

***Euplotes* telomerase contains an La motif protein produced by apparent translational frameshifting**

**Stefan Aigner, Joachim Lingner¹,
Karen J. Goodrich, Cheryl A. Grosshans,
Andrej Shevchenko², Matthias Mann^{2,3} and
Thomas R. Cech⁴**

Department of Chemistry and Biochemistry and Howard Hughes Medical Institute, University of Colorado, Boulder, CO 80309-0215, USA and ²European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69012 Heidelberg, Germany

¹Present address: Swiss Institute for Experimental Cancer Research (ISREC), 155 Chemin des Boveresses, CH-1066 Epalinges, Switzerland

³Present address: Protein Interaction Laboratory, University of Southern Denmark – Odense, Campusvej 55, DK-5230 Odense M, Denmark

⁴Corresponding author
e-mail: Thomas.Cech@Colorado.Edu

Telomerase is the ribonucleoprotein enzyme responsible for the replication of chromosome ends in most eukaryotes. In the ciliate *Euplotes aediculatus*, the protein p43 biochemically co-purifies with active telomerase and appears to be stoichiometric with both the RNA and the catalytic protein subunit of this telomerase complex. Here we describe cloning of the gene for p43 and present evidence that it is an authentic component of the telomerase holoenzyme. Comparison of the nucleotide sequence of the cloned gene with peptide sequences of the protein suggests that production of full-length p43 relies on a programmed ribosomal frameshift, an extremely rare translational mechanism. Anti-p43 antibodies immunodeplete telomerase RNA and telomerase activity from *E.aediculatus* nuclear extracts, indicating that the vast majority of mature telomerase complexes in the cell are associated with p43. The sequence of p43 reveals similarity to the La autoantigen, an RNA-binding protein involved in maturation of RNA polymerase III transcripts, and recombinant p43 binds telomerase RNA *in vitro*. By analogy to other La proteins, p43 may function in chaperoning the assembly and/or facilitating nuclear retention of telomerase.

Keywords: La autoantigen/nuclear retention/RNA chaperone/telomerase/translational frameshift

Introduction

The linearity of eukaryotic chromosomes poses a challenge to the cellular DNA replication machinery: their ends cannot be replicated by conventional DNA polymerases, a situation that leads to gradual loss of genetic information in the absence of a replenishing mechanism (Watson, 1972; Olovnikov, 1973; Singer and Gottschling, 1994; Blasco *et al.*, 1997; Nakamura *et al.*, 1997). For most eukaryotes, the solution to this end-replication problem is the enzyme telomerase, which synthesizes telomeric DNA, repetitive sequences at the chromosome ends.

Telomerase is a ribonucleoprotein (RNP) complex (Greider and Blackburn, 1987), consisting minimally of a catalytic protein subunit and an RNA subunit. While this RNA is transcribed by RNA polymerase III (pol III) in ciliates (Greider and Blackburn, 1989; Shippen-Lentz and Blackburn, 1990; Yu *et al.*, 1990; Lingner *et al.*, 1994), it is a pol II transcript in the yeast *Saccharomyces cerevisiae* (Chapon *et al.*, 1997) and probably in vertebrates (Hinkley *et al.*, 1998; Chen *et al.*, 2000). Using a portion of its intrinsic RNA moiety as the template for synthesis of telomeric DNA repeats (Greider and Blackburn, 1989), telomerase functions as a specialized reverse transcriptase (RT). In fact, the catalytic protein subunit of telomerase (TERT, telomerase reverse transcriptase) is structurally and evolutionarily related to other RTs (Eickbush, 1997; Lingner *et al.*, 1997a; Nakamura and Cech, 1998). TERT and the RNA subunit have been identified in organisms including ciliates, yeasts and mammals (for review see Bryan and Cech, 1999), supporting the idea that these two components constitute the functionally conserved telomerase core.

