complex. *J. Mol. Biol.* 208: 65-78.

Dernburg, A.F., J.W. Sedat, W.Z. Cande, and H.W. Bass. 1995. Cytology of telomeres. In *Telomeres* (ed. E.H. Blackburn and C.W. Greider), pp. 295-338. Cold Spring Harbor Press, Cold Spring Harbor.

Dignam, J.P., R.M. Lebovitz, and R.G. Roeder. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucl. Acids Res.* 11: 1475-1489.

Dionne, I. and R.J. Wellinger. 1996. Cell cycle-regulated generation of single-stranded G-rich DNA in the absence of telomerase. *Proc. Natl. Acad. Sci. USA* 93: 13902-13907.

England, B.P., A. Admon, and R. Tjian. 1991. Cloning of Drosophila transcription factor Adf-1 reveals homology to Myb oncoproteins. *Proc. Natl. Acad. Sci. USA* 89: 683-687.

Fairall, L., S.D. Harrison, A.A. Travers, and D. Rhodes. 1992. Sequence-specific DNA binding by a two zinc-finger peptide from the Drosophila melanogaster Tramtrack protein. *J Mol Biol* 226: 349-366.

Falzon, M., J.W. Fewell, and E.L. Kuff. 1993. EBP-80, a transcription factor closely resembling the human autoantigen Ku, recognizes single- to double-strand transitions in DNA. *J. Biol. Chem.* 268: 10546-10552.

Fan, X. and C.M. Price. 1997. Coordinate regulation of G- and C strand length during new telomere synthesis. *Mol. Biol. Cell.* 8: 2145-2155.

Fang, G. and T.R. Cech. 1991. Molecular cloning of telomere-binding protein genes from Stylonychia mytilis. *Nucl. Acids Res.* 19: 5515-5518.

Fang, G. and T.R. Cech. 1993a. The b subunit of Oxytricha telomere-binding protein promotes G-quartet formation by telomeric DNA. *Cell* 74: 875-885.

Fang, G. and T.R. Cech. 1993b. Oxytricha telomere-binding protein: DNA-dependent dimerization of the alpha and β subunits. *Proc. Natl. Acad. Sci. USA* 90: 6056-6060.

Fang, G., J.T. Gray, and T.R. Cech. 1993. Oxytricha telomere-binding protein: separable DNA-binding and dimerization domains of the alpha-subunit. *Genes Dev.* 7: 870-882.

Farr, C., J. Fantes, P. Goodfellow, and H. Cooke. 1991. Functional reintroduction of human telomeres into mammalian cells. *Proc. Natl. Acad. Sci. USA* 88: 7006-7010.

Feldbrugge, M., M. Sprenger, K. Hahlbrock, and B. Weisshaar. 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.

Feldmann, H., L. Driller, B. Meier, G. Mages, J. Kellermann, and E.L. Winnacker. 1996. HDF2, the second subunit of the Ku homologue from Saccharomyces cerevisiae. *J Biol Chem* 271: 27765-27769.

Feldmann, H. and E.L. Winnacker. 1993. A putative homologue of the human autoantigen Ku from Saccharomyces cerevisiae. *J. Biol. Chem.* 268: 12895-12900.

Feng, J., W.D. Funk, S.-S. Wang, S.L. Weinrich, A.A. Avilion, C.-P. Chiu, R.R. Adams, E. Chang, R.C. Allsopp, J. Yu, S. Le, M.D. West, C.B. Harley, W.H. Andrews, C.W. Greider, and B. Villeponteau. 1995. The RNA component of human telomerase. *Science* 269: 1236-1241.

Friedberg, E.C., G.C. Walker, and W. Siede. 1995. DNA Repair and Mutagenesis. ASM Press, Washington, D.C.

Froelich-Ammon, S.J., B.A. Dickinson, J.M. Bevilacqua, S.C. Schultz, and T.R. Cech. 1998. Modulation of telomerase activity by telomere DNA-binding proteins in Oxytricha. *Genes Dev* 12: 1504-1514.

Garner, M.M. and M.B. Burg. 1994. Macromolecular crowding and confinement in cells exposed to hypertonicity. *Am. J. Physiol. Cell Physiol.* 266: 877-892.

Gartenberg, M.R., C. Ampe, T.A. Steitz, and D.M. Crothers. 1990. Molecular characterization of the GCN4-DNA complex. *Proc. Natl. Acad. Sci. USA* 87: 6034-6038.

Garvik, B., M. Carson, and L. Hartwell. 1995. Single-stranded DNA arising at telomeres in cdc13 mutants may constitute a specific signal for the RAD9 checkpoint. *Mol. Cell. Biol.* 15: 6128-6138.

Gehring, W.J., Y.Q. Qian, M. Billeter, K. Furukubo-Tokunaga, A.F. Schier, D. Resendez-Perez, M. Affolter, G. Otting, and K. Wuthrich. 1994. Homeodomain-DNA recognition. *Cell* 78: 211-223.

Giffin, W., H. Torrance, D.J. Rodda, G.G. Prefontaine, L. Pope, and R.J. Hache. 1996. Sequence-specific DNA binding by Ku autoantigen and its effects on transcription. *Nature* 380: 265-268.

Gilson, E., M. Roberge, R. Giraldo, D. Rhodes, and S.M. Gasser. 1993. Distortion of the DNA double helix by RAP1 at silencers and multiple telomeric binding sites. *J. Mol. Biol.* 231: 293-310.

