A Human Telomeric Protein
Laura Chong, Bas van Steensel, Dominique Broccoli, Hediye Erdjument-Bromage, John Hanish, Paul Tempst, Titia de Lange†

Telomeres are multifunctional elements that shield chromosome ends from degradation and end-to-end fusions, prevent activation of DNA damage checkpoints, and modulate the maintenance of telomeric DNA by telomerase. A major protein component of human telomeres has been identified and cloned. This factor, TRF, contains one Myb-type DNA-binding repeat and an amino-terminal acidic domain. Immunofluorescent labeling shows that TRF specifically colocalizes with telomeric DNA in human interphase cells and is located at chromosome ends during metaphase. The presence of TRF along the telomeric TTAGGG repeat array demonstrates that human telomeres form a specialized nucleoprotein complex.

Human chromosomes carry a long terminal array of double-stranded TTAGGG hexamers that are maintained by telomerase. Telomeric DNA is thought to form a protective nucleoprotein cap through its association with telomere-specific proteins (1). Because the loss of telomere function can induce cell cycle arrest and genome instability, the telomeric complex is likely to be required in all human cells. Changes in the structure and function of human telomeres are thought to play a role in malignant transformation and cellular senescence (2, 3).

Protein components of the telomeric complex have been identified in ciliates and in yeast, but not in vertebrate systems (1). Quests for vertebrate telomeric proteins had previously yielded a single candidate factor that could potentially bind along the length of the telomeric TTAGGG repeat array (4–6). This protein, TRF (telomeric repeat binding factor), associates with double-stranded TTAGGG repeat arrays in vitro and displays strong specificity for vertebrate telomeric DNA (4, 5). TRF does not bind to single-stranded telomeric sequences and does not require the proximity of a DNA terminus for its interaction (4). The activity is expressed in nuclei from human, monkey, rodent, and chicken cells, which all carry TTAGGG repeat arrays at their chromosome ends (4). Here, we show that TRF is a protein component of human telomeres.

Human TRF (hTRF) activity can be detected in HeLa cell nuclear extracts on

---

Fig. 1. Purification and identification of the 60-kD hTRF protein. (A) Specific DNA affinity chromatography of hTRF. Partially purified HeLa hTRF was applied to a column containing restriction fragments with the sequence [TTAGGG]n coupled to Streptavidin-agarose (7). Input, flow-through (FT), and the indicated KCl fractions were assayed for hTRF binding activity with the use of a [TTAGGG]12 gel-shift probe. (B) Coomassie blue staining pattern of purified hTRF. The 60-kD TRF band is indicated. β-Casein was added to enhance hTRF activity in purified preparations (7). The asterisk at the right indicates a ~100-kD protein that is present in some of the hTRF preparations. Marker proteins (M) were prestained. (C) Recovery of hTRF activity by elution of the 60-kD protein from SDS-PAGE. Proteins from a gel similar to the one shown in (B) were eluted (5, 9), and hTRF activity was assayed by gel shift with a [TTAGGG]12 probe. Lanes 1 to 11 contain proteins isolated from successive gel slices covering the 120- to 20-kD range. Lane 7 contains proteins from the 55- to 65-kD range. (D) Analysis of hTRF tryptic peptides by chemical sequencing and laser-desorption mass spectrometry (10). Amino acids in lowercase were tentatively assigned; "x" indicates that no identification could be made (27). IY indicates calculated initial sequencing yields; m/z is the experimental mass of the peptide, [MH+] denotes the theoretical average isotopic mass of the peptide (plus one proton), calculated from the cDNA-derived sequence (Fig. 3C). Mox refers to methionine sulfoxide (singly oxidized methionine).
the basis of its ability to alter the mobility of a double-stranded DNA fragment containing the sequence [TTAGGG]$_4$. (4) Using this assay, we purified HeLa hTRF to near homogeneity by ion-exchange chromatography, by elution from columns containing nonspecific Escherichia coli DNA, and by fractionation over specific telomeric DNA resins (7) (Fig. 1A). Three independent preparations of purified hTRF contained a protein in the 60-kd apparent molecular mass range (Fig. 1B), which copurified with hTRF activity over a column containing double-stranded TTAGGG repeats (8). A ~100-kd protein was present in some but not all purified hTRF preparations (Fig. 1B) (8). Elution of the 60-kd protein from SDS-polyacrylamide gel electrophoresis (PAGE) (9) resulted in partial recovery of hTRF activity (Fig. 1C, lane 7), which indicates that a 60-kd polypeptide is sufficient for the formation of the hTRF complex with TTAGGG repeat probes. Amino acid sequences were obtained for trypsin peptides derived from the 60-kd band (10) (Fig. 1D), one of which (T29) revealed sequence identity to an anonymous partial complementary DNA (cDNA) sequence in the GenBank database. On the basis of this nucleotide sequence, cDNAs were isolated from a HeLa cell library, sequenced, and found to contain an open reading frame (ORF) encoding all sequenced peptides (11), as discussed below.

