MOLECULAR AND CELLULAR BIOLOGY, Nov. 1992, p. 4834–4843 Vol. 12, No. 11
0270-7306/92/114834-10$02.00/0

A Mammalian Factor That Binds Telomeric TTAGGG Repeats In Vitro

ZHONG ZHONG,[1] LILY SHIUE,[2] SHAWN KAPLAN,[1] AND TITIA DE LANGE[1]*

The Rockefeller University, New York, New York 10021,[1] and Department of Microbiology and Immunology, University of California, San Francisco, California 94143[2]

Received 12 June 1992/Returned for modification 21 July 1992/Accepted 5 August 1992

We have identified a DNA-binding activity with specificity for the TTAGGG repeat arrays found at mammalian telomeres. This factor, called TTAGGG repeat factor (TRF), is present in nuclear extracts of human, mouse, and monkey cells. TRF from HeLa cells was characterized in detail by electrophoretic mobility shift assays. It binds double-stranded TTAGGG repeats in linear and circular DNAs. Single-stranded repeats are not recognized. The optimal site for TRF appears to contain more than six contiguous TTAGGG repeats. Tandem arrays of TAGGG, TTTAGGG, TTTTAGGG, TTGGGG, and TTAGGC repeats do not bind TRF well, indicating that TRF preferentially recognizes the telomeric repeat sequence present at mammalian chromosome ends. The apparent molecular mass of this factor, based on recovery of TRF from sodium dodecyl sulfate-polyacrylamide gels, is approximately 50 kDa. We suggest that TRF binds along the length of mammalian telomeres.

The ends of linear eukaryotic chromosomes are protected from degradation and fusion by a specialized terminal structure called the telomere (see references 7 and 72 for reviews). Telomeres typically contain an array of short (5- to 8-bp) repeats, which are usually G rich in the strand that extends to the 3′ end of the chromosome. Although details of telomeric DNA are known in many systems, it is not clear how telomeres stabilize natural chromosome ends.

Loss of telomeric DNA by degradation and incomplete replication may be balanced by telomerase. Telomerase has an unusual reverse transcriptase activity, which adds single-stranded telomeric repeats to 3′ DNA ends by using an internal RNA template (26–28, 44, 59, 70, 71; see reference 6 for a review). Although telomere synthesis by telomerase may be important to counteract deletion of telomeric sequences, this activity does not explain other telomere attributes such as resistance to ligation and recombination.

The protective activity of telomeres could depend in part on structural features of telomeric repeats. It is likely that all G-rich telomeric repeats can adopt a stable folded configuration in vitro by G-G base pairing (32, 35, 55, 60, 64, 68), and single-stranded G-rich protrusions are a conserved feature of ciliate macronuclear telomeres (31, 37, 48). Whether telomeric foldback structures actually protect DNA ends and, indeed, whether they occur in vivo are not known.

Several lines of evidence indicate that telomeric DNA is complexed to nonhistone proteins. The abundant macronuclear telomeres of hypotrichous ciliates are encapsulated by a protein complex that recognizes the sequence and structure of these termini (20, 23–25, 49–51). In *Saccharomyces cerevisiae*, multiple binding sites for the repressor/activator protein RAP1 occur within each telomere (5, 10, 40). Genetic and biochemical evidence suggests that RAP1 interacts with telomeres in vivo (12, 41, 65, 69). The chromatin structure of telomeres in *S. cerevisiae* and *Tetrahymena* species bears evidence of a large DNA-protein complex (8, 69). Finally, human telomeric DNA is attached to the nuclear matrix, presumably by interacting with a nonhistone factor (15). On the basis of these observations, telomeres may be viewed as unique nucleoprotein complexes.

The structural conservation of telomeric DNA and the fact that telomere healing in *S. cerevisiae* occurs on DNA ends with heterologous telomeric repeats (see references 7 and 72 for reviews) suggest that telomere function is highly conserved. According to this view, essential telomere-protein interactions are expected to be conserved as well. For example, the demonstration of telomerase in several ciliates as well as in mammalian cells stressed its fundamental role in telomere replication (24, 44, 58, 71). It is not yet known whether other telomere factors are similarly conserved in widely diverged eukaryotes. To address this question, we have undertaken a search for telomere binding factors in mammalian cells.

Mammalian telomeres contain the repeat sequence TTAGGG in long tandem arrays which vary from 2–30 kb at human chromosome ends to as much as 100 kb at the ends of mouse chromosomes (3, 16, 29, 30, 36, 42, 45, 63). By electrophoretic mobility shift assays (EMSA) we have identified a conserved mammalian DNA-binding factor with in vitro specificity for double-stranded TTAGGG repeats. According to our biochemical characterization, this candidate telomere protein could bind along the length of mammalian telomeres.

MATERIALS AND METHODS

Cell lines. HeLa suspension cells were grown in spinner flasks in Joklik's medium–5% iron-supplemented bovine calf serum (Hyclone). Monkey COS-7 (kidney) cells were grown in roller bottles in Dulbecco modified Eagle medium–10% fetal calf serum (Hyclone). Mouse J55S (plasmacytoma) cells were grown in roller bottles in Joklik's medium–10% bovine calf serum. Human Namalwa (B) cells were grown in roller bottles in RPMI 1640–10% fetal calf serum. The media were supplemented with antibiotics, nonessential amino acids, glutamine, and 10 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid (HEPES)–KOH (pH 7.4). Cells were harvested at $<2 \times 10^5$ cells per ml.

Nuclear extracts. Nuclear extracts were prepared at 4°C by

* Corresponding author.

TABLE 1. Structures of DNA probes[a]

3merA	CTGAAGCTTGCATGCCTGCAGGTCGACTCTAGAGATCGAATTCGAGCCC**GGGTTAGGGTTAGGGTTAGGG**CCCCTCTAGA GTCGACGATCCCGGGTACCGAGCTCGAATTCATCG
6merA	GGTCGACTCTAGAGATCGAATTCGAGCCC**GGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGG**CCCTCTAGAGT CGACGATCCCGGGTACCGAGCTCGAATTCATCG
12merA	TCGACTCTAGAGATCGAATTCGAGCCC**GGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTT** **AGGGTTAGGGTTAGGGTTAGGG**CCCCTCTAGAGTCGA
3merB	GATCTGATATCATCGATGAATTCGAGCTC**GGTTAGGGTTAGGGTTAGGG**ATCCTCTAGAGTCGA
6merB	CGATGAATTCGAGCT**GGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGG**ATC
6merB*	GATCTGATATCATCGATGAATTCGAGCTC**GGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGG**ATCCTCTAGATCG A
12merB	CGATGAATTCGAGCTC**GGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTT** **AGGG**CTA
12merB*	GATCTGATATCATCGATGAATTCGAGCTC**GGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGT** **TAGGGTTAGGGTTAGGGTTAGGG**ATCCTCTAGAGTCGA
3merC	TCGAGCAGCTGAAGCTTGCATGCCTGCAGGTCGACTCTAGAGATCGAATTCGAGCCC**GGGTTAGGGTTAGGGTTAGGG**CC CTCTAGAGTCGACGATCCCGGGTACCGAGCTCGAATTCATCGATGATATCTGACT
3+3mer	TCGAGCAGGAATTCGAGCCC**GGGTTAGGGTTAGGGTTAGGG**CCCCTCTAGAGTCGACCTGAAGCTTGCATGCCTGCAGGT CGACTCTAGAGATCGAATTCGAGCCC**GGGTTAGGGTTAGGGTTAGGG**CCCCTCTAGAGTCGACGATCCCGGGTAC
3+inv3mer	TCGAGCAGGTCGACTCTAGAGGGG**CCCTAACCCTAACCCTAACCC**GGGCTCGAATTCCTGAAGCTTGCATGCCTGCAGGT CGACTCTAGAGATCGAATTCGAGCCC**GGGTTAGGGTTAGGGTTAGGG**CCCCTCTAGAGTCGACGATCCCGGGTAC
6merC	TCGAGCAGCTGAAGCTTGCATGCCTGCAGGTCGACTCTAGAGATCGAATTCGAGCCC**GGGTTAGGGTTAGGGTTAGGGTT** **AGGGTTAGGGTTAGGG**CCCCTCTAGAGTCGACGATCCCGGGTACCGAGCTCGAATTCATCGATGATATCTGATC