Giraldo, R. and D. Rhodes. 1994. The yeast telomere-binding protein RAP1 binds to and promotes the formation of NDA quadruplexes in telomeric DNA. *EMBO J.* 13: 2411-2420.

Goldstein, S. 1990. Replicative senescence: the human fibroblast comes of age. Science 249: 1129-1133

Gottschling, D.E. and T.R. Cech. 1984. Chromatin structure of the molecular ends of Oxytricha macronuclear DNA: Phased nucleosomes and a telomeric complex. *Cell* 38: 501-510.

Gottschling, D.E. and V.A. Zakian. 1986. Telomere proteins: Specific recognition and protection of the natural termini of Oxytricha macronuclear DNA. *Cell* 47: 195-205.

Graham, I.R. and A. Chambers. 1994. Use of a selection technique to identify the diversity of binding sites for the yeast RAP1 transcription factor. *Nucl. Acids Res.* 22: 124-130.

Gravel, S., M. Larrivee, P. Labrecque, and R.J. Wellinger. 1998. Yeast Ku as a Regulator of Chromosomal DNA End Structure. *Science* 280: 741-744.

Gray, J.T., D.W. Celander, C.M. Price, and T.R. Cech. 1991. Cloning and expression of genes for the Oxytricha telomere-binding protein: Specific subunit interactions in the telomeric complex. *Cell* 67: 807-814.

Greenstein, D., S. Hird, R.H. Plasterk, Y. Andachi, Y. Kohara, B. Wang, M. Finney, and G. Ruvkun. 1994. Targeted mutations in the Caenorhabditis elegans POU homeo box gene

ceh- 18 cause defects in oocyte cell cycle arrest, gonad migration, and epidermal differentiation. *Genes Dev* 8: 1935-1948.

Greider, C.W. 1991. Telomerase is processive. *Mol. Cell. Biol.* 11: 4572-4580.

Greider, C.W. 1995. Telomerase biochemistry and regulation. In *Telomeres* (ed. E.H. Blackburn and C.W. Greider), pp. 35-69. Cold Spring Harbor Press, Cold Spring Harbor.

Greider, C.W. 1996. Telomere length regulation. *Ann. Rev. Biochem.* 65: 337-365.

Greider, C.W. and E.H. Blackburn. 1985. Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. *Cell* 43: 405-413.

Greider, C.W. and E.H. Blackburn. 1987. The telomere terminal transferase of tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell* 51: 887-898.

Greider, C.W. and E.H. Blackburn. 1989. A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis. *Nature* 337: 331-337.

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

Grunstein, M. 1997. Molecular model for telomeric heterochromatin in yeast. *Curr Opin Cell Biol* 9: 383-387.

Gu, Y., S. Jin, Y. Gao, D.T. Weaver, and F.W. Alt. 1997. Ku70-deficient embryonic stem cells have increased ionizing radiosensitivity, defective DNA end-binding activity, and inability to support V(D)J recombination. *Proc. Natl. Acad. Sci. USA* 94: 8076-8081.

Gualberto, A., R.M. Patrik, and K. Walsh. 1992. Nucleic acid specificity of a vertebrate telomere-binding protein: evidence for G-G base pair recognition at the core binding site. *Genes Dev.* 6: 815-824.

Hammond, P.W., T.N. Lively, and T.R. Cech. 1997. The anchor site of telomerase from Euplotes aediculatus revealed by photo-cross-linking to single- and double-stranded DNA primers. *Mol. Cell. Biol.* 17: 296-308.

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

Hardy, C.F.J., L. Sussel, and D. Shore. 1992. A RAP1-interacting protein involved in transcriptional silencing and telomere length regulation. *Genes Dev.* 6: 801-814.

Harley, C.B., A.B. Futcher, and C.W. Greider. 1990. Telomeres shorten during ageing of human fibroblasts. *Nature* 345: 458-460.

Harrington, L., T. McPhail, V. Mar, W. Zhou, R. Oulton, M.B. Bass, I. Arruda, and M.O. Robinson. 1997a. A mammalian telomerase-associated protein. *Science* 275: 973-977.

Harrington, L., W. Zhou, T. McPhail, R. Oulton, D.S. Yeung, V. Mar, M.B. Bass, and M.O. Robinson. 1997b. Human telomerase contains evolutionarily conserved catalytic and structural subunits. *Genes Dev.* 11: 3109-3115.

Harrington, L.A. and C.W. Greider. 1991. Telomerase primer specificity and chromosome healing. *Nature* 353: 451-454.

Hastie, N.D., M. Dempster, M.G. Dunlop, A.M. Thompson, D.K. Green, and R.C. Allshire. 1990. Telomere reduction in human colorectal carcinoma and with ageing. *Nature* 346: 866-868.

Hayflick, L. and P.S. Moorhead. 1961. The serial cultivation of human diploid cell strains. *Exp. Cell Res.* 25: 585-621.

Hecht, A., T. Laroche, S. Strahl-Bolsinger, S.M. Gasser, and M. Grunstein. 1995. Histone H3 and H4 N-termini interact with SIR3 and SIR4 proteins: a molecular model for the formation of heterochromatin in yeast. *Cell* 80: 583-592.

Hecht, A., S. Strahl-Bolsinger, and M. Grunstein. 1996. Spreading of transcriptional repressor SIR3 from telomeric heterochromatin. *Nature* 383: 92-96.

Henderson, E. 1995. Telomere DNA structure. In *Telomeres* (ed. E.H. Blackburn and C.W. Greider), pp. 11-35. Cold Spring Harbor Press, Cold Spring Harbor.