In contrast, less is known about the roles of additional factors present in telomerase holoenzymes, which appear highly variable across species. Biochemical purification of *Tetrahymena thermophila* telomerase yielded two proteins, p80 and p95 (Collins *et al.*, 1995), which bind *in vitro* to telomerase RNA and to telomeric DNA, respectively (Collins and Gandhi, 1998; Gandhi and Collins, 1998). Mammalian p80 homologs, termed TEP1, were subsequently found (Harrington *et al.*, 1997; Nakayama *et al.*, 1997). Our knowledge of the possible functions of these proteins is limited to *in vitro* binding of p80 and p95. Telomerase from *S.cerevisiae* contains two accessory protein factors, Est1p and Est3p, which were identified in genetic screens (Lundblad and Szostak, 1989; Lendvay *et al.*, 1996) and are unrelated by sequence to p80 or p95. Little is known about the function of Est3p, but a distinct role for Est1p is emerging: this protein binds both telomeric DNA (Virta-Pearlman *et al.*, 1996) and the RNA subunit (Zhou *et al.*, 2000), and appears to aid telomerase in locating and/or positioning itself at the telomere (Evans and Lundblad, 1999). While these four telomerase subunits are dispensable for *in vitro* core enzymatic activity (Lingner *et al.*, 1997b; Collins and Gandhi, 1998; Bryan *et al.*, 2000), at least Est1p and Est3p are essential for yeast telomerase function *in vivo* (Lendvay *et al.*, 1996).

Other protein subunits have been implicated in the assembly of the telomerase holoenzyme. Telomerase activity from human (Weinrich *et al.*, 1997; Beattie *et al.*, 1998) and the ciliate *T.thermophila* (Collins and Gandhi, 1998) can be reconstituted by combining the purified RNA component with TERT synthesized in rabbit reticulocyte lysates. However, this *in vitro* assembly of the telomerase RNP requires the contribution of factors supplied by the reticulocyte extract (Holt *et al.*, 1999; Licht and Collins, 1999). In the human case, these have been shown to

include the molecular chaperones p23 and Hsp90, which appear to remain bound in the active holoenzyme (Holt *et al.*, 1999). *Saccharomyces cerevisiae* telomerase RNA binds the same Sm proteins that direct the transport and assembly of small nuclear RNPs (snRNPs), and in the absence of the binding site for these proteins the accumulation of telomerase is severely reduced (Seto *et al.*, 1999).

We now investigate another telomerase accessory factor, p43. This protein was first identified by biochemical purification of active telomerase from the hypotrichous ciliate *Euplotes aediculatus* (Lingner and Cech, 1996). The molecular mass of the isolated complex was consistent with an RNP stoichiometry of one molecule each of p123 (the *Euplotes* TERT), the RNA subunit and p43, which appeared as a doublet on SDS-polyacrylamide gels (Lingner and Cech, 1996). However, the possibility remained that p43 might merely co-purify with telomerase and not represent an authentic subunit. We now report cloning of the gene encoding p43. We show that this protein is associated with most or all active telomerase and that it is related to the La class of proteins, which are known to bind the oligouridylylate stretch at the 3' end of pol III transcripts and to function in RNP biogenesis. Most pol III transcripts lose their 3'-Us and the La proteins, and are exported to the cytoplasm. Ciliate telomerase RNAs retain the 3'-Us in their mature form, so our finding that one of these RNAs remains stably associated with La or a La-related protein provides new insight regarding how nuclear retention of some telomerases may be achieved.

Results

Cloning of the gene for p43

Biochemical purification of *E. aediculatus* telomerase revealed a 43 kDa doublet whose two components were suggested to be related to each other by partial proteolysis or post-translational modification (Lingner and Cech, 1996). This doublet of bands was excised from an SDS-polyacrylamide gel and digested with trypsin. Nano-electrospray tandem mass spectrometric analysis of the resultant peptides allowed us to determine full-length sequences of six tryptic peptides and partial sequences of an additional six tryptic peptides (Figure 1).

A DNA probe was generated by PCR using genomic DNA as a template and degenerate oligonucleotide primers with sequences based on a part of peptide T_4 (Figure 1B) and a part of peptide t_5 . Since macronuclear chromosomes of hypotrichous ciliates contain only a single gene, screening of an *E. aediculatus* genomic DNA library with this probe yielded the entire macronuclear chromosome harboring the p43 gene (Figure 2A). Conceptual translation of this DNA sequence revealed 11 of the peptides obtained by mass spectrometry (Figures 1 and 2B), thus validating that we had indeed cloned the gene coding for p43 and providing some support for the notion that both components of the doublet are encoded by the same gene.