[a] Sequences are given from 5′ to 3′. TTAGGG repeats are in boldface type. All probes are double stranded restriction fragments with blunt ends.

extraction of nuclei with 0.4 M KCl by the method of Dignam et al. (17). Typically, 10^{10} HeLa cells yielded 60 ml of nuclear extract with 6 mg of protein per ml according to the Bradford assay (Bio-Rad). Extracts were dialyzed against 100 mM KCl in buffer D (20 mM HEPES–KOH [pH 7.9], 1 mM β-mercaptoethanol, 1 mM phenylmethylsulfonyl fluoride, 20% glycerol) and stored at −80°C. For the experiment whose results are shown in Fig. 1B, HeLa nuclei were extracted with 0.4, 0.6, 0.8, or 1.0 M KCl. These extracts were assayed for TTAGGG repeat factor (TRF) activity on a per-nucleus basis.

Cell elutriation, nocodazole treatment, and preparation of whole-cell extracts. The whole-cell extracts from HeLa cells in different stages of the cell cycle were generously provided by Neil Segil, The Rockefeller University. All procedures are described by Segil et al. (54).

DNA probes and competitors. EMSA DNA probes are described in Table 1. These sequences are synthetic DNAs cloned into the polylinker of pSP73 (Promega). All assays were performed with gel-purified, labeled restriction fragments from these plasmids. The fragments were isolated from 5% polyacrylamide gels after end labeling with Klenow enzyme and [α-^{32}P]dCTP. Competition experiments were performed with either synthetic oligonucleotides or covalently closed circular plasmids. The structure of the oligonucleotide competitors is as follows:

ds(TTAGGG)$_7$ 5′-(TTAGGG)$_7$CATCGATGGG-3′
3′-CATGCTACCC(AATCCC)$_7$-5′

ds(TTGGGG)$_7$ 5′-(TTGGGG)$_7$CATCGATGGG-3′
3′-CATGCTACCC(AACCCC)$_7$-5′

The oligonucleotides ss(TTAGGG)$_7$ and ss(CCCTAA)$_7$ are the top and bottom strand of ds(TTAGGG)$_7$, respectively. Synthetic oligonucleotides were purified by denaturing (urea) polyacrylamide gel electrophoresis and annealed by incubation at 65°C for 2 h.

In the construction of plasmids with telomeric repeat arrays, the following starting materials were used. The TTAGGG repeat arrays of 25 and 135 units were derived from clones isolated from human libraries (15a, 16). The *Ascaris* (TTAGGC)$_{25}$ repeat array is the *Bam*HI-*Sac*I fragment from plasmid pTElS5 (46), which also contains a region of 150 bp of subtelomeric DNA. The *Chlamydomonas* (TTTTAGGG)$_{35}$ repeat array is derived from pCT17 (47) as a *Hin*cII-*Xba*I fragment that also contains 250-bp subtelomeric sequences. The *Arabidopsis* (TTTAGGG)$_{55}$ repeat array is derived from pAtT4 (53) as a *Not*I-*Sac*I fragment. The *Tetrahymena* (TTGGGG)$_{\sim 85}$ repeat array is derived from p45 (56) as a *Bam*HI fragment. The *Giardia* (TAGGG)$_{\sim 110}$ repeat array is derived from pGC1ΔN (38) as a *Sma*I-*Eco*RV fragment. The telomeric repeat arrays were inserted in the polylinker of pSP73 (Promega) next to a *Sal*I-*Xho*I fragment from pMC1neopolA (Stratagene), which contains the neomycin resistance gene. All telomeric repeat arrays were inserted in the same orientation, and the plasmids were sequenced to verify the composition of the telomeric repeat array. The name of each plasmid refers to the sequence of the telomeric repeats it carries preceded by the approximate number of repeats in the insert. For instance, pSXneo85 TTGGGG contains an array of TTGGGG repeats of approximately 500 bp. Plasmid DNAs were isolated from CsCl-ethidium bromide gradients. The nonspecific competitor DNA was chromosomal DNA from *Escherichia coli* cleaved with *Hae*III or sheared by sonication and purified by extraction with phenol and precipitation with ethanol.

Electrophoretic mobility shift assays. The optimal assay conditions for TRF were determined by varying the probe concentrations, KCl and Mg^{2+} concentrations, pH, temperature, incubation time, competitor DNA, and gel conditions

(data not shown). The following EMSA conditions were used throughout these studies. Labeled DNA probe (1 ng and ~20 cps; representing at least a 10-fold excess to the binding capacity of TRF) was incubated with crude nuclear extract (typically 1 μl with ~6 μg of protein) or partially purified TRF (1 μl with ~2.5 μg of protein) in 20 μl of EMSA buffer (20 mM HEPES-KOH [pH 7.9], 150 mM KCl, 1 mM $MgCl_2$, 0.1 mM EDTA, 0.5 mM dithiothreitol, 5% glycerol, 4% Ficoll [molecular weight 400,000]) with 1 to 3 μg of *Hae*III-cleaved *E. coli* chromosomal DNA. After incubation for 30 min at 22°C, the mixture was fractionated on a 5% acrylamide gel (29:1 acrylamide:bisacrylamide) run in 1× TBE (89 mM Tris base, 89 mM boric acid, 2 mM EDTA) at 130 V for 2.5 h at 22°C. Gels were dried under vacuum at 80°C and autoradiographed for 16 h at 22°C.

Estimation of TRF abundance. To estimate the abundance of TRF in HeLa cells, we quantitated EMSA results by using a PhophorImage (Molecular Dynamics). In these experiments variable amounts of crude HeLa nuclear extract were allowed to bind an excess of the *12merA* probe (~6 × 10^9 molecules per lane). The percentage of probe bound by TRF was measured and corrected for background values derived from competition experiments. The abundance of TRF per cell was then estimated from the number of bound *12merA* molecules and the number of nuclei used to derive the extract (~1.6 × 10^6/μl).

Partial purification of TRF. HeLa nuclear extract was passed through DEAE-Sepharose (Sigma) in buffer D–100 mM KCl. The flowthrough and first wash fractions contained TRF and were fractionated on phosphocellulose (P11; Whatman) in buffer D containing stepwise-increased KCl concentrations. TRF activity elutes at 0.6 M KCl. In a different protocol, crude nuclear extract was fractionated by gel filtration on Sephacryl S-300 (Pharmacia). The void fractions containing TRF activity were applied to a column with ds(TTAGGG)$_7$ DNA coupled to CNBr-activated Sepharose CL4B (Pharmacia). This column was run in buffer D with stepwise-increased concentrations of KCl. TRF elutes at 0.5 M KCl. Partially purified TRF was dialyzed against buffer D–100 mM KCl and stored at −80°C.