The hTRF cDNA hybridizes to two mRNAs of ~1.8 and ~3.0 kb that are expressed in a variety of human tissues (Fig. 2).

Sequence analysis of three overlapping hTRF cDNAs derived from the larger mRNA (11) revealed an ORF encoding a 439-amino acid protein (Fig. 3). The predicted molecular mass of this protein is 50,341 daltons, which is 10 kD smaller than the apparent molecular mass of purified HeLa hTRF. In vitro transcription and translation of the cloned cDNA produced a protein of the same size as purified HeLa hTRF (60 kD) (Fig. 3A), which indicated anomalous migration during SDS-PAGE. To verify that the cloned gene represented hTRF, we used in vitro-expressed protein in mobility-shift assays with a [TTAGGG]$_{12}$ probe. The in vitro-expressed protein formed a complex with the telomeric DNA probe that comigrated with the largest of three closely migrating gel-shift complexes formed with HeLa hTRF (Fig. 3B). The two smaller hTRF complexes detected in HeLa extracts were never observed with in vitro-expressed hTRF. Whether the additional HeLa hTRF complexes resulted from modification in vivo or from alteration of the hTRF protein during isolation (for instance, by partial proteolytic degradation or dephosphorylation) is not known. Competition experiments with cloned telomeric DNA in circular plasmids showed that the cloned hTRF protein and HeLa hTRF have the same sequence specificity and, as expected (4), do not require a DNA end for binding (Fig. 3B). For both the HeLa and cloned hTRF activities, the strongest competition was observed with a plasmid containing a TTAGGG repeat array; a plasmid with TTAGGG repeat competed to a lesser extent, and no competition occurred with a plasmid with TTAGGC repeats (Fig. 3B). These results demonstrated that the cloned cDNA encoded hTRF.

Comparison with the sequence information in the databases indicated that hTRF is a novel protein with three previously recognized sequence motifs (Fig. 3, C and D). hTRF contains two overlapping nucleoplasmin-type nuclear localization signals around position 350 and an NH$_2$-terminal region that is rich in aspartic and glutamic acid residues (11). The acidic domain of hTRF extends over a segment of 52 amino acids with 44% acidic residues and a calculated isoelectric point (pl) of 3.0. Although clus-
Fig. 4. Telomeric localization of TRF in human cells. (A through C) Colocalization of a FLAG epitope-tagged mTRF protein with telomeric DNA in interphase (15, 16). In (A), HeLa cells were transiently transfected with FLAG-mTRF expressed from the CMV promoter and labeled with an anti-FLAG mAb (M5) and FITC-conjugated donkey antibody to mouse IgG (green). In (B), telomeric DNA was visualized in the same nuclei by fluorescent in situ hybridization of digoxigenin-labeled [pC(CUA)₅]₈ DNA, followed by sheep anti-digoxigenin and TRITC-conjugated antibody to sheep IgG (red). In (C) is a superimposition of the images in (A) and (B). White and yellow indicate colocalization of the signals in (A) and (B). (D) Metaphase chromosomes of a HeLa cell expressing [HAl₅] epitope-tagged mTRF, labeled with an HA mAb (12C5) and a FITC-labeled goat antibody to mouse IgG (green) (15, 16). In each panel, the DNA was stained with DAPI (blue). Scale bars, 5 μm.

The nuclear and chromosomal location of TRF was determined by immunofluorescence microscopy of epitope-tagged mouse TRF (mTRF) protein expressed in transiently transfected HeLa cells (15, 16). Figure 4 shows that Met-Asp-Tyr-Lys-Asp-Asp-Asp-Lys (FLAG)-tagged mTRF displayed a punctate pattern in interphase nuclei. A similar speckled distribution was found for hemagglutinin [HAl₅]-tagged mTRF (17). A minority of the transfected cells showed homogeneous nuclear staining in addition to the speckled pattern, possibly because of greater amounts of mTRF expression (17). No specific patterns emerged in control experiments with untransfected cells or after omission of the primary antibody (17). Dual labeling experiments revealed that the speckled mTRF distribution coincided with telomeric DNA detected by TTAGGG repeat-specific fluorescent in situ hybridization (Fig. 4, B and C). All telomeric loci were found to contain mTRF, and, vice versa, all mTRF speckles were associated with telomeric DNA.