Gene structure

Peptides could be matched with the macronuclear DNA sequence in all three reading frames, suggesting the presence of introns. Therefore, we sequenced p43 cDNA prepared from poly(A)⁺-selected mRNA and compared it with the genomic sequence. The seven introns found in the

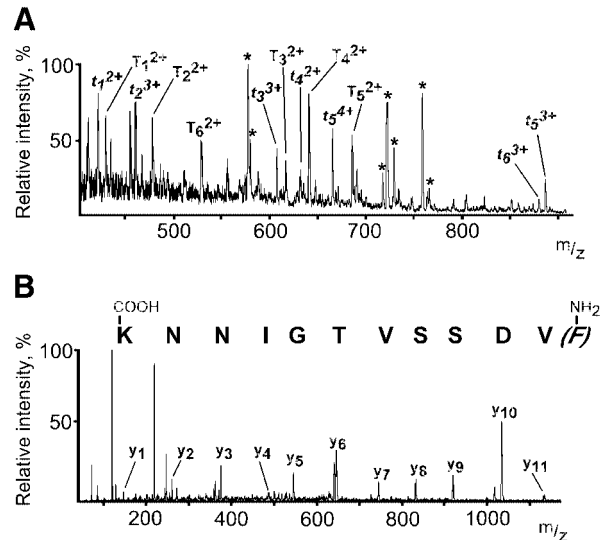


Fig. 1. Sequencing of p43 by nano-electrospray tandem mass spectrometry. *m/z*, mass-to-charge ratio. (A) Mass spectrum of the unfractionated digest of the p43 doublet band. Trypsin autolysis products are marked with asterisks. Tryptic peptides whose entire sequences were determined are designated with T. Tryptic peptides sequenced partially and retrospectively matched to the full-length sequence are designated with *t*. (B) Mass spectrum of product ions resulting from fragmentation of peptide precursor ion T_4^{2+} from (A). Interpretation of these mass spectra relative to those obtained from esterified peptides allowed assignment of *y*-ions and sequence determination.

p43 transcript (Figure 2A) make it the most intron-rich *Euplotes* mRNA known to date (Hoffman *et al.*, 1995). The introns are very short, ranging in size from 26 to 92 nucleotides (nt) (mean 43 nt) and are very AT rich (mean 75%, range 66–84%); protein-coding regions of the p43 gene contain 67.7% AT, as is common with genes from these protozoa (Hoffman *et al.*, 1995). In good agreement with the previously identified consensus for hypotrichs (Hoffman *et al.*, 1995), the introns are flanked by the 5' and 3' splice site sequences 5'/GTAag...TAG/3', where capital and lower case letters indicate bases conserved in 100% and >70% of the seven introns, respectively, and solidi represent splice sites. The telomere-proximal 5' and 3' non-coding regions are extremely short (55 and 34 nt, respectively) and AT rich (76 and 88%, respectively), as commonly observed in hypotrich chromosomes (Hoffman *et al.*, 1995). These flanking sequences contain good matches to the core element of the *Euplotes* chromosomal breakage sequence (E-Cbs) 5'-ATTGAA-3' (Klobutcher *et al.*, 1998; data not shown), which directs the excision of macronuclear chromosomes from the micronuclear genome. The ends of the macronuclear DNA molecule are demarcated by the double-stranded telomeric sequence 5'-(C₄A₄)₄-3' (Figure 2A).

Indication of a programmed +1 frameshift in p43 translation

Analysis of the cDNA sequence did not reveal a single open reading frame (ORF) capable of encoding all 11 peptides that were identified by mass spectrometry and are assignable to p43. Instead, an upstream ORF predicts an 88 amino acid polypeptide containing one of

the sequenced peptides (peptide t_1 , Figure 2B). A second potential protein-coding region is found in a reading frame that is shifted by +1 relative to the first ORF and that partially overlaps with it (Figure 3). Translation of this region predicts a 352 amino acid protein (with a calculated molecular mass of 40.9 kDa) containing the remaining 10 peptides. However, no ATG start codon can be found upstream of these 10 peptides in the second ORF. To rule out a sequencing error as the source of the apparent

interruption of a single ORF, we amplified and sequenced cDNA from a different RNA preparation, which yielded identical results. In addition, the macronuclear chromosome encoding p43, when conceptually spliced using the splice site consensus described above, predicts the transcript sequence that was obtained experimentally. The fact that the change in reading frame is predicted from both genomic and cDNA also excludes the possibility of insertional or deletional RNA-editing events leading to the observed discontinuity of the ORF. Thus, translation of p43 is likely to rely on a ribosomal frameshift at a site between peptides t_1 and t_3 (Figure 3).