Recovery of TRF from SDS-PAGE. The procedure for TRF recovery was essentially as described by Baeuerle and Baltimore (4). Partially purified TRF (200 μl of P11 0.6 M KCl fraction containing 500 μg of protein) was fractionated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (9% acrylamide) until the bromophenol blue dye had migrated 20 cm. An initial experiment established that TRF migrates in the 40- to 60-kDa range. In the experiment shown in Fig. 6, gel slices (0.5 cm) in the range from 40 to 65 kDa were crushed and soaked overnight in 0.5 ml of 50 mM Tris-HCl (pH 7.9)–0.1% SDS–0.1 mg of bovine serum albumin per ml–1 mM dithiothreitol–0.2 mM EDTA–0.1 mM phenylmethylsulfonyl fluoride–2.5% glycerol. The extracted protein was precipitated with 4 volumes of acetone (at −20°C), washed with cold methanol, and dried. The pellets were dissolved in 10 μl of 8 M urea and subsequently diluted with 125 μl of 20 mM Tris-HCl (pH 7.6)–100 mM KCl–2 mM dithiothreitol–10 mM phenylmethylsulfonyl fluoride. After 18 h of renaturation at 4°C, 8 μl of renatured fraction was mixed with 10 μl of 2× concentrated EMSA buffer and incubated on ice for 2.5 h. Subsequently 1 μg of sonicated *E. coli* DNA and 1 ng of probe were added. The mixture (final volume, 20 μl) was incubated for 30 min at 22°C and assayed on a gel as described above. The molecular weights of the isolated fractions were derived from a calibration curve obtained with Bio-Rad molecular weight standards (low- and high-molecular-weight range).

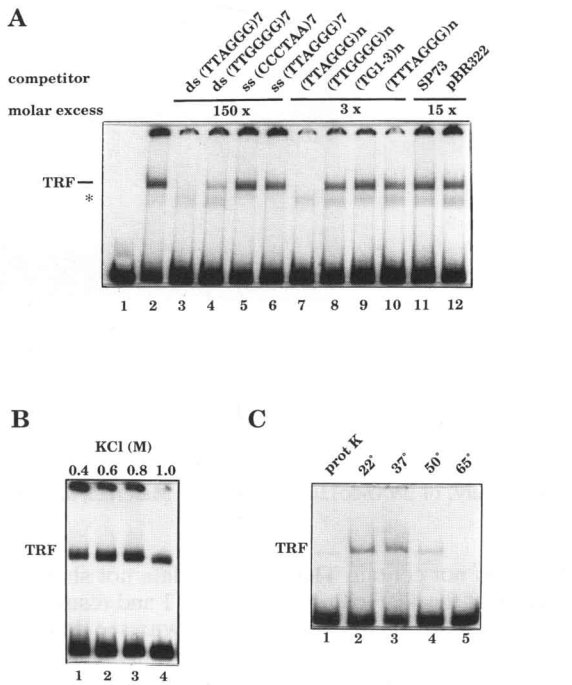

FIG. 1. Identification of a TRF in HeLa nuclear extract. (A) EMSA of crude HeLa nuclear extract in the presence of various competitor DNAs with *12merA* as a probe. See Table 1 for the structure of *12merA* and Materials and Methods for the composition of competitor DNAs. Lanes: 2, no competitor; 3, 150-fold molar excess of ds(TTAGGG)$_7$; 4, 150-fold molar excess of ds(TTGGGG)$_7$; 5, 150-fold molar excess of ss(CCCTAA)$_7$; 6, 150-fold molar excess of ss(TTAGGG)$_7$; 7, 3-fold molar excess of pSXneo135TTAGGG; 8, 3-fold molar excess of pSXneo85TTGGGG; 9, 3-fold molar excess of pYt103 (containing 81 bp of TG_{1-3} repeats [57]); 10, 3-fold molar excess of pSXneo55TTTAGGG; 11, 15-fold molar excess of pSP73; 12, 15-fold molar excess of pBR322. Each reaction contained 3 μg of *E. coli* chromosomal DNA cleaved with *Hae*III. Lane 1 did not contain extract. The position of the TRF complex is indicated; the asterisk indicates a nonspecific binding activity. (B) TRF activity in four nuclear extracts made in the presence of increasing KCl concentrations. Identical numbers of HeLa nuclei were extracted in parallel with 0.4 M (lane 1), 0.6 M (lane 2), 0.8 M (lane 3), and 1.0 M (lane 4) KCl and assayed for TRF activity on *12merA* probe after dialysis of each extract to 100 mM KCl. (C) Effect of proteinase K treatment and incubation at elevated temperatures on TRF. In lane 1, HeLa extract was digested with proteinase K at 50 μg/ml for 30 min at 37°C prior to the gel shift assay. In the other lanes, HeLa nuclear extract was incubated for 30 min at 22°C (lane 2), 37°C (lane 3), 50°C (lane 4), and 65°C (lane 5) immediately prior to the mobility shift assay.

RESULTS

Identification of TRF. We used a cloned DNA restriction fragment containing 12 tandemly arranged TTAGGG repeats (*12merA* [Table 1]) as the probe in EMSA to search for a sequence-specific binding activity in extracts of HeLa cell nuclei. In the experiment in Fig. 1A, this probe is bound by two factors, TRF and a nonspecific DNA-binding factor. TRF is extracted from HeLa cell nuclei with 0.4 M KCl; extraction with 0.6, 0.8, or 1 M KCl does not release additional TRF activity (Fig. 1B). Cytoplasmic (S100) ex-

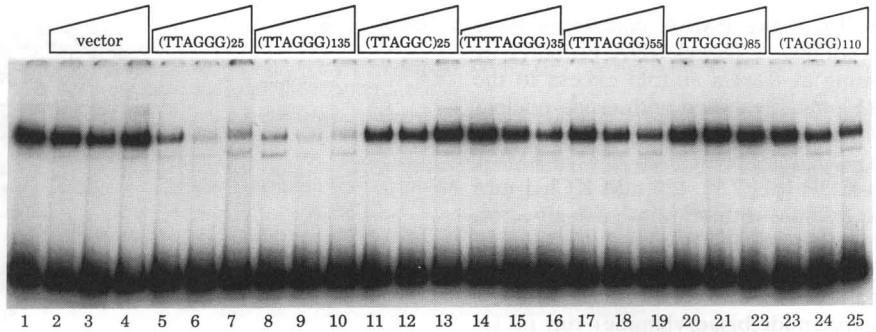

FIG. 2. Sequence specificity of TRF. Partially purified TRF (P11 0.6 M KCl fraction [see Materials and Methods]) was allowed to bind *12merA* in the presence of plasmid DNAs with homologous or heterologous telomeric repeat arrays. Lanes: 1, no extract; 2 to 4, 3×, 9×, and 27× molar excess, respectively, of pSXneo; 5 to 7, 3×, 9×, and 27× molar excess, respectively, of pSXneo25TTAGGG; 8 to 10, 3×, 9×, and 27× molar excess, respectively, of pSXneo135TTAGGG; 11 to 13, 3×, 9×, and 27× molar excess, respectively, of pSXneo25TTAGGC; 14 to 16, 3×, 9×, and 27× molar excess, respectively, of pSXneo35TTTTAGGG; 17 to 19, 3×, 9×, and 27× molar excess, respectively, of pSXneo55TTTAGGG; 20 to 22, 3×, 9×, and 27× molar excess, respectively, of pSXneo85TTGGGG; 23 to 25, 3×, 9×, and 27× molar excess, respectively, of pSXneo110TAGGG. See Materials and Methods for the structures of the plasmids used in these experiments.

tracts do not contain TRF activity (data not shown). Quantitation of the TRF complexes in Fig. 1 and results of extract titration experiments (data not shown; see Materials and Methods) suggest that HeLa cells contain at least 1,500 copies of TRF per nucleus. In the experiment in Fig. 1C, TRF is inactivated by incubation at 65°C or by treatment with proteinase K, indicating that the activity depends on a proteinaceous component.

The competition experiments in Fig. 1A identify TRF as a double-stranded TTAGGG repeat binding factor. Synthetic single-stranded telomeric repeat DNAs do not compete for TRF at a 150-fold molar excess (Fig. 1A). By contrast, the TRF complex, but not the nonspecific complex, is competed for with a 150-fold molar excess of unlabeled synthetic DNA with seven dsTTAGGG repeats [ds(TTAGGG)$_7$; see Materials and Methods]. Similarly, a circular plasmid with an array of 135 TTAGGG repeats competes for TRF at a threefold molar ratio of plasmid to probe (equivalent to a ~30-fold molar ratio of TTAGGG repeats). The effective competition of TTAGGG repeats in a circular context indicates that TRF does not require a DNA end for binding. In addition, the nucleotides adjacent to the TTAGGG repeats in the *12merA* probe are probably not recognized by TRF, since these sequences are not found in the competitor DNAs.