Henry, Y.A.L., A. Chambers, J.S.H. Tsang, A.J. Kingsman, and S.M. Kingsman. 1990. Characterisation of the DNA binding domain of the yeast RAP1 protein. *Proc. Natl. Acad. Sci. USA* 18: 2617-2623.

Hicke, B.J., D.W. Celander, G.H. MacDonald, C.M. Price, and T.R. Cech. 1990. Two versions of the gene encoding the 41-kilodalton subunit of the telomere binding protein of Oxytricha nova. *Proc. Natl. Acad. Sci. USA* 87: 1481-1485.

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

Hope, I.A. and K. Struhl. 1987. GCN4, a eukaryotic transcriptional activator protein, binds as a dimer to target DNA. *EMBO J.* 6: 2781-2784.

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

Inoue, K. and C.J. Sherr. 1998. Gene expression and cell cycle arrest mediated by transcription factor DMP1 is antagonized by D-type cyclins through a cyclin-dependent-kinase- independent mechanism. *Mol Cell Biol* 18: 1590-1600.

Ishikawa, F., M.J. Matunis, G. Dreyfuss, and T.R. Cech. 1993. Nuclear proteins that bind the pre-mRNA 3' splice site sequence r(UUAG/G) and the human telomeric DNA sequence $d(TTAGGG)_n$. *Mol. Cell. Biol.* 13: 4301-4310.

Jackson, S.P. 1996. The recognition of DNA damage. *Curr. Opin. Genet. Dev.* 6: 19-25.

Jhappan, C., H.C. Morse, 3rd, R.D. Fleischmann, M.M. Gottesman, and G. Merlino. 1997. DNA-PKcs: a T-cell tumour suppressor encoded at the mouse scid locus. *Nat Genet* 17: 483-486.

Jorgensen, C.I., B.H. Kallipolitis, and P. Valentin-Hansen. 1998. DNA-binding characteristics of the Escherichia coli CytR regulator: a relaxed spacing requirement between operator half-sites is provided by a flexible, unstructured interdomain linker. *Mol Microbiol* 27: 41-50.

Kang, C., X. Zhang, R. Ratliff, R. Moyzis, and A. Rich. 1992. Crystal structure of four-stranded Oxytricha telomeric DNA. *Nature* 356: 126-131.

Kim, N.W., M.A. Piatyszek, K.R. Prowse, C.B. Harley, M.D. West, P.L.C. Ho, G.M. Coviello, W.E. Wright, S.L. Weinrich, and J.W. Shay. 1994. Specific association of human telomerase activity with immortal cells and cancer. *Science* 266: 2011-2015.

Kipling, D. and H.J. Cooke. 1990. Hypervariable ultra-long telomeres in mice. *Nature* 347: 400-402.

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

Klempnauer, K.-H. and A.W. Sippel. 1987. The highly conserved amino-termina region of the protein encoded by the v-myb oncogene functions as a DNA-binding domain. *EMBO J.* 6: 2719-2725.

Klobutcher, L.A., M.T. Swanton, P. Donini, and D.M. Prescott. 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.

Kolquist, K.A., L.W. Ellisen, C.M. Counter, M. Meyerson, L.K. Tan, R.A. Weinberg, D.A. Haber, and W.L. Gerald. 1998. Expression of TERT in early premalignant lesions and a subset of cells in normal tissues [see comments]. *Nat Genet* 19: 182-186.

Konig, P., L. Fairall, and D. Rhodes. 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.

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

Konig, P. and D. Rhodes. 1997. Recognition of telomeric DNA. *Trends Biochem. Sci.* 22: 43-47.

Konkel, L.M., S. Enomoto, E.M. Chamberlain, P. McCune-Zierath, S.J. Iyadurai, and J. Berman. 1995. A class of single-stranded telomeric DNA-binding proteins required for Rap1p localiztion in yeast nuclei. *Proc. Natl. Acad. Sci. USA* 92: 5558-5562.

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

Kotlarz, D., A. Fritsch, and H. Buc. 1986. Variations of intramolecular ligation rates allow the detection of protein-induced bends in DNA. *EMBO J.* 5: 799-803.

Krauskopf, A. and E.H. Blackburn. 1996. Control of telomere growth by interactions of RAP1 with the most distal telomeric repeats. *Nature* 383: 354-357.

Kurokawa, R., V.C. Yu, A. Naar, S. Kyakumoto, Z. Han, S. Silverman, M.G. Rosenfeld, and C.K. Glass. 1993. Differential orientations of the DNA-binding domain and carboxy-terminal dimerization interface regulate binding site selection by nuclear receptor heterodimers. *Genes Dev* 7: 1423-1435.

Kyrion, G., K.A. Boakye, and A.J. Lustig. 1992. C-terminal truncation of RAP1 results in the deregulation of telomere size, stability, and function in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 12: 5159-5173.

LaBranche, H., S. Dupuis, Y. Ben-David, M.R. Bani, R.J. Wellinger, and B. Chabot. 1998. Telomere elongation by hnRNP A1 and a derivative that interacts with telomeric repeats and telomerase [see comments]. *Nat Genet* 19: 199-202.

Laroche, T., S.G. Martin, M. Gotta, H.C. Gorham, F.E. Pryde, E.J. Louis, and S.M. Gasser. 1998. Mutation of yeast Ku genes disrupts the subnuclear organization of telomeres [In Process Citation]. *Curr Biol* 8: 653-656.