Examination of the DNA sequence allowed us to map the frameshift site further (Figure 3). Reading frame zero (i.e. the reading frame of the initiating methionine) encodes peptide t_1 and a stop codon farther downstream, while reading frame +1 encodes a stop codon and peptide t_3 farther downstream. The frameshift event is, therefore, theoretically restricted to occur between these two stop codons. A shift in the reading frame forward by one nucleotide could occur at any one of four sites, resulting in four possible protein products differing at positions 86, 87 and 88 (Figure 3). Since we were unable to obtain peptides that actually span the frameshift site, we cannot distinguish between the four locations and respective translation products. Fortunately, however, the frameshift occurs within the conserved La motif of p43. Alignment of p43 with La proteins from a set of evolutionarily diverse organisms (Figure 4, top, and see below) reveals a highly conserved phenylalanine residue at a position equivalent to amino acid 87 in p43 (marked with an asterisk in Figure 4), followed by a less conserved residue (asparagine or lysine). This alignment argues in favor of a frameshift at site 3 or 4 (Phe₈₇Asn₈₈ or Phe₈₇Lys₈₈, respectively), because sites 1 and 2 both give rise to a non-conserved leucine residue at position 87. For the sequence analyses described below, and to construct a frameshift-corrected bacterial expression plasmid, we arbitrarily chose asparagine to be at position 88.

The frameshifted full-length translation product is predicted to be a basic protein (theoretical pI = 9.35) of 437 amino acids with a molecular mass of 51 kDa. As described below, the discrepancy between the apparent molecular mass of p43 on SDS gels as observed in our

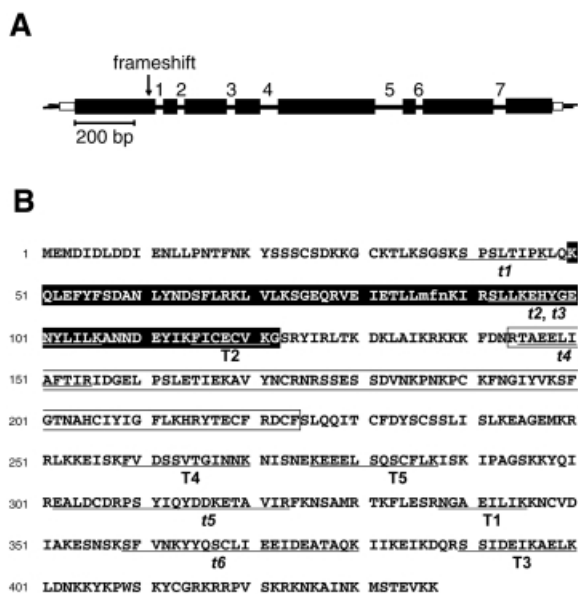


Fig. 2. p43 gene and protein. (A) Structure of the macronuclear chromosome containing the p43 gene. Protein-coding regions are depicted by solid black boxes, introns are numbered and shown as lines. 5' and 3' non-coding regions are shown as empty boxes proximal to the telomeres, which are represented with their 3'-terminal single-stranded extensions. The location of the putative translational frameshift site is marked by an arrow. (B) Predicted amino acid sequence of the p43 gene product. Peptides sequenced *de novo* from purified protein and assigned to p43 are shown underlined and are labeled as in Figure 1A. Peptide T_6 could not be assigned to p43. Amino acids 86–88, whose identity could not be determined unambiguously due to the putative frameshift, are in lower case letters. Shown here is the sequence assuming the frameshift event at site 3 (see Figure 3). The La motif and the putative RRM are shaded and boxed, respectively.