The chromosomal distribution of TRF was determined by expression of a fusion protein with two tandem HA epitopes at its NH₂-terminus. In metaphase chromosomes of stably transfected HeLa cells, [HAl₅]-tagged mTRF was predominantly detected at chromosome ends (Fig. 4D). Mouse TRF appeared to be a common feature of all chromosome ends. Occasionally, telomeres without mTRF could be found, but the absence of mTRF did not appear to be specific for any one chromosome end and is most likely a result of difficulties in detection of the protein in metaphase spreads. Discrete localization of mTRF to internal loci was not observed. In some metaphase spreads, weak staining along the axis of all chromosomes accompanied the telomeric signals (17). However, the strongest signals were invariably observed at chromosome ends.

These results demonstrate that TRF is a telomeric protein in vivo. TRF occupies chromosome ends both in interphase and metaphase, consistent with its role in the telomeric complex. Indirect evidence had previously suggested that mammalian chromosome ends contain one or more telomere-specific proteins (3–6, 18). For instance, human telomeres display an altered chromatin structure and bind to the nuclear matrix (18). Moreover, formation of new human telomeres upon DNA transfection occurs only with telomere seeds that contain precise TTAGGG repeat arrays (5). The failure of other (closely related) repetitive sequences to form new telomeres suggested that telomere healing involved a specific telomeric DNA-binding protein (5). However, telomeric proteins had not been isolated from human cells or other vertebrates (1, 6).
permeabilized in 0.5% Nonidet P-40 in phosphate-buffered saline, and labeled with mouse mAb 12CA5 followed by FITC-labeled goat antibody to mouse IgG (Jackson ImmunoResearch Labs). DNA was stained with DAPI (0.2 μg/ml). Images of interphase nuclei were obtained with a Zeiss Axioptan microscope equipped with a Kodak DCS-200 digital camera. Images were noise-filtered with a 3 × 3 median filter, corrected for background with nontransfected cells as a reference, and merged to obtain triple labeling images using Adobe Photoshop. Chromosome spreads were photographed on Kodak Gold II 400 ISO film and converted to digital images with a Nikon slide scanner, after which FITC and DAPI images were corrected for background and superimposed.

17. B. van Steensel and T. de Lange, data not shown.

27. Abbreviations for the amino acid residues are as follows: Ala, C; Asp, D; Asn, E; Glu, F; His, G; Ile, I; Lys, K; Leu, L; Met, M; Thr, N; Asp, P; Pro, Q; Gln, R; Arg, S; Ser, T; Thr, Trp; and Y, Tyr.
28. We thank B. Stillman and S. Bell for information about CFC1 and N. Andrews, R. Benesva, A. Bianchi, R. Bose, R. Darnell, J. Feng, R. Lang, M. Lui, A. Lustig, N. Segil, O. Siblon, S. Smith, and J. Young for their contributions and advice. Supported by grants from the Lucille P. Markey Charitable Trust and NIH (GM49048) to T.d.L., grants from the National Cancer Institute (P50 CA68748-29) and NSF (BIR-9240123) to P.T., and an Irma T. Hirschi Career Scientist Award to P.T. B.v.S. is the recipient of a Human Frontier Science Program postdoctoral fellowship.

3 October 1995; accepted 1 November 1995
results now demonstrate directly that telomeres form a complex with a telomere-specific protein, hTRF. In agreement with a role for hTRF at telomeres, the sequence specificity of hTRF matches the sequence requirements for telomere formation in human cells (6). On the basis of the data presented here, we suggest that TRF binding is required for telomere function at the ends of chromosomes in human and other mammalian cells.

TRF is a double-stranded telomeric DNA-binding factor. Double-stranded telomeric DNA-binding proteins have previously been implicated in telomere function in budding yeasts. In Saccharomyces cerevisiae, Raplp binds along the double-stranded telomeric TTAGGG tracts, where it regulates telomere length, chromosome stability, and telomeric position effects (19). Genetic alteration of the telomeric sequence of Kluyveromyces lactis has implicated a double-stranded telomeric DNA-binding protein in the regulation of telomerase (20). Although hTRF is not a Raplp homolog, we note that the recently determined structure of a Raplp1-DNA complex reveals two Myb-type IIT motifs (21). Further structural and functional comparisons between these telomeric factors should be of interest. Several observations suggest that telomere maintenance in mammals, as in yeast, is subject to homeostasis (6). For instance, mammalian telomeres are maintained at constant length over the generations, and different mammals show species-specific telomere length (22). Moreover, telomerase expression in immortalized human cells does not lead to unlimited telomere elongation (23). A simple model to explain such phenomena invokes a factor (such as TRF) that binds along the length of the telomeric repeat array and modulates telomerase-mediated telomere elongation (6, 20).