Larson, D.D., E.A. Spangler, and E.H. Blackburn. 1987. Dynamics of telomere length variation in Tetrahymena thermophila. *Cell* 50: 477-483.

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

Lee, M.S. and E.H. Blackburn. 1993. Sequence-specific DNA primer effects on telomerase polymerization activity. *Mol. Cell. Biol.* 13: 6586-6599.

Lee, M.S., R.C. Gallagher, J. Bradley, and E.H. Blackburn. 1993. In vivo and in vitro studies of telomeres and telomerase. *Cold Spring Harbor Symposia on quantitative biology* LVIII: 707-718.

Lendvay, T.S., D.K. Morris, J. Sah, B. Balasubramanian, and V. Lundblad. 1996. Senescence mutants of Saccharomyces cerevisiae with a defect in telomere replication identify three additional EST genes. *Genetics* 144: 1399-1412.

Levis, R.W., R. Ganesan, K. Houtchens, L.A. Tolar, and F.-M. Sheen. 1993. Transposons in place of telomeric repeats at a Drosophila telomere. *Cell* 75: 1083-1093.

Li, B. and A.J. Lustig. 1996. A novel mechanism for telomere size control in Saccharomyces cerevisiae. *Genes Dev.* 10: 1310-1326.

Liang, F., P.J. Romanienko, D.T. Weaver, P.A. Jeggo, and M. Jasin. 1996. Chromosomal double-strand break repair in Ku80-deficient cells. *Proc. Natl. Acad. Sci. USA* 93: 8829-8933.

Lieber, M.R., U. Grawunder, X. Wu, and M. Yaneva. 1997. Tying loose ends: roles of Ku and DNA-dependent protein kinase in the repair of double-strand breaks. *Curr Opin Genet Dev* 7: 99-104.

Lin, J.-J. and V.A. Zkian. 1996. The Saccharomyces CDC13 protein is a single-strand TG1-3 telomeric DNA binding protein in vitro that affects telomere behavior in vivo. *Proc. Natl. Acad. Sci. USA* 93: 13760-13765.

Lin, J.J. and V.A. Zakian. 1994. Isolation and characterization of two Saccharomyces cerevisiae genes that encode proteins that bind to (TG1-3)n single strand telomeric DNA in vitro [published erratum appears in Nucleic Acids Res 1994 Dec 11;22(24):5516]. *Nucleic Acids Res* 22: 4906-4913.

Lingner, J. and T.R. Cech. 1996. Purification of telomerase from Euplotes aediculatus: requirement of a primer 3' overhang. *Proc. Natl. Acad. Sci. USA* 93: 10712-10717.

Lingner, J., T.R. Cech, T.R. Hughes, and V. Lundblad. 1997a. Three Ever Shorter Telomere (EST) genes are dispensable for in vitro yeast telomerase activity. *Proc. Natl. Acad. Sci. USA* 94: 11190-11195.

Lingner, J., T.R. Hughes, A. Shevchenko, M. Mann, V. Lundblad, and T.R. Cech. 1997b. Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science* 276: 561-567.

Lipsick, J.S. 1996. One billion years of Myb. *Oncogene* 13: 223-235.

Liu, Z. and B.-K. Tye. 1991. A yeast protein that binds to vertebrate telomeres and conversed yeast telomeric junctions. *Genes Dev.* 5: 49-59.

Longtine, M.S., N.M. Wilson, M.E. Petracek, and J. Berman. 1989. A yeast telomere binding activity binds to two related telomere sequence motifs anid is indistinguishable from RAP1. *Curr. Genet.* 16: 225-239.

Louis, E.J. and J.E. Haber. 1992. The structure and evolution of subtelomeric Y' repeats in Saccharomyces cerevisiae. *Genetics* 131: 559-574.

Ludérus, M.E.E., B. van Steensel, L. Chong, O.C.M. Sibon, F.F.M. Cremers, and T. de Lange. 1996. Structure, subnuclear distribution, and nuclear matrix association of the mammalian telomeric complex. *J. Cell. Biol.* 135: 867-883.

Lugert, T. and W. Werr. 1994. A novel DNA-binding domain in the Shrunken initiator-binding protein (IBP1). *Plant Mol. Biol.* 25: 493-506.

Lundblad, V. and E.H. Blackburn. 1993. An alternative pathway for yeast telomere maintenance rescues est1- senescence. *Cell* 73: 347-360.

Lundblad, V. and J.W. Szostak. 1989. A mutant with a defect in telomere elongation leads to senescence in yeast. *Cell* 57: 633-643.

Lustig, A.J., S. Kurtz, and D. Shore. 1990. Involvement of the silencer and UAS binding protein RAP1 in regulation of telomere length. *Science* 250: 549-553.

Lustig, A.J., C. Liu, C. Zhang, and J.P. Hanish. 1996. Tethered Sir3p nucleates silencing at telomeres and internal loci in Saccharomyces cerevisiae. *Mol. Cell. Biol.* 16: 2483-2495.

Lydall, D. and T. Weinert. 1995. Yeast checkpoint genes in DNA damage processing: implications for repair and arrest. *Science* 270: 1488-1491.

Makarov, V., Y. Hirose, and J.P. Langmore. 1997. Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell* 88: 657-666.

Makarov, V.L., S. Lejnine, J. Bedoyan, and J.P. Langmore. 1993. Nucleosomal organization of telomere-specific chromatin in Rat. *Cell* 73: 775-787.

Mantell, L.L. and C.W. Greider. 1994. Telomerase activity in germline and embryonic cells of Xenopus. *EMBO J.* 13: 3211-3217.