Circular plasmids carrying heterologous telomeric repeats from *Tetrahymena* spp. (TTGGGG), *S. cerevisiae* (TG$_{1-3}$), and *Arabidopsis* spp. (TTTAGGG) have little effect on TRF complex formation, and a synthetic ds(TTGGGG)$_7$ array is a less effective competitor than ds(TTAGGG)$_7$ (Fig. 1A). This initial characterization indicates that TRF preferentially binds TTAGGG repeat arrays over heterologous telomeric DNA; a more extensive analysis of the sequence specificity of TRF is presented below.

Heterologous telomeric repeats do not bind TRF. To determine the sequence specificity of TRF, we performed a series of competition experiments with circular plasmids carrying vertebrate and heterologous telomeric repeats. The vector backbone of all constructs is identical and does not affect TRF complex formation (Fig. 2). To establish the appropriate assay conditions, we allowed TRF to bind to an excess of *12merA* probe in the presence of a wide concentration range of plasmids carrying either 25 or 135 TTAGGG repeats (data not shown). These initial experiments indicated that both TTAGGG repeat plasmids compete out all TRF binding to *12merA* at a ninefold molar ratio of plasmid to probe (Fig. 2). The residual band detected under these competition conditions is attributed to a minor nonspecific binding activity in this TRF fraction. At a 3-fold molar excess, which is equivalent to a 6- and 30-fold molar excess of repeats for the (TTAGGG)$_{25}$ and (TTAGGG)$_{135}$ plasmids, respectively, the plasmids decrease the level of TRF-*12merA* complex approximately 10-fold (Fig. 2). The effect of plasmids with heterologous telomeric repeat arrays was therefore similarly examined at 3-, 9-, and 27-fold molar excess. Although the lengths of the heterologous telomeric repeat stretches vary, the range is covered by the plasmids carrying 25 and 135 TTAGGG repeats that serve as a positive control in this experiment. As shown in Fig. 2, plasmids carrying telomeric DNA inserts from *Ascaris* spp. [(TTAGGC)$_{25}$ (46)], *Chlamydomonas* spp. [(TTTTAGGG)$_{35}$ (47)], higher plants [(TTAGGG)$_{55}$ (53)], *Tetrahymena* spp. [(TTGGGG)$_{-85}$ (56)], and *Giardia* spp. [(TAGGG)$_{-110}$ (38)] have little effect on TRF complex formation. At the highest molar ratio (27-fold), TTTTAGGG, TTTAGGG, and TAGGG repeat arrays compete slightly (Fig. 2).

Since TRF does not bind TTGGGG and TTAGGC repeats effectively, it seems that TRF recognizes TTAGGG repeats in part by interactions with nucleotides at positions 3 and 6 (with the first T of TTAGGG arbitrarily chosen as position 1). Furthermore, the competition data indicate that TRF recognizes additional features of TTAGGG repeats since changes in the number of T residues dramatically reduce complex formation.

Optimal TRF binding requires at least six TTAGGG repeats. Although the data presented in Fig. 2 establish that TRF binds the *12merA* probe in a sequence-specific manner, they do not reveal the length of the site that is recognized by TRF. To address this question, we generated a family of EMSA probes with variable numbers of TTAGGG repeats (Table 1). As shown in Fig. 3A, three probes with 3, 6, or 12 complete TTAGGG repeats (*3merA*, *6merA*, and *12merA*, respectively) were used to determine the apparent affinity of TRF for TTAGGG repeat arrays of different lengths. The result shows that TRF binds *12merA* better than *6merA*, while the *3merA* probe binds very little TRF. A minor amount of a specific complex migrating more slowly than the nonspecific complex can be discerned in long exposures, but the binding of TRF to the 21-bp telomeric repeat region in *3merA* is clearly suboptimal.