Human telomeres undergo programmed shortening in the soma (16, 24). When grown in vitro, human cells enter a growth crisis at a stage when telomeres appear critically shortened and chromosome end fusions are frequent (23). Similarly, loss of telomeric DNA in human tumors has been suggested to contribute to genome instability in cancer (2, 25). A mechanism to restore and maintain telomeres (such as activation of telomerase) may be required for tumor progression and cellular immortalization (26). Our findings raise the possibility that some of the deleterious consequences of telomere attrition may be caused by a failure of chromosome ends to bind protective telomeric proteins, including hTRF.

REFERENCES AND NOTES


11. hTRF cDNAs were isolated from a HeLa cDNA library (Stratagene) by screening with a 33-nucleotide DNA probe that represents the region of homology of 29 peptide sequence to an anonymous cDNA in GenBank (accession number Z19253). Three overlapping cDNAs, together spanning 2684 bp, were sequenced (GenBank accession number U47073). The sequence contains a single long ORF of 439 amino acids starting with an AAATGAG initiating codon, which is favorable for translation initiation (M. Kochel and C. Weisblum, unpublished data). The predicted end of the TRF cDNA contains an AATAAA polyadenylation (polyA) addition signal 18 bp upstream of the polyA tail. The TRF ORF amino acid sequence was compared to the nonredundant data base with the use of the program BLASTP (10). The nuclear targeting signals are of the nucleoplasmin type (J. Robbins, S. M. DiWorth, R. A. Laskey, C. Dingwall, Cell 64, 615 (1991); alignment to the Myb prototypical nucleoporin with the program GUSAT [C. Grochulski, C. R. 1.5; a SeqEd 1.01 display and Rao rules for conservative amino acid changes were used.


15. A cDNA for the mouse homolog of hTRF was isolated by cross-hybridization to the human cDNA (D. Broccoli, L. Huang, T. de Lange, unpublished data). The ORF in this cDNA was fused to a FLAG epitope. HeLa cells with long telomeres in the 15- to 30-kb size range (T. de Lange, EMBO J. 11, 717 (1992) were transfected with a mammalian expression vector containing the FLAG-hTRF gene expressed from the cytomegalovirus (CMV) promoter and grown for 24 hours on Acantho Coated substrate. Cells were washed in phosphate-buffered saline, fixed in 2% formaldehyde, permeabilized in phosphate-buffered saline for 10 min, and treated with 0.5% Nonidet P-40 in phosphate-buffered saline for 20 min. In situ hybridization was carried out essentially as described (D.C. M. Sibon et al., Histochemistry 101, 223 (1994)) with the use of a digoxigenin-labeled [CCUCUA]3 repeat RNA probe made by in vitro transcription of plasmid pTHS (19) using 11-digoxigenin-5'-triphosphosphate-labeled mR9 ribonucleotides with yeast antibody to digoxigenin (anti-digoxigenin) (Boehringer) and mouse monoclonal antibody (mAb) M6 (Eastman Kodak) against the FLAG epitope, followed by Alexa fluor 488 conjugated to tetramethyl rhodamine isothiocyanate (TRITC)-conjugated donkey antibody to sheep immunoglobulin G (lgG) and fluorescein isothiocyanate (FITC)-conjugated donkey antibody to mouse IgG (both from Jackson ImmunoResearch Labs). DNA was stained with 4',6-diamino-2-phenylindole (DAPI, 0.2 µg/ml). Cells were embedded in p-phenylene diamine (1 mg/ml) (Sigma) and 90% glycerol in phosphate-buffered saline for 20 min. After the omission of the [CCUCUA]3 repeat RNA probe or anti-digoxigenin, or when nuclear DNA was not denatured, only the FITC signal was found. In untransfected cells or after omission of mAb M5, only the TRITC signal was observed. Formalin-fixed, paraffin-embedded sections were prepared essentially as described (D.A. Compton, T. J. Yen, D. W. Cleveland, J. Cell Biol. 112, 1083 (1991)). HeLa cells were stably transfected by calcium phosphate-buffered saline (pH 7.4). After omission of the [CCUCUA]3 repeat RNA probe or anti-digoxigenin, or when nuclear DNA was not denatured, only the FITC signal was found. In untransfected cells or after omission of mAb M5, only the TRITC signal was observed. Formalin-fixed, paraffin-embedded sections were prepared essentially as described (D.A. Compton, T. J. Yen, D. W. Cleveland, J. Cell Biol. 112, 1083 (1991)). HeLa cells were stably transfected by calcium phosphate-buffered saline (pH 7.4). After omission of the [CCUCUA]3 repeat RNA probe or anti-digoxigenin, or when nuclear DNA was not denatured, only the FITC signal was found. In untransfected cells or after omission of mAb M5, only the TRITC signal was observed.