Marcand, S., S.W. Buck, P. Moretti, E. Gilson, and D. Shore. 1996. Silencing of genes at nontelomeric sites in yeast is controlled by sequestration of silencing factors at telomeres by Rap 1 protein. *Genes Dev.* 10: 1297-1309.

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

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

McClintock, B. 1941. The stability of broken ends of chromosomes in zea mays. *Genetics* 26: 234-282.

McClintock, B. 1942. The fusion of broken ends of chromosomes following nuclear fusion. *Proc. Natl. Acad. Sci. USA* 28: 458-463.

McCormick, R.J., T. Badalian, and D.E. Fisher. 1996. The leucine zipper may induce electrophoretic mobility anomalies without DNA bending. *Proc Natl Acad Sci U S A* 93: 14434-14439.

McCormick-Graham, M.,Haynes, W. J., and Romero, D. P. 1997. Variable telomeric repeat synthesis in Paramecium tetraurelia is consistent with misincorporation by telomerase. EMBO J. 16: 3233-3242

McEachern, M.J. and E.H. Blackburn. 1995. Runaway telomere elongation caused by telomerase RNA gene mutations. *Nature* 376: 403-409.

McEachern, M.J. and E.H. Blackburn. 1996. Cap-prevented recombination between terminal telomeric repeat arrays (telomere CPR) maintains telomeres in Kluyveromyces lactis lacking telomerase. *Genes Dev.* 10: 1822-1834.

McElligott, R. and R.J. Wellinger. 1997. The terminal DNA structure of mammalian chromosomes. *EMBO J.* 16: 3705-3714.

McKay, S.J. and H. Cooke. 1992. hnRNP A2/B1 binds specifically to single stranded vertebrate telomeric repeat $TTAGGG_n$. *Nucl. Acids Res.* 20: 6461-6464.

Melek, M., E.C. Green, and D.E. Shippen. 1996. Processing of nontelomeric 3' ends by telomerase: Default template alignment and endonucleotide cleavage. *Mol. Cell. Biol.* 16: 3437-3445.

Meyerson, M., C.M. Counter, E.N. Eaton, L.W. Ellisen, P. Steiner, S.D. Caddle, L. Ziaugra, R.L. Beijersbergen, M.J. Davidoff, Q. Liu, S. Bacchetti, D.A. Haber, and R.A. Weinberg. 1997. hEST2, the putative human telomerase catalytic subunit gene, is up-regulated in tumor cells and during immortalization. *Cell* 90: 785-795.

Milne, G.T., S. Jin, K.B. Shannon, and D.T. Weaver. 1996. Mutations in two Ku homologs define a DNA end-joining repair pathway in Saccharomyces cerevisiae. *Mol Cell Biol* 16: 4189-4198.

Mimori, T. and J.A. Hardin. 1986. Mechanism of interaction between Ku protein and DNA. *J. Biol. Chem.* 261: 10375-10379.

Mimori, T., J.A. Hardin, and J.A. Steitz. 1986. Characterization of the DNA-binding protein antigen Ku recognized by autoantibodies from patients with rheumatic disorders. *J. Biol. Chem.* 261: 2274-2278.

Moretti, P., K. Freeman, L. Coodly, and D. Shore. 1994. Evidence that a complex of SIR proteins interacts with the silencer and telomere-binding protein RAP1. *Genes Dev.* 8: 2257-2269.

Morin, G.B. 1989. The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. *Cell* 59: 521-529.

Morin, G.B. 1991. Recognition of a chromosome truncation site associated with alpha-thalassaemia by human telomerase. *Nature* 353: 454-456.

Morris, D.K. and V. Lundblad. 1997. Programmed translational frameshifting in a gene required for yeast telomere replication. *Curr. Biol.* 7: 969-976.

Moyzis, R.K., J.M. Buckingham, L.S. Cram, M. Dani, L.L. Deaven, M.D. Jones, J. Meyne, R.L. Ratliff, and J.-R. Wu. 1988. A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. *Proc. Natl. Acad. Sci. USA* 85: 6622-6626.

Muller, H.J. 1938. The remaking of chromosomes. *The Collecting Net - Woods Hole* 13: 181-195.

Murchie, A.I.H. and D.M.J. Lilley. 1994. Tetraplex folding of telomere sequences and the inclusion of adenine bases. *EMBO* 13: 993-1001.

Nakamura, T.M. and T.R. Cech. 1998. Reversing time: origin of telomerase. *Cell* 92: 587-590.

Nakamura, T.M., G.B. Morin, K.B. Chapman, S.L. Weinrich, W.H. Andrews, J. Lingner, C.B. Harley, and T.R. Cech. 1997. Telomerase catalytic subunit homologs from fission yeast and human. *Science* 277: 955-959.

Nakayama, J., M. Saito, H. Nakamura, A. Matsuura, and F. Ishikawa. 1997. TLP1: A gene encoding a prein component of mammalian telomerase is a novel member of WD repeats family. *Cell* 88: 875-884.

Nimmo, E.R., G. Cranston, and R.C. Allshire. 1994. Telomere-associated chromosome breakage in fission yeast results in variegated expression of adjacent genes. *EMBO J.* 13: 3801-3811.

Nimmo, E.R., A.L. Pidoux, P.E. Perry, and R.C. Allshire. 1998. Defective meiosis in telomere-silencing mutants of Schizosaccharomyces pombe. *Nature* 392: 825-828.