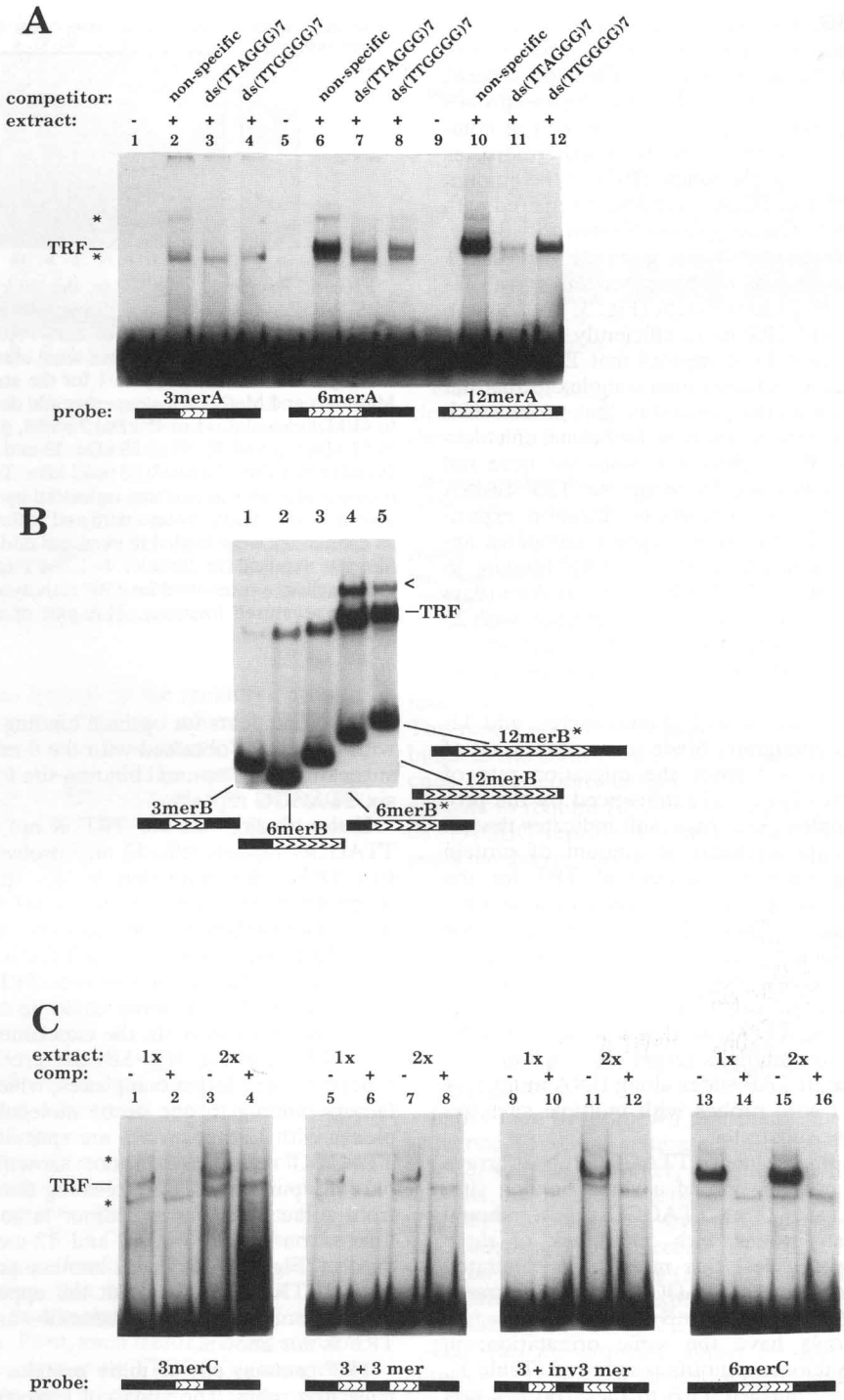

FIG. 3. TRF binds preferentially to long contiguous TTAGGG repeat arrays. (A) Mobility shift assays were performed with crude HeLa nuclear extract and three identically sized probes with 3 (*3merA*; lanes 1 to 4), 6 (*6merA*; lanes 5 to 8), or 12 (*12merA*; lanes 9 to 12) TTAGGG repeats (see Table 1 for structures of the probes). Lanes 1, 5, and 9 do not contain nuclear extract. In lanes 3, 7, and 11, TRF is competed for with a 150-fold molar excess of ds(TTAGGG)$_7$ (see Materials and Methods). In lanes 4, 8, and 12, TRF is competed for with a 150-fold molar excess of ds(TTGGGG)$_7$ (see Materials and Methods). The position of the TRF complex is indicated; the asterisks indicate a nonspecific DNA-binding activity (lower band) and a larger complex of unknown composition. (B) Partially purified HeLa TRF (by gel filtration and specific DNA affinity chromatography [see Materials and Methods]) was incubated with *3merB* (lane 1), *6merB* (lane 2), *6merB** (lane 3), *12merB* (lane 4), and *12merB** (lane 5). The structure of the probes is given in Table 1. Next to lane 5, the position of the TRF complex in this lane is indicated; the caret points to larger complexes that may represent multimeric TRF binding (see text). (C) Partially purified TRF (0.6 M KCl fraction from phosphocellulose [see Materials and Methods]) was incubated with *3merC* (lanes 1 to 4), *3+3mer* (lanes 5 to 8), *3+inv3mer* (lanes 9 to 12), and *6merC* (lanes 13 to 16). Either 1.5 or 3.0 μl of TRF was added to the incubations, as indicated above the lanes. In the even lanes, TRF is competed for with a fourfold molar excess of pSXneo135TTAGGG (see Materials and Methods). The asterisks are defined in the legend to panel A.

As shown in Fig. 3B, a similar result was obtained with partially purified factor and a second set of probes with variable numbers of telomeric repeats (*3merB*, *6merB*, *6merB**, *12merB*, and *12merB** [Table 1]). Again, probes with 12 TTAGGG repeats formed a larger amount of complex than did the probes with 6 repeats. Furthermore, as seen in Fig. 3A, the 3-mer probe binds TRF but the amount of complex formed with this DNA was minor compared with that for the other probes. The complexes formed with 3-mer, 6-mer, and 12-mer probes specifically compete with ds(TTAGGG)$_7$, and the heterologous telomeric repeats in ds(TTGGGG)$_7$ DNA compete to a lesser extent (Fig. 3; not shown). Since 12-mer probes bind TRF more efficiently than probes with six TTAGGG repeats do, it appears that TRF requires at least six TTAGGG repeats for optimal complex formation.

Our data do not exclude the possibility that probes with more than 12 TTAGGG repeats show an additional enhancement of TRF binding. For technical reasons we have not been able to address this issue by using our TRF EMSA conditions. We note that in competition titration experiments, a plasmid with 25 TTAGGG repeats competes approximately fourfold more efficiently for TRF binding to *12merA* than a plasmid with 12 TTAGGG repeats does (data not shown). Whether the more efficient competition with 25 TTAGGG repeats is due to presentation of a larger site or other effects (such as the presence of multiple sites) cannot be established at present.

The TRF complexes formed with 3-mer, 6-mer, and 12-mer probes in Fig. 3A comigrate. Since these probes are of identical size (Table 1) and since the migration rate of DNA-protein complexes is strongly influenced by the protein moiety in the complex (33), this result indicates that all these complexes contain an identical amount of protein factor. Therefore, the increased avidity of TRF for the longer telomeric repeat arrays cannot be due to cooperative binding of multiple factors. Two other explanations for the enhanced binding to longer repeat arrays are conceivable. First, the TRF-binding site could span more than six TTAGGG repeats. Alternatively, the enhanced binding of TRF to probes with long TTAGGG repeat arrays could be due to the presence of multiple target sites within one molecule. For instance, if TRF slides along DNA molecules in search of its target site, probes with multiple sites are expected to be a better substrate.

If the enhanced binding to longer TTAGGG repeat arrays is due to redundant presentation of a short binding site, probes with two blocks of three TTAGGG repeats should bind TRF better than probes with one block of three TTAGGG repeats do. To test this model, we generated probes with two blocks of three TTAGGG repeats separated by 64 bp of unrelated sequence. In the *3+3mer* probe the two TTAGGG repeat arrays have the same orientation; in *3+inv3mer* one of the telomere motifs is inverted (Table 1). The binding of partially purified TRF to these probes was compared with complex formation on *3merC* and *6merC*, which are similar in length to the probes with two sets of TTAGGG repeat arrays but have contiguous arrays of three and six TTAGGG repeats, respectively (Table 1). The result in Fig. 3C shows that TRF binds the *3+3mer*, *3+inv3mer*, and *3merC* probes to a similar extent. Although the amount of complex formed on probes with one or two blocks of three repeats varies slightly within a 2-fold range, each probe binds approximately 10-fold less TRF than *6merC* does. This result indicates that the presence of multiple binding sites within one probe does not enhance TRF binding significantly. Therefore it appears that TRF requires contiguous TTAGGG repeats for optimal binding. These data, together with the results obtained with the 6-mer and 12-mer probes, suggest that the optimal binding site for TRF covers at least six TTAGGG repeats.

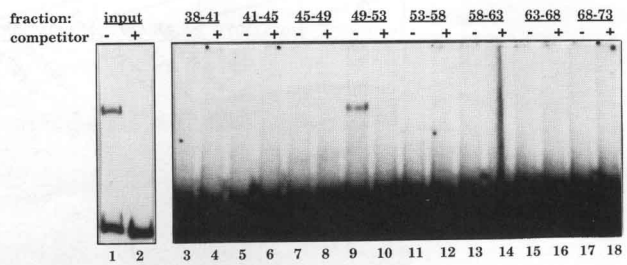

FIG. 4. Recovery of TRF in the 50-kDa range fraction from SDS-PAGE. Protein (500 μg) of partially purified TRF (P11 0.6 M KCl fraction) was subjected to SDS-PAGE, and proteins in the indicated molecular mass regions were eluted and assayed for TRF activity on *12merA* (see Table 1 for the structure of the probe and Materials and Methods for experimental details). Lanes: 3 and 4, 38 to 41 kDa; 5 and 6, 41 to 45 kDa; 7 and 8, 45 to 49 kDa; 9 and 10, 49 to 53 kDa; 11 and 12, 53 to 58 kDa; 13 and 14, 58 to 63 kDa; 15 and 16, 63 to 68 kDa; 17 and 18; 68 to 73 kDa. The molecular mass of the proteins in each fraction was estimated by comparison with molecular mass standards. Assays with and without pSXneo135TTAGGG as competitor were loaded in even and odd lanes, respectively. The film was exposed for 2 weeks. In lanes 1 and 2, the input P11 0.6 M KCl fraction was assayed for TRF activity in the presence of urea as for the renatured fractions. This part of the gel was exposed for 16 h.

If the binding site for TRF is not much larger than six TTAGGG repeats, the 12-mer probes could accommodate two TRFs. We note that in the absence of cooperative interactions, the simultaneous attachment of two factors to the 12-mer probes is expected to be infrequent. In addition, since TRF requires at least six TTAGGG repeats for optimal binding, only 12-mer probes with TRF bound at one end of the TTAGGG repeat array could be expected to accommodate a second factor. In the experiment with partially purified TRF shown in Fig. 3B, the probes with 12 TTAGGG repeats acquire larger complexes, which could represent two factors binding to one probe molecule. These larger complexes with 12-mer probes are specifically competed for by TTAGGG repeats (data not shown) and are seen with partially purified TRF, suggesting that they are not derived from some other factor. Minor larger complexes are also formed on 3-mer, 6-mer, and 12-mer probes with crude extract (Fig. 3A) and with another preparation of partially purified TRF (Fig. 3C), but the appearance of these faint bands is not uniformly reproducible, and their relationship to TRF is not known.

