Human telomeres are composed of long arrays of TTAGGG repeats that form a nucleoprotein complex required for the protection and replication of chromosome ends. One component of human telomeres is the TTAGGG repeat binding factor 1 (TRF1), a ubiquitously expressed protein, related to the proto-oncogene Myb, that is present at telomeres throughout the cell cycle. Recent evidence has implicated TRF1 in the control of telomere length. TRF1 is proposed to be an inhibitor of telomerase, acting in cis to limit the elongation of individual chromosome ends. Here we report the cloning of TRF2, a distant homologue of TRF1 that carries a very similar Myb-related DNA-binding motif. Like TRF1, TRF2 was ubiquitously expressed, bound specifically to duplex TTAGGG repeats in vitro and localized to all human telomeres in metaphase chromosomes. TRF2 was shown to have an architecture similar to that of TRF1 in that it carries a C-terminal Myb motif and a large TRF1-related dimerization domain near its N terminus. However, the dimerization domains of TRF1 and TRF2 did not interact, suggesting that these proteins exist predominantly as homodimers. While having similar telomere binding activity and domain organization, TRF2 differed from TRF1 in that its N terminus was basic rather than acidic, and TRF2 was much more conserved than TRF1. The results indicate that the TTAGGG repeat arrays at the ends of human and mouse chromosomes bind to two related proteins. Because TRF1 and TRF2 showed significant differences, we suggest that these factors have distinct functions at telomeres.

An anonymous cDNA fragment encoding a TRF1-related Myb motif was reported in the database. This sequence information was used in a combination of polymerase chain reaction (PCR) and hybridization strategies to isolate the full-length human and mouse cDNAs representing this protein (called TRF2, for TTAGGG repeat binding factor 2; see Methods). The human TRF2 cDNA hybridized to a 3.1-kb mRNA with the same ubiquitously expressed pattern as TRF1 (Fig. 1a).

Conceptual translation of the mouse and human TRF2 cDNAs revealed two open reading frames, starting with Kozak consensus initiation sites, that predicted two closely related proteins of approximately 500 amino acids. In agreement with the slightly longer primary sequence of TRF2, in vitro translated human and mouse TRF2 proteins displayed a larger apparent molecular weight than human and mouse TRF1 (Fig. 1b). Alignment of the predicted sequences of TRF1 and TRF2 showed that the strongest conservation between these proteins is in their C-terminal Myb domains (Fig. 1c,d). In addition, a moderate level of conservation (27% identity) was observed in the region from positions 45 to 245 in TRF2, coinciding with the previously identified dimerization domain of TRF1 (Fig. 1d; ref. 6). Thus, TRF2 had a similar overall architecture as TRF1 with a C-terminal Myb-related DNA-binding motif and a putative dimerization domain in the N-terminal half of the protein (Fig. 1d). However, a striking difference was that while both mouse and human TRF1 have an acidic amino

The Rockefeller University, 1230 York Avenue, New York, New York 10021, USA. Correspondence should be addressed to T.D.L. e-mail: delange@rockvax.rockefeller.edu

nature genetics volume 10 october 1997

231
[TTAGGG]$_n$ binding activity of Trf2AN-49 in the presence of six different duplex telomeric sequences used as competitor DNAs (Fig. 2c). The competition experiments showed that the sequence preference of Trf2AN-49 is indistinguishable from that of TRF1, both proteins binding much better to TTAGGG repeats than to TTAGGC, TTAGGGGG and TTAGG repeats. Similar to TRF1 (L.C. and T.D.L., unpublished observations), Trf2AN-49 bound weakly to the TTTGGGG repeats from Tetrahymena telomeres (Fig. 2c, lanes 5,6). As Trf2 AN-49 binding is competed for by circular plasmids carrying TTAGGG repeat arrays (Fig. 2c), there is no in vitro requirement for a DNA end near the TRF2 binding site. Competition experiments also showed that, like TRF1 (ref. 1), Trf2AN-49 failed to bind to single-stranded [TTAGGG]$_6$ and [CCC- TAA]$_6$ oligonucleotides (Fig. 2c, lanes 15-18). These results were consistent with the high degree of sequence similarity of the Myb motifs in TRF1 and TRF2 and suggested that TRF2, like TRF1, could bind specifically to telomeric DNA in vivo.