Nugent, C.I., G. Bosco, L.O. Ross, S.K. Evans, A.P. Salinger, J.K. Moore, J.E. Haber, and V. Lundblad. 1998. Telomere maintenance is dependent on activities required for end repair of double-strand breaks [In Process Citation]. *Curr Biol* 8: 657-660.

Nugent, C.I., T.R. Hughes, N.F. Lue, and V. Lundblad. 1996. Cdc13p: a single-strand telomeric DNA-binding protein with a dual role in yeast telomere maintenance. *Science* 274: 249-252.

Nussenzweig, A., C. Chen, V.d.C. Soares, M. Sanchez, K. Sokol, M.C. Nussenzweig, and G.C. Li. 1996. Requirement for Ku80 in growth and immunoglobulin V(D)J recombination. *Nature* 382: 551-555.

Ogata, K., H. Hojo, S. Aimoto, T. Nakai, H. Nakamura, A. Sarai, S. Ishii, and Y. Nishimura. 1992. Solution structure of a DNA-binding unit of Myb: a helix-turn-helix-related motif with conserved tryptophans forming a hydrophobic core. *Proc. Natl. Acad. Sci. USA* 89: 6428-6432.

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

Ohi, R., D. McCollum, B. Hirani, G.J. Den Haese, X. Zhang, J.D. Burke, K. Turner, and K.L. Gould. 1994. The Schizosaccharomyces pombe cdc5+ gene encodes an essential protein with homology to c-Myb. *EMBO J.* 13: 471-483.

Olovnikov, A.M. 1973. A theory of marginotomy. *J. Theor. Biol.* 41: 181-190.

Paillard, S. and F. Strauss. 1991. Analysis of the mechanism of interaction of simian Ku protein with DNA. *Nucl. Acids Res.* 19: 5619-5624.

Palladino, F., T. Laroche, E. Gilson, A. Axelrod, L. Pillus, and S.M. Gasser. 1993. SIR3 and SIR4 proteins are required for the positioning and integrity of yeast telomeres. *Cell* 75: 543-555.

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

Petracek, M.E., L.M. Konkel, M.L. Kable, and J. Berman. 1994. A Chlamydomonas protein that binds single-stranded G-strand telomere DNA. *EMBO J.* 13: 3648-3658.

Pluta, A.F. and V.A. Zakian. 1989. Recombination occurs during telomere formation in yeast. *Nature* 337: 429-433.

Polotnianka, R.M., J. Li, and A.J. Lustig. 1998. The yeast Ku heterodimer is essential for protection of the telomere against nucleolytic and recombinational activities [In Process Citation]. *Curr Biol* 8: 831-834.

Porter, S.E., P.W. Greenwell, K.B. Ritchie, and T.D. Petes. 1996. The DNA-binding protein Hdf1p (a putative Ku homologue) is required for maintaining normal telomere length in Saccharomyces cerevisiae. *Nucl. Acids Res.* 24: 582-585.

Prescott, D.M. 1994. The DNA of ciliated protozoa. *Microbiol Rev* 58: 233-267.

Prescott, J. and E.H. Blackburn. 1997. Functionally interacting telomerase RNAs in the yeast telomerase complex. *Genes Dev.* 11: 2790-2800.

Price, C.M. 1990. Telomere structure in Euplotes crassus: Characterization of DNA-protein interactions and isolation of a telomere-binding protein. *Mol. Cell. Biol.* 10: 3241-3431.

Price, C.M. and T.R. Cech. 1987. Telomeric DNA-protein interactions of Oxytricha macronuclear DNA. *Genes Dev.* 1: 783-793.

Price, C.M. and T.R. Cech. 1989. Properties of the telomeric DNA-binding protein from Oxytricha nova. *Biochem.* 28: 769-774.

Prowse, K. and C.W. Greider. 1995. Developmental and tissue specific regulation of mouse telomerase and telomere length. *Proc. Natl. Acad. Sci. USA* 92: 4818-4822.

Raghuraman, M.K., C.J. Dunn, B.J. Hicke, and T.R. Cech. 1989. Oxytricha telomeric nucleoprotein complexes reconstituted with synthetic DNA. *Nucl. Acids Res.* 17: 4235-4252.

Rastinejad, F., T. Perlmann, R.M. Evans, and P.B. Sigler. 1995. Structural determinants of nuclear receptor assembly on DNA direct repeats [see comments]. *Nature* 375: 203-211.

Rathmell, W.K. and G. Chu. 1994. A DNA end-binding factor involved in double-strand break repair and V(D)J recombination. *Mol. Cell. Biol.* 14: 4741-4747.

Riethman, H.C., R.K. Moyzis, J. Meyne, D.T. Burke, and M.V. Olson. 1989. Cloning human telomeric DNA fragments into Saccharomyces cerevisiae using a yeast-artificial-chromosome vector. *Proc. Natl. Acad. Sci. USA* 86: 6240-6244.

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

Runge, K.W. and V.A. Zakian. 1996. TEL2, an essential gene required for telomere length regulation and telomere position effect in Saccharomyces cerevisiae. *Mol. Cell. Biol.* 16: 3094-3105.

Sandell, L. and V. Zakian. 1993. Loss of a yeast telomere: arrest, recovery, and chromosome loss. *Cell* 75: 729-741.

Sauer, R.T., D.L. Smith, and A.D. Johnson. 1988. Flexibility of the yeast alpha 2 repressor enables it to occupy the ends of its operator, leaving the center free. *Genes Dev* 2: 807-816.