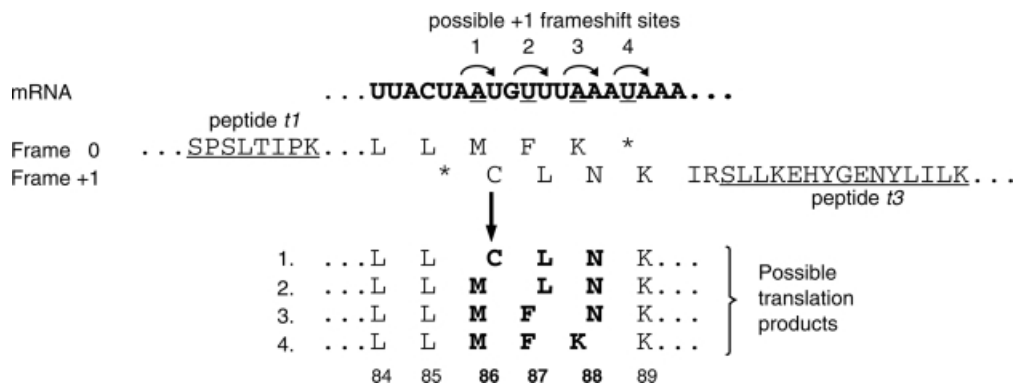


Fig. 3. Close-up of the frameshift region. The mRNA sequence as deduced from cDNA is shown in bold at the top. The four theoretically possible +1 frameshift events are illustrated by curved arrows above the 'skipped' nucleotides, which are underlined. Conceptual translations in the reading frame of peptide t_1 (designated frame 0) and peptide t_3 (designated frame +1) are aligned below the mRNA sequence using the one-letter amino acid code and asterisks to denote stop codons. The translation products originating from each of the four possible +1 frameshifts are given below the vertical arrow. Bold letters signify amino acids whose identity depends on the actual site of frameshifting.

La motif

H.s. 19 HQEYFFGCDNLRDKFLKEQT-KLDEGWVPLEIMIKFN
 D.m. 56 RQVEYFFGDANLRDKFLREQHGKNEIDGWVPLSVLYTFK
 S.p. 69 KQVEFYFSDTNLPHDKFLWTTGK-NDGWVPIOTTIANFK
 S.c. 35 KQVEFYFSEFNFPMDRFLRTTAEK-NDGWVPISTIATFN
 E.a. 49 KQLEFYFSDANLYNDSFLRKLVLKSGEQRVEIETLLMFN

H.s. RLNRLTLD-F--NVIVEALSKSKAELMEISEDKTKIRR 91
 D.m. RLASLSTD-LSE--IIVAAINKSEEGLEVEISEDKLSLRR 129
 S.p. RMRRFQP--L-E-AIVNALRKS-PELLEVEDEAGEKVR 149
 S.c. RMKKYRP--V-DK-VIEALR-S-SEILEVSDGENVVKR 104
 E.a. KIRSELKSHYGENYLI--LKANNDEYIKFI--CECVKG 122

RRM motif

RNP2 RNP1
 H.s. 111 RSVY-IKGF... (29 amino acids)... AFKGSIFVVF 158
 D.m. 149 RTAY-AKGF... (34 amino acids)... KFKGSIFLAF 201
 S.p. 154 RSVY-CKGF... (32 amino acids)... KFKGSVFVVF 204
 S.c. 123 RTLA-VMNF... (38 amino acids)... KFNCTVVLVVF 179
 E.a. 144 RTAEELIAR... (37 amino acids)... KFNCTYVKSF 200

Fig. 4. Relationship of p43 to La proteins. Parts of the predicted protein sequence of p43 (E.a.) are aligned with the respective regions of La proteins from human (H.s.), *D.melanogaster* (D.m.), the fission yeast *S.pombe* (S.p.) and *S.cerevisiae* (S.c.) (Chambers *et al.*, 1988; Yoo and Wolin, 1994; van Horn *et al.*, 1997). Amino acids in black shading are similar or identical in the majority of the five proteins (see Materials and Methods). Numbers to the left and right of each of the sequences refer to their start and end amino acid positions, respectively. Top: alignment of the La motif regions. The three amino acids in p43 whose identity is uncertain due to lack of information on the exact location of the frameshift are underlined. The sequence shown here assumes the frameshift event at site 3 (see Figure 3). The phenylalanine residue conserved in La proteins is marked by an asterisk. Bottom: alignment of parts of the putative RRM region. The RNP2 and RNP1 submotifs are overlined.

previous report (Lingner and Cech, 1996) and its predicted molecular mass can be explained by proteolysis during purification.

p43 is an La protein homolog

A database search with the BLAST algorithm using the full-length p43 sequence revealed significant, albeit weak, homology to the La autoantigen protein family. These proteins are known from a variety of evolutionarily distant organisms and have been implicated in several aspects of RNP assembly and maturation of pol III transcripts. Authentic La proteins share amino acid identities and similarities of ~12–27 and ~26–43%, respectively (Yoo and Wolin, 1994; van Horn *et al.*, 1997; data not shown), and p43 is 11–13% identical and 22–29% similar to the La homologs from *Homo sapiens*, *Drosophila melanogaster*, *S.cerevisiae* and *Schizosaccharomyces pombe* (data not shown).