To determine whether TRF2 is a component of the mammalian telomeric complex, full-length FLAG-tagged mouse Trf2 was expressed in transiently transfected HeLa cells. Indirect immunofluorescence with a monoclonal antibody to the FLAG epitope (M2) revealed a punctate pattern in interphase nuclei (Fig. 3a), as expected if the epitope-tagged Trf2 protein specifically localized to telomeres. Labelling of the telomeres in the same nuclei by fluorescence in situ hybridization with a TTAGGG repeat specific RNA probe (Fig. 3b) revealed that most telomeric loci coincided with an Trf2 signal (Fig. 3c), indicative of a telomeric localization of the epitope-tagged protein. However, some of the Trf2 signals were not obviously associated with telomeric DNA, and some of the telomeric loci did not contain detectable Trf2 (Fig. 3d and data not shown). The reason for this partial co-localization of TRF2 with telomeric DNA is not clear at this stage. Because a similar pattern is observed for the endogenous TRF2 protein (D.B., B. van Steensel & T.D.L., unpublished observations), the subnuclear localization of Trf2 in Fig. 3 is unlikely to represent an artifact of the transfection system. Transfected Trf2 was also demonstrable at the ends of mitotic chromosomes. All HeLa metaphase chromosomes were found to contain a Trf2 signal at their termini (Fig. 3d), and no other location for the FLAG-tagged Trf2 protein was noted. Similarly, FLAG-tagged Trf2AN-49 was found to localize to HeLa chromosome ends, and we observed the localization of epitope-tagged and GFP-tagged Trf2 to telomeres in transfected NIH3T3 cells (data not shown). Thus, according to these indirect methods, TRF2 appears to be an integral component of the telomeric complex associated with telomeric DNA in interphase and mitosis. The telomeric localization of

---

**Fig. 1** Characterization of the primary sequence and expression of TRF2. a, Northern blot (Clontech) containing poly-adenylated RNAs from the indicated human tissues probed for TRF2, TRF1 and &beta;-actin expression. TRF1 probes also detect a second 1.8 kb mRNA of unknown structure that is not shown here. For TRF2, only a single mRNA is detected. b, SDS-PAGE of 35S-methionine-labelled in vitro translation products obtained with cDNAs encoding mouse and human TRF1 and TRF2 (as indicated above the lanes). c, Alignment of the primary sequences of mouse and human TRF1 and TRF2. Identical amino acids are highlighted. The likely positions of the three helices constituting the Myb domain (as inferred from the NMR structure of &epsilon;Myb) are indicated below the sequence. d, Schematic representation of the domain structure of TRF1 and TRF2 and their sequence similarity. The data on TRF1 are from Broccoli et al. A small segment of the TRF1 sequences that is absent in a shorter form of the protein, probably due to alternative splicing, is indicated in the TRF1 sequence (ALT SPL) (J. Karlsceder and T.D.L., unpublished observations). The asterisk in c denotes an alanine residue that was found to be absent in one of the HeLa cDNAs we sequenced.
TRF2 in metaphase cells was corroborated by detection of the endogenous TRF2 protein with a TRF2-specific antibody (D.B., B. van Steensel & T.d.L., unpublished observations; T. Bilaud & E. Gilson, personal communication).

TRF1 is a homodimer in solution and requires two Myb motifs to bind to telomeric DNA in vitro and in vivo. With the yeast two-hybrid assay, the homodimerization domain of TRF1 had been mapped to the region from position 66 to 263 (see Fig. 4a, ref. 6). As TRF2 displayed considerable sequence similarity to the dimerization domain of TRF1 (Fig. 1c), we wanted to determine whether the putative dimerization domain of TRF2 can interact with itself in the yeast two-hybrid assay. Interaction of LexA and GAD fusion proteins bearing the TRF2 dimerization domain resulted in readily detectable β-galactosidase activity in the same range observed for TRF1–TRF1 interaction (Fig. 4a), suggesting that TRF2 can form homodimers with a similar architecture to that of TRF1. The similar migration rate of TRF1 and TRF2ΔN-49 complexes in gel-shift assays (Fig. 2b) suggests that both proteins bind to DNA as a dimer of approximately the same mass. Furthermore, addition of FLAG antibody to TRF2ΔN-49 gel shifts resulted in two supershifted complexes (Fig. 2b)—in agreement with the presence of two FLAG epitopes in each DNA-bound TRF2 complex.