Schuck, P. 1997. Use of surface plasmon resonance to probe the equilibrium and dynamic aspects of interactions between biological macromolecules. *Annu Rev Biophys Biomol Struct* 26: 541-566.

Sen, D. and W. Gilbert. 1988. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* 334: 364-366.

Shampay, J. and E.H. Blackburn. 1988. Generation of telomere-length heterogeneity in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. USA* 85: 534-538.

Shampay, J., J.W. Szostak, and E.H. Blackburn. 1984. DNA sequences of telomeres maintained in yeast. *Nature* 310: 154-157.

Sheng, H., Z. Hou, T. Schierer, D.L. Dobbs, and E. Henderson. 1995. Identification and characterization of a putative telomere binding protein from Tetrahymena thermophila. *Mol. Cell. Biol.* 15: 1144-1153.

Shippen, D.E., E.H. Blackburn, and C.M. Price. 1994. DNA bound by the Oxytricha telomere protein is accessible to telomerase and other DNA polymerases. *Proc. Natl. Acad. Sci. USA* 91: 405-409.

Shippen-Lentz, D. and E.H. Blackburn. 1990. Functional evidence for an RNA template in telomerase. *Science* 247: 546-552.

Shore, D. 1994. RAP1: a protean regulator in yeast. *Trends Gen.* 10: 408-412.

Shore, D. 1995. Telomere Position Effects and Transcriptional Silencing in the Yeast *Saccharomyces cerevisiae*. In *Telomeres* (ed. E.H.B.a.C.G. Greider), pp. 139-191. CSHL Press.

Shore, D., J. Langowski, and R.L. Baldwin. 1981. DNA flexibility studied by covalent closure of short fragments into circles. *Proc Natl Acad Sci U S A* 78: 4833-4837.

Shore, D. and K. Nasmyth. 1987. Purification and cloning of a DNA binding protein from yeast that binds to both silencer and activator elements. *Cell* 51: 721-732.

Siede, W., A.A. Friedl, I. Dianova, F. Eckardt-Schupp, and E.C. Friedberg. 1996. The Saccharomyces cerevisiae Ku autoantigen homologue affects radiosensitivity only in the absence of homologous recombination. *Genetics* 142: 91-102.

Sinclair, D.A., K. Mills, and L. Guarente. 1998. Molecular mechanisms of yeast aging. *Trends Biochem Sci* 23: 131-134.

Singer, M.S. and D.E. Gottschling. 1994. TLC1: template RNA component of Saccharomyces cerevisiae telomerase. *Science* 266: 404-409.

Sitlani, A. and D.M. Crothers. 1996. Fos and jun do not bend th AP-1 recognition site. *Proc. Natl. Acad. Sci. USA* 93: 3248-3252.

Smith, F.W. and J. Feigon. 1992. Quadruplex structure of Oxytricha telomeric DNA oligonucleotides. *Nature* 356: 164-168.

Smith, F.W. and J. Feigon. 1993. Strand orientation in the DNA quadruplex formed from the Oxytricha telomere repeat oligonucleotide d($G_4T_4G_4$) in solution. *Biochem.* 32: 8682-8692.

Smith, S. and T. de Lange. 1997. TRF1, a mammalian telomeric protein. *Trends in Genetics* 13: 21-26.

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

Stokes, D.G. and R.P. Perry. 1995. DNA-binding and chromatin localization properties of CHD1. *Mol. Cell. Biol.* 15: 2745-2753.

Sugawara, N. and J.W. Szostak. 1986. Telomeres of *Schizosaccharomyces pombe*. *Yeast (Suppl.)* 1: S373.

Sundquist, W.I. and A. Klug. 1989. Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. *Nature* 342: 825-829.

Swirnoff, A.H. and J. Milbrandt. 1995. DNA-binding specificity of NGFI-A and related zinc finger transcription factors. *Mol Cell Biol* 15: 2275-2287.

Szostak, J.W. and E.H. Blackburn. 1982. Cloning yeast telomeres on linear plasmid vectors. *Cell* 29: 245-255.

Tanikawa, J., T. Yasukawa, M. Enari, K. Ogata, Y. Nishimura, S. Ishii, and A. Sarai. 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.

Tommerup, H., A. Dousmanis, and T. de Lange. 1994. Unusual chromatin in human telomeres. *Mol. Cell. Biol.* 14: 5777-5785.

Tsukamoto, Y., J. Kato, and H. Ikeda. 1996. Hdf1, a yeast Ku-protein homologue, is involved in illegitimate recombination, but not in homologous recombination. *Nucl. Acids Res.* 24: 2067-2072.

Tsukamoto, Y., J. Kato, and H. Ikeda. 1997a. Budding yeast Rad50, Mre11, Xrs2, and Hdf1, but not Rad52, are involved in the formation of deletions on a dicentric plasmid. *Mol Gen Genet* 255: 543-547.

Tsukamoto, Y., J. Kato, and H. Ikeda. 1997b. Silencing factors participate in DNA repair and recombination in Saccharomyces cerevisiae [see comments]. *Nature* 388: 900-903.

Tuteja, N., R. Tuteja, A. Ochem, P. Taneja, N.W. Huang, A. Simoncsits, S. Susic, K. Rahman, L. Marusic, J. Chen, J. Zhang, S. Wang, S. Pongor, and A. Falaschi. 1994. Human DNA helicase II: a novel DNA unwidning enzyme identified as the Ku autoantigen. *EMBO J.* 13: 4991-5001.