TRF contains one or more proteins of ~50 kDa apparent molecular mass. The apparent molecular mass of TRF was deduced by recovery of active factor after fractionation on SDS-polyacrylamide gels (4). TRF was partially purified through DEAE-Sepharose and phosphocellulose columns (see Materials and Methods) and electrophoresed through a preparative denaturing gel. Proteins eluted from different gel slices were renatured and tested for TRF activity in mobility shift assays with *12merA* as a probe. Figure 4 shows that TRF activity is recovered in a single fraction containing proteins with an apparent molecular mass of 49 to 53 kDa. The same result was obtained in a second independent fractionation experiment (data not shown). The recovered activity appears to be authentic TRF on the basis of two

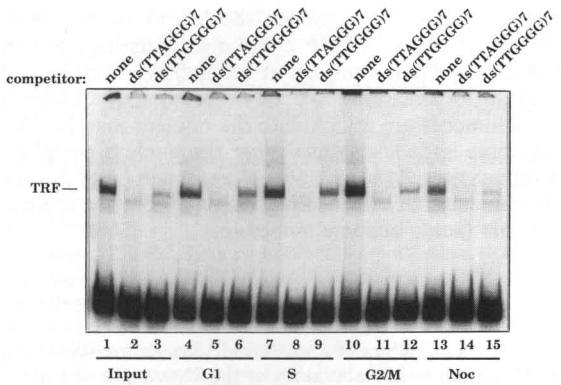

FIG. 5. TRF expression in the cell cycle. Whole-cell extracts from HeLa cells (lanes 1 to 3), HeLa cells arrested with nocodazole (Noc) (lanes 13 to 15), and elutriated HeLa fractions in G_1 (lanes 4 to 6), S (lanes 7 to 9), and G_2 (lanes 10 to 12) were assayed for TRF activity by using *12merA* (Table 1) as a probe. Incubations were run without competitor DNA (lanes 1, 4, 7, 10, and 13), with a 150-fold excess of ds(TTAGGG)$_7$ (lanes 2, 5, 8, 11, and 14), and with a 150-fold excess of ds(TTGGGG)$_7$ (lanes 3, 6, 9, 12, and 15). The amount of extract assayed for each cell population was normalized to represent equivalent amounts of cellular protein.

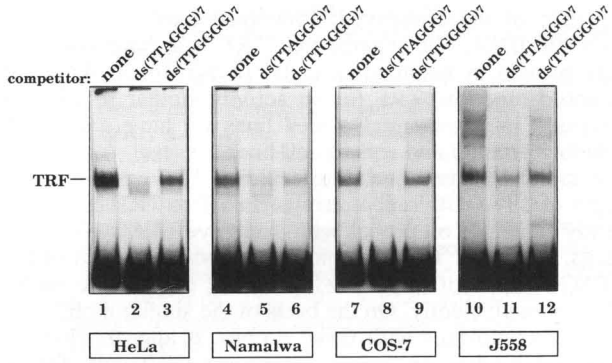

FIG. 6. TRF-like activity in different mammalian cell lines. Nuclear extracts from human HeLa (lanes 1 to 3), human Namalwa (lanes 4 to 6), monkey COS-7 (lanes 7 to 9), and mouse J558 (lanes 10 to 12) cells were assayed for TRF activity by mobility shift of the *12merA* probe. Incubations were performed without competitor DNA (lanes 1, 4, 7, and 10), with a 150-fold excess of ds(TTAGGG)$_7$ (lanes 2, 5, 8, and 11), and with a 150-fold excess of ds(TTGGGG)$_7$ (lanes 3, 6, 9, and 12). The amount of extract assayed for each cell line was normalized to represent equivalent amounts of nuclear protein.

criteria. First, the complex formed by the renatured protein comigrates with unfractionated TRF and not with the nonspecific DNA-binding factor in the extract. Second, the recovered binding activity is specifically competed for a threefold molar excess of a plasmid with 135 TTAGGG repeats. This result indicates that TRF contains one or more proteins with an apparent molecular mass of ~50 kDa.

TRF is expressed in all cell cycle stages. Whole-cell extracts were prepared from HeLa cells in G_1, S, and G_2 phases isolated by centrifugal elutriation. In addition, proteins were extracted from nocodazole-treated HeLa cells arrested in mitosis. TRF activity was detected in each of these extracts (Fig. 5). In this experiment, the activity varied somewhat with the highest level of TRF in G_2. However, in other experiments the increase in G_2 was not as prominent. We conclude that TRF is present throughout the cell cycle. Whether TRF undergoes a subtle cell cycle modulations cannot be assessed at present.

TRF is expressed in human, primate, and rodent cell lines. Extracts from four mammalian cell lines were examined for telomeric repeat binding activity. The experiment in Fig. 6 shows that monkey kidney cells (COS-7), mouse plasmacytoma cells (J558), and a human Burkitt's lymphoma cell line (Namalwa) express a factor that binds to *12merA*. This binding activity is similar or identical to that of HeLa TRF on the basis of two criteria. First, each factor is competed for by ds(TTAGGG)$_7$ and in each case ds(TTGGGG)$_7$ competes to a lesser extent, as is seen for HeLa TRF. Second, in each case the complex comigrates with the HeLa TRF complex. TRF-like activity is also expressed by a human granulocyte-macrophage precursor cell line (HL-60), human histiocytic leukemia cells, and mouse erythroleukemia cells (data not shown).

DISCUSSION

Telomeres are thought to function as a nucleoprotein complex consisting of telomeric repeats and telomere-binding factors. Although the sequence and structure of telomeric DNA has been characterized in many systems, much less is known about the protein components of telomeres. This report describes a mammalian nuclear factor, TRF, that binds to the conserved TTAGGG repeat arrays found at the ends of vertebrate chromosomes. Optimal binding of TRF appears to require DNA probes with at least six contiguous dsTTAGGG repeats. In the human genome, TTAGGG repeats hybridize to a few chromosome-internal sites as well as at telomeres. The sequences of two of these loci demonstrate that these sites are composed mostly of non-TTAGGG repeat DNA. They contain scattered TTAGGG repeats but no tandemly repeated stretches longer than 24 bp (34, 67). According to the in vitro DNA-binding characteristics of TRF, this factor would not be expected to bind to such short TTAGGG repeat arrays efficiently. Thus, in human chromosomes optimal TRF-binding sites occur predominantly at telomeres.

It has long been known that many mammals contain the sequence TTAGGG reiterated within alpha-satellite repeats (21, 22, 42, 62). However, although the hexamer TTAGGG prevails, many other variants of this sequence are present (22, 62). Furthermore, several cloned alpha-satellites lack the tandem arrays of precise TTAGGG hexamers that are required for TRF binding (21). On the basis of this sequence information, we do not expect these alpha-satellite repeats to be efficiently bound by TRF.

TRF has an exquisite ability to detect the differences between vertebrate telomeric repeats and the sequences of heterologous telomeres. Subtle changes in the (TTAGGG)$_n$ recognition site impair binding substantially. For instance, changes in the number of T residues or single-base substitutions dramatically decrease complex formation. The innermost ~0.5 kb of human telomeres is composed of G-rich repeats that are similar but not identical to the TTAGGG repeats (2, 9, 14). According to our data, TRF should not bind to this region but instead should associate with the major part of human telomeres, where long arrays of precise TTAGGG repetitions occur. TTAGGG repeats are the only known conserved sequence of vertebrate telomeres, and it is generally assumed that telomere function is conferred by this DNA. Thus, by these criteria TRF shows specificity for the

sequence of the conserved, functional part of vertebrate telomeric DNA. We suggest that TRF is a candidate telomere protein in mammalian cells. In agreement with this proposed function, TRF (or an activity similar to TRF) is expressed in all mammalian cell lines we have examined, including primate and rodent cell lines.