We next used the yeast two-hybrid assay to determine whether the dimerization domains of TRF1 and TRF2 can interact. In a large number of trials, we failed to detect convincing interactions between these proteins in two settings (Fig. 4a). No activation occurred when the dimerization domain of TRF2 was probed for interactions with full-length TRF1. In addition, no significant interactions were observed when the dimerization domain of TRF2 was probed for interactions with TRF2ΔN-49.
observed between fusion proteins containing the isolated TRF1 and TRF2 dimerization domains. The weak β-galactosidase activity detected with these proteins is probably due to the inherent slight transactivation activity of the TRF1 dimerization domain (Fig. 4a), which has been noted previously.6

These data indicated that whereas TRF2 can interact with itself, the TRF1–TRF2 interaction is either absent or much weaker. Absence of heterodimer formation was further corroborated by co-immunoprecipitation of in vitro translation products (Fig. 4b). In these experiments, FLAG–TRF1 could be immunoprecipitated together with GFP–TRF1, and FLAG–Trf2 was found to associate with an untagged version of Trf2, consistent with homodimerization by both proteins. However, FLAG–mTrf1 failed to form a stable association with Trf2 (Fig. 4b). The lack of heterodimerization is also consistent with the gel-shift assays on extracts from cells over-expressing Trf2ΔN-49 (Fig. 2b). In these experiments, expression of an excess of Trf2ΔN-49 did not reduce the presence of TRF1 homodimer in the extract, indicating that Trf2ΔN-49 did not form heterodimers with TRF1. The unaltered level of the TRF1 gel-shift complex was most clearly demonstrated in reactions in which the Trf2ΔN-49 complex was supershifted with the FLAG antibody (Fig. 2b, lane 4). In similar experiments executed with FLAG-tagged TRF1, heterodimerization between the endogenous TRF1 and the transfected protein is consistently detected (B. van Steensel and T.d.L., unpublished observations).

These results demonstrate that human telomeres contain two distinct, distantly related telomeric DNA-binding proteins, TRF1 and TRF2, which have several features in common. They both carry a C-terminal Myb-type DNA binding domain and can form homodimers through interactions in an N-terminal dimerization domain. Both proteins show specificity for the duplex TTAGGG repeats typical of vertebrate telomeres, and they have the same expression pattern. However, TRF1 and TRF2 are clearly distinguishable, based on the difference in the amino-acid composition of their N terminus and their rates of evolution, suggesting that these two factors may have distinct functions. As TRF1 has been shown to be a negative regulator of telomere maintenance, it should be of interest to examine the role of TRF2 in this process.

Unlike the situation in human cells, in the yeast Saccharomyces cerevisiae, the duplex part of the telomere is bound to a single DNA-binding protein, Rap1p. Although Rap1p is structurally and functionally similar to TRF1 (refs 5, 7, 12–14), these two proteins show little or no amino acid sequence similarity. Recent data suggest that Rap1p bound to telomeres exists in two alternative states, one associated with the Sir proteins and another complexed with Rap1 proteins (D. Shore, personal communication). The two forms of Rap1p are proposed to have different functions at telomeres. The presence of two distinct telomeric proteins on human telomeres could similarly reflect two different tasks for duplex telomeric binding factors.
Methods
Cloning and sequence analysis. The sequence of the 3′ UTR of a TRF1-related EST (GenBank T58911) was used in a nested PCR strategy to clone a human TRF2 cDNA fragment from a HeLa cDNA library (Stratagene). Additional TRF2 cDNAs were isolated by hybridization screening from a Namalwa 3 and a breast-cancer cell line 16 library (gifts of R.G. Roeder, RU, and C. Blobel, Memorial Sloan Kettering Cancer Center, respectively). The latter library yielded a cDNA that carried the complete open reading frame (designated mTRF2-16.1). This cDNA contains 1,282 bp of 3′ UTR but appears to lack a canonical poly(A) addition sequence. The cDNA derived from HeLa cells and the mTRF2-16.1 cDNAs differed at amino-acid position 433, where the HeLa cDNA encoded two alanine residues while the 16.1 cDNA encoded only a single alanine (Fig. 1c). A mouse cDNA (designated kTRF2-26) carrying the full open reading frame of TRF2 was isolated from a mouse brain cDNA library (Stratagene) by hybridization to the human TRF2 cDNA. Sequences were determined on both strands by means of duplex templates and have been deposited in the GenBank database. A recent comparison to the databases indicated that the TRF2 proteins are not homologous to previously identified proteins other than TRF1 and other Myb-related factors. Alignment of the TRF1 and TRF2 sequences was achieved with Clustal 1.5, with a gap opening penalty of 20 and a gap extension penalty of 0.1, and the results were displayed with SeqVu 1.01.