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

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

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

Vaziri, H. and S. Benchimol. 1998. Reconstitution of telomerase activity in normal human cells leads to elongation of telomeres and extended replicative life span. *Curr. Biol.* 8: 279-282.

Vermeesch, J.R. and C.M. Price. 1994. Telomeric DNA sequence and structure following de novo telomere synthesis in Euplotes crassus. *Mol. Cell. Biol.* 14: 554-566.

Virta-Pearlman, V., D.K. Morris, and V. Lundblad. 1996. Est1 has the properties of a single-stranded telomere end-binding protein. *Genes Dev.* 10: 3094-3104.

Wang, H. and E.H. Blackburn. 1997. De novo telomere addition by Tetrahymena telomerase in vitro. *Embo J* 16: 866-879.

Wang, W., R. Skopp, M. Scofield, and C. Price. 1992. Euplotes crassus has genes encoding telomere-binding proteins and telomere-binding protein homologs. *Nucl. Acids Res.* 20: 6621-6629.

Wang, Y. and D.J. Patel. 1993. Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex. *Structure* 1: 263-282.

Wang, Y. and D.J. Patel. 1995. Solution structure of the Oxytricha telomeric repeat d[G4(T4G4)3] G- tetraplex. *J Mol Biol* 251: 76-94.

Watson, J.D. 1972. Origin of concatemeric T7 DNA. *Nature* 239: 197-201.

Weinrich, S.L., R. Pruzan, L. Ma, M. Ouellette, V. Tesmer, S.E. Holt, A.G. Bodnar, S. Lichtsteiner, N.W. Kim, J.B. Trager, R.D. Taylor, R. Carlos, W.H. Andrews, W.E. Wright, J.W. Shay, C.B. Harley, and G.B. Morin. 1997. Reconstitution of human telomerase withthe template RNA component hTR and the catalytic protein subunit hTRT. *Nature Genetics* 17: 498-502.

Wellinger, R.J., A.J. Wolf, and V.A. Zakian. 1993. Saccharomyces telomeres acquires single-strand TG_{1-3} tails late in S-phase. *Cell* 72: 51-60

Wellinger, R.J., K. Ethier, P. Labrecque, and V.A. Zakian. 1996. Evidence for a new step in telomere maintenance. *Cell* 85: 423-433.

Weston, K. 1992. Extension of the DNA binding consensus of the chicken c-Myb and v-Myb proteins. *Nucleic Acids Res* 20: 3043-3049.

Wilkie, A.O.M., J. Lamb, P.C. Harris, R.D. Finney, and D.R. Higgs. 1990. A truncated human chromosome 16 associated with a thalassaemia is stabilized by addition of telomeric repeat (TTAGGG)n. *Nature* 346: 868-871.

Williamson, J.R., M.K. Raghuraman, and T.R. Cech. 1989. Monovalent cation-induced structure of telomeric DNA: The G-quartet model. *Cell* 59: 871-880.

Wilm, M., A. Shevchenko, T. Houthaeve, S. Breit, L. Schweigerer, T. Fotsis, and M. Mann. 1996. Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature* 379: 466-469.

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

Wong, M.W., R.W. Henry, B. Ma, R. Kobayashi, N. Klages, P. Matthias, M. Strubin, and N. Hernandez. 1998. The large subunit of basal transcription factor SNAPc is a Myb domain protein that interacts with Oct-1. *Mol Cell Biol* 18: 368-377.

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

Wotton, D. and D. Shore. 1997. A novel Rap1p-interacting factor, Rif2p, cooperates with Rif1p to regulate telomere length in Saccharomyces cerevisiae. *Genes Dev* 11: 748-760.

Wright, J.H., D.E. Gottschling, and V.A. Zakian. 1992. Saccharomyces telomeres assume a non-nucleosomal chromatin structure. *Genes Dev.* 6: 197-210.

Wright, J.H. and V.A. Zakian. 1995. Protein-DNA interactions in soluble telosomes from Saccharomyces cerevisiae. *Nucl. Acids Res.* 23: 1454-1460.

Wright, W.E., V.M. Tesmer, K.E. Huffman, S.D. Levene, and J.W. Shay. 1997. Normal human chromosomes have long G-rich telomeric overhangs at one end. *Genes Dev.* 11: 2801.

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

Yoo, S. and W.S. Dynan. 1998. Characterization of the RNA binding properties of Ku protein. *Biochemistry* 37: 1336-1343.

Yu, G.-L., J.D. Bradley, L.D. Attardi, and E.H. Blackburn. 1990. In vivo alteration of telomere sequences and senescence caused by mutated Tetrahymena telomerase RNAs. *Nature* 344: 126-132.

Zahler, A.M., J.R. Williamson, T.R. Cech, and D.M. Prescott. 1991. Inhibition of telomerase by G-quartet DNA structures. *Nature* 350: 718-720.

Zhang, L. and L. Guarente. 1996. The C6 zinc cluster dictates asymmetric binding by HAP1. *Embo J* 15: 4676-4681.

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

Zhu, C., M.A. Bogue, D.S. Lim, P. Hasty, and D.B. Roth. 1996. Ku86-deficient mice exhibit severe combined immunodeficiency and defective processing of V(D)J recombination intermediates. *Cell* 86: 379-389.

Zijlmans, J. M., Martens, U. M., Poon, S. S., Raap, A. K., Tanke, H. J., Ward, R. K., and P. M. Lansdorp. 1997. Telomeres in the mouse have large inter-chromosomal variations in the number of T2AG3 repeats. Proc. Natl. Acad. Sci. USA 94: 7423-8