In addition to its sequence specificity, TRF can detect the length of TTAGGG repeat arrays. In vitro, TRF binding is greatly enhanced on probes with extensive TTAGGG repeat arrays. Although TRF can bind DNA with 3 copies of the $(TTAGGG)_n$ motif, probes with 6 or 12 repeats are recognized more efficiently. On the basis of the similar mobility of the TRF complexes with these probes, it appears that the enhanced binding to longer repeat arrays is not due to cooperative interactions of multiple factors. Furthermore, probes with six contiguous TTAGGG repeats are a better substrate for TRF than are probes with two separate blocks of three repeats. Therefore, the improved interaction with longer repeat arrays is not due to redundant representation of minimal sites within one molecule. One possibility is that TRF binding is dependent on a DNA structure that is more easily generated with longer TTAGGG repeat stretches. Since the DNA probes used in this study migrate according to their molecular weight in acrylamide gels, we have no evidence for bending in TTAGGG repeat stretches, but other structures are not excluded. Our data are most easily explained if the optimal site for the binding of a single TRF is composed of more than six contiguous TTAGGG repeats, i.e., more than 36 bp in length. In a recent compilation of recognition sites for DNA-binding factors, no site of this length was listed (19). Most of the binding sites are shorter than 15 bp, with a few exceptions of around 20 bp. Therefore, the optimal recognition sequence for TRF may be unusually long.

The interactions of TRF with TTAGGG repeat arrays of variable length is best explained by a binding factor with multiple DNA-binding domains. For example, TRF might be a complex of several monomers that each bind to one or a few repeats. Increasing saturation of these domains may then account for the enhanced binding to longer repeat arrays. A similar binding mode has been proposed for the interaction of heat shock factors with their repetitive recognition site (see reference 61 for a review). In addition, it is conceivable that a single TRF polypeptide contains multiple DNA-binding domains which each recognize one or a few TTAGGG repeats. Additive effects of these domains could contribute to enhanced binding on long contiguous TTAGGG repeat arrays. For example, TRF could be similar to TFIIIA, which recognizes a site of more than 50 bp with multiple Zn fingers (18, 43).

Since TRF recognizes TTAGGG repeats irrespective of the proximity of a DNA terminus, this factor could potentially bind along the length of telomeric repeat arrays. Factors that could bind along the length of the telomere have been found in *S. cerevisiae* and *Physarum* spp. (5, 10, 12, 13, 39–41, 65). The DNA-binding characteristics of TRF and the apparent molecular mass of this factor do not suggest any obvious biochemical similarity to (candidate) telomere proteins in other organisms.

If TRF is indeed a component of the nucleoprotein complex at mammalian chromosome ends, the binding of this factor could contribute to any of the functions ascribed to telomeres. One possibility is that TRF interacts with telomerase to modulate the replication of telomeres. Such a regulator has been proposed on the basis of the phenotype of telomerase mutants in *Tetrahymena* spp. (70), and the telomere protein RAP1 regulates the length telomeres in *S. cerevisiae* (12, 41, 65). TRF could also constitute a component of a protective complex at telomeres or mediate their subnuclear localization. We have previously shown that human telomeres are attached to the nuclear matrix (15) and chromosome ends are located near the nuclear periphery in many eukaryotes (1, 11, 52, 66). Thus, it will be of interest to examine the subnuclear localization of TRF when antibodies against this factor become available.

ACKNOWLEDGMENTS

We are particularly grateful to Harold Varmus for allowing us to initiate this work in his laboratory at the University of California, San Francisco. We thank Fritz Muller, University of Freibourg; Judith Berman, University of Minnesota; Eric Richards, Cold Spring Harbor Laboratories; Sylvie Le Blancq, Columbia University; and Janis Shampay, University of California, San Francisco, for gifts of plasmids. We thank Grant Hartzog, University of California, San Francisco, for essential advice and other insights. Neil Segil and Nat Heintz, The Rockefeller University, are gratefully acknowledged for their gift of whole-cell extracts from HeLa cells in different stages of the cell cycle. Judith Yanowitz is thanked for technical assistance. In addition, we thank Lisa Dailey, Maria Cardenas, and the members of the laboratory of Bob Roeder, all at The Rockefeller University, for advice and reagents. Our colleagues at The Rockefeller University; Art Lustig, Memorial Sloan Kettering Cancer Center; and David Shore, Columbia University, are thanked for helpful comments on the manuscript.

T.d.L. is a Lucille P. Markey Scholar, and this work was supported by the Lucille P. Markey Charitable Trust.

REFERENCES

1. **Agard, D. A., and J. W. Sedat.** 1983. Three-dimensional architecture of a polytene nucleus. Nature (London) **302:**676–681.
2. **Allshire, R. C., M. Dempster, and N. D. Hastie.** 1989. Human telomeres contain at least three types of G-rich repeat distributed non-randomly. Nucleic Acids Res. **17:**4611–4627.
3. **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 (London) **332:**656–659.
4. **Baeuerle, P. A., and D. Baltimore.** 1988. Activation of DNA-binding activity in an apparently cytoplasmic precursor of the NF-κB transcription factor. Cell **53:**211–217.
5. **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.
6. **Blackburn, E. H.** 1990. Telomeres and their synthesis. Science **249:**489–490.
7. **Blackburn, E. H.** 1991. Structure and function of telomeres. Nature (London) **350:**569–573.
8. **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.
9. **Brown, W. R. A., P. J. MacKinnon, A. Villasante, N. Spurr, V. J. Buckle, and M. J. Dobson.** 1990. Structure and polymorphism of human telomere-associated DNA. Cell **63:**119–132.
10. **Buchman, A. R., W. J. Kimmerly, J. Rine, and R. D. Kornberg.** 1988. 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.
11. **Chung, H.-M. M., C. Shea, S. Fields, R. N. Taub, L. H. T. van der Ploeg, and D. B. Tse.** 1990. Architectural organisation in the interphase nucleus of the protozoan Trypanosoma brucei: location of telomeres and mini-chromosomes. EMBO J. **9:**2611–2619.
12. **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.

13. **Coren, J. S., E. M. Epstein, and V. M. Vogt.** 1991. Characterization of a telomere-binding protein from *Physarum polycephalum*. Mol. Cell. Biol. **11:**2282–2290.
14. **Cross, S. H., J. Lindsey, J. Fantes, S. McKay, N. McGill, and H. Cooke.** 1990. The structure of a subterminal repeated sequence present on many human chromosomes. Nucleic Acids Res. **18:**6649–6657.
15. **de Lange, T.** 1992. Human telomeres are attached to the nuclear matrix. EMBO J. **11:**717–724.

15a. **de Lange, T.** Unpublished data.