Expression studies. TRF2 containing a N-terminal FLAG epitope tag was constructed by cloning of an NotI–Apal fragment from Trf2-26 representing the entire open reading frame into a modified pRC/CMV expression vector (Invitrogen) carrying the FLAG epitope 5′ of the cloning site. The resulting construct contained ten amino acids derived from the pBluescript polylinker and an additional 22 amino acids from the Trf2 5′ UTR as well as the Trf2 start codon. FLAG-tagged Trf2AN49 was constructed by PCR amplification of the region between amino acid 49 and the stop codon and cloning of the purified PCR product into pRC/CMV. HeLa cells were transfected with the constructs by electroporation and processed for immunofluorescence of metaphase spreads or for simultaneous detection of telomeric DNA by FISH and TRF2 by immunofluorescence as described previously.27,28 FLAG-tagged protein was detected with monoclonal antibody M2 (Kodak), followed by FITC-labelled donkey anti-mouse. For FISH, digoxigenin-labelled [CCCUAA]37 DNA was detected with a sheep anti-digoxigenin antibody (Boehringer) and a TRITC-conjugated donkey anti-sheep IgG. Whole-cell extracts were prepared from transfected cells in a buffer containing 400 mM KCl and 0.2% NP-40 as described7 and used under standard conditions2,7 in electrophoretic mobility shift assays for duplex [TATTGGG]12 binding activity.6

Yeast 2-hybrid analysis. Constructs for 2-hybrid analysis11 containing the region between amino acids 45 and 246, representing the putative dimerization domain of TRF2, were built by PCR amplification and cloning into vectors pBTM116 (ref. 17) and pACT2 (Clontech) to create GAL and LexA fusions, respectively. The other constructs used as well as the procedures followed for yeast transformation and analysis of β-galactosidase activity were described previously.4

Co-immunoprecipitation experiments. The human and mouse TRF2 constructs were used previously described2,4,6. The FLAG-tagged Trf2 and untagged Trf2 were translated from cDNAs described above.35S-methionine–labelled proteins2,24,26 were immunoprecipitated with the M2 FLAG antibody in buffer D (20 mM Heps-KOH pH 7.9, 100 mM KCl, 20% glycerol, 0.2 mM EDTA, 0.2 mM EGTA, 0.1% NP40, 0.1% Triton X-100, 0.5 mM DTT, 0.5 mM PMSF), followed by four washes in buffer D with detergents and two washes in buffer D without detergent. Pellets were suspended in SDS/PAGE loading buffer and resolved on a 9% SDS-polyacrylamide gel.

GenBank accession numbers. T58911 for cDNA fragment encoding TRF1-related Myb motif; for sequences in Fig. 1: TRF1, U40605; TRF1-AS (shorter form), AF003001; Trf1, U65586; TRF2, AF002999; Trf2, AF003000.

Acknowledgements
We thank R. Darnell, R. Roeder and C. Blobel for generous gifts of cDNA libraries. We thank D. Shure, E. Gilson and the late T. Bialaid for communication of unpublished results. This work was supported by grants from the NIH and the Rita Allen Foundation. D.B. is supported by a Merck fellowship. A.S. is supported by NIH grant ST32GM07739.

Received 3 March; accepted 1 July 1997.