16. **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 telomeres. Mol. Cell. Biol. **10:**518–527.
17. **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. Nucleic Acids Res. **11:**1475–1489.
18. **Engelke, D. R., S.-Y. Ng, B. S. Shastry, and R. G. Roeder.** 1980. Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. Cell **19:**717–728.
19. **Faisst, S., and S. Meyer.** 1992. Compilation of vertebrate-encoded transcription factors. Nucleic Acids Res. **20:**3–26.
20. **Fang, G., and T. R. Cech.** 1991. Molecular cloning of telomere-binding protein genes from Stylonychia mytilis. Nucleic Acids Res. **19:**5515–5518.
21. **Fanning, T.** 1987. Origin and evolution of a major feline satellite DNA. J. Mol. Biol. **197:**627–634.
22. **Fry, K., and W. Salser.** 1977. Nucleotide sequences of HS-alpha satellite DNA from kangaroo rat Dipodomys ordii and characterization of similar sequences in other rodents. Cell **12:**1069–1084.
23. **Gottschling, D. E., and T. R. Cech.** 1984. Chromatin structure of the molecular ends of Oxytricha macronuclear DNA: phased nucleosome and a telomeric complex. Cell **38:**501–510.
24. **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.
25. **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.
26. **Greider, C. W., and E. H. Blackburn.** 1985. Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. Cell **43:**405–413.
27. **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.
28. **Greider, C. W., and E. H. Blackburn.** 1989. A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis. Nature (London) **337:**331–337.
29. **Harley, C. B., A. B. Futcher, and C. W. Greider.** 1990. Telomeres shorten during ageing of human fibroblasts. Nature (London) **345:**458–460.
30. **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 (London) **346:**866–868.
31. **Henderson, E. R., and E. H. Blackburn.** 1989. An overhanging 3′ terminus is a conserved feature of telomeres. Mol. Cell. Biol. **9:**345–348.
32. **Henderson, E. R., C. C. Hardin, S. K. Walk, I. Tinoco, Jr., and E. H. Blackburn.** 1987. Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine-guanine base pairs. Cell **51:**899–908.
33. **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.
34. **Ido, J. W., A. Baldini, D. C. Ward, S. T. Reeders, and R. A. Wells.** 1991. Origin of human chromosome 2: an ancestral telomere-telomere fusion. Proc. Natl. Acad. Sci. USA **88:**9051–9055.
35. **Kang, C., X. Zhang, R. Ratliff, R. Moyzis, and A. Rich.** 1992. Crystal structure of four-stranded Oxytricha telomeric DNA. Nature (London) **356:**126–131.
36. **Kipling, D., and H. J. Cooke.** 1990. Hypervariable ultra-long telomeres in mice. Nature (London) **347:**400–402.
37. **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.
38. **Le Blancq, S. M., R. S. Kase, and L. H. T. van der Ploeg.** 1991. Analysis of a Giardia lamblia rRNA encoding telomere with (TAGGG)n as the telomere repeat. Nucleic Acids Res. **19:**5790.
39. **Liu, A., and B.-K. Tye.** 1991. A yeast protein that binds to vertebrate telomeres and conserved yeast telomeric junctions. Genes Dev. **5:**49–59.
40. **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 and is indistinguishable from RAP1. Curr. Genet. **16:**225–239.
41. **Lustig, A. J., S. Kurtz, and D. Shore.** 1990. Involvement of the silencer and USA binding protein RAP1 in regulation of telomere length. Science **250:**549–553.
42. **Meyne, J., R. L. Ratliff, and R. K. Moyzis.** 1989. Conservation of the human telomere sequence (TTAGGG)n among vertebrates. Proc. Natl. Acad. Sci. USA **86:**7049–7053.
43. **Miller, J., A. D. McLachlan, and A. Klug.** 1985. Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. EMBO J. **4:**1609–1614.
44. **Morin, G. B.** 1989. The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. Cell **59:**521–529.
45. **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.
46. **Muller, F., C. Wicky, A. Spicher, and H. Tobler.** 1991. New telomere formation after developmentally regulated chromosomal breakage during the process of chromatin diminution in Ascaris lumbricoides. Cell **67:**815–822.
47. **Petracek, M. E., P. A. Lefebvre, C. D. Silflow, and J. Berman.** 1990. Chlamydomonas telomere sequences are A+T rich but contain three consecutive G · C base pairs. Proc. Natl. Acad. Sci. USA **87:**8222–8226.
48. **Pluta, A. F., B. P. Kaine, and B. B. Spear.** 1982. The terminal organization of macronuclear DNA in Oxytricha fallax. Nucleic Acids Res. **10:**8145–8154.
49. **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:**3421–3431.
50. **Raghuraman, M. K., and T. R. Cech.** 1989. Assembly and self-association of Oxytricha telomeric nucleoprotein complexes. Cell **59:**719–728.
51. **Raghuraman, M. K., C. J. Dunn, B. J. Hicke, and T. R. Cech.** 1989. Oxytricha telomeric nucleoprotein complexes reconstituted with synthetic DNA. Nucleic Acids Res. **17:**4235–4253.
52. **Rawlins, D. J., M. I. Highett, and P. J. Shaw.** 1991. Localization of telomeres in plant interphase nuclei by in situ hybridization and 3D confocal microscopy. Chromosoma **100:**424–431.
53. **Richards, E. J., and F. M. Ausubel.** 1988. Isolation of a higher eukaryotic telomere from Arabidopsis thaliana. Cell **53:**127–136.
54. **Segil, N., S. Boseman Roberts, and N. Heintz.** 1991. Mitotic phosphorylation of the oct-1 homeodomain and regulation of oct-1 DNA binding activity. Science **254:**1814–1816.
55. **Sen, D., and W. Gilbert.** 1988. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. Nature (London) **334:**364–366.
56. **Shampay, J., and E. H. Blackburn.** 1989. Tetrahymena micronuclear sequences that function as telomeres in yeast. Nucleic Acids Res. **17:**3247–3260.
57. **Shampay, J., J. W. Szostack, and E. H. Blackburn.** 1984. DNA sequences of telomeres maintained in yeast. Nature (London) **310:**154–157.

58. **Shippen-Lentz, D., and E. H. Blackburn.** 1989. Telomere terminal transferase activity from *Euplotes crassus* adds large numbers of TTTTGGGG repeats onto telomeric primers. Mol. Cell. Biol. **9:**2761–2764.
59. **Shippen-Lentz, D., and E. H. Blackburn.** 1990. Functional evidence for an RNA template in telomerase. Science **247:**546–552.
60. **Smith, F. W., and J. Feigon.** 1992. Quadruplex structure of Oxytricha telomeric DNA oligonucleotides. Nature (London) **356:**164–168.
61. **Sorger, P. K.** 1991. Heat shock factors and the heat shock response. Cell **65:**363–366.
62. **Southern, E. M.** 1970. Base sequence and evolution of guinea-pig alpha-satellite DNA. Nature (London) **227:**794–798.
63. **Starling, J. A., J. Maule, N. D. Hastie, and R. C. Allshire.** 1990. Extensive telomere repeat arrays in mouse are hypervariable. Nucleic Acids Res. **18:**6881–6888.
64. **Sundquist, W. I., and A. Klug.** 1989. Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. Nature (London) **342:**825–829.
65. **Sussel, L., and D. Shore.** 1991. Separation of transcriptional activation and silencing functions of the RAP1-encoded repressor/activator protein 1: isolation of viable mutants affecting both silencing and telomere length. Proc. Natl. Acad. Sci. USA **88:**7749–7753.
66. **van Dekken, H., D. Pinkel, J. Mullikin, B. Trask, G. van den Engh, and J. Gray.** 1989. Three-dimensional analysis of the organization of human chromosome domains in human and human-hamster interphase nuclei. J. Cell Sci. **94:**299–306.
67. **Wells, R. A., G. G. Germino, S. Krishna, V. J. Buckle, and S. T. Reeders.** 1990. Telomere-related sequences at interstitial sites in the human genome. Genomics **8:**699–704.
68. **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.
69. **Wright, J. H., D. E. Gottschling, and V. A. Zakian.** 1992. Saccharomyces telomeres assume a non-nucleosomal chromatin structure. Genes Dev. **6:**197–210.
70. **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 (London) **344:**126–132.
71. **Zahler, A. M., and D. M. Prescott.** 1988. Telomere terminal transferase activity in the hypotrichous ciliate Oxytricha nova and a model for replication of the ends of linear DNA molecules. Nucleic Acids Res. **16:**6953–6972.
72. **Zakian, V. A.** 1989. Structure and function of telomeres. Annu. Rev. Genet. **23:**579–604.