Based on sequence analysis, authentic La proteins can be divided into three regions (Yoo and Wolin, 1994): (i) a characteristic, highly conserved La motif located at the N-terminus and comprising ~70 amino acids; (ii) a loosely conserved RNA recognition motif (RRM), which is also found in a variety of other RNA-binding proteins and contains two signature submotifs, RNP2 and RNP1 (Bandziulis *et al.*, 1989; Kenan *et al.*, 1991; Birney *et al.*, 1993; Burd and Dreyfuss, 1994); and (iii) a highly charged, usually weakly conserved C-terminus. p43 appears to meet all three sequence criteria. Figure 4 (top) shows an alignment of the La motifs of La homologs with the corresponding region of p43. Within this region, p43 shares 25–31% identity and 51–54% similarity with

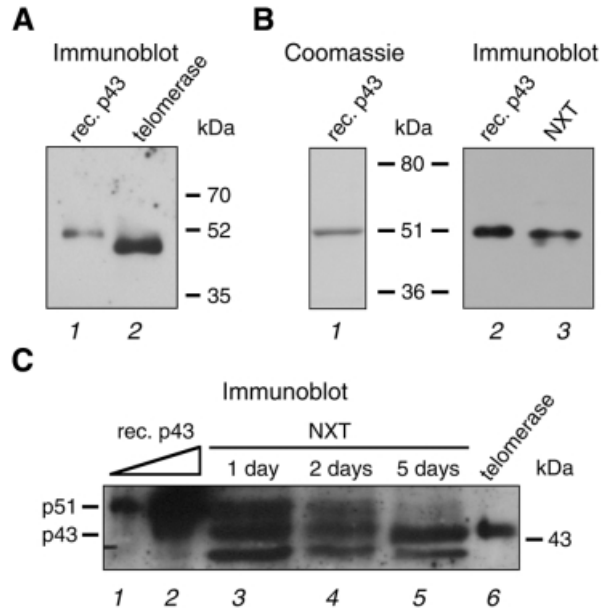


Fig. 5. Affinity-purified α -p43 antibodies against recombinant p43 recognize a ~51 kDa protein in *E.aediculatus* nuclear extract but a ~43 kDa protein in purified telomerase. (A) Purified recombinant His₆-tagged p43 (lane 1) and 0.4 pmol of telomerase purified by heparin chromatography and affinity chromatography (lane 2) were denatured and separated on a 4–20% polyacrylamide gel, followed by immunoblot analysis with α -p43 antibodies. (B) Recombinant p43 (lanes 1 and 2) and nuclear extract (NXT, lane 3) were electrophoresed as in (A), followed by either Coomassie Blue staining (left) or immunoblot analysis (right). Recombinant p43 migrates with slightly decreased mobility relative to endogenous p43 due to its His₆ tag, which adds ~2 kDa (data not shown). (C) Recombinant p43 (lanes 1 and 2), nuclear extract that was stored at 4°C for 1 day (lane 3), 2 days (lane 4) or 5 days (lane 5), or 0.04 pmol of purified telomerase (lane 6) were denatured and electrophoresed on an 8% polyacrylamide gel and subjected to immunoblot analysis. p51 and p43 designate the proteins migrating at apparent mol. wts of 51 and 43 kDa, respectively.

previously characterized La proteins, compared with 36–57% identity and 60–73% similarity among the La motifs of established La proteins. C-terminal to the La motif, p43 contains weak yet discernible RNP2 and RNP1 signature submotifs whose spacing (37 amino acids) is consistent with these submotifs being part of an RRM (Figure 4, bottom). Lastly, the C-terminal third of p43 contains ~40% charged residues, a value similar to that of the comparably sized human La protein (Figure 2B; Chambers *et al.*, 1988). In summary, these data suggest that p43 is related to authentic La proteins in terms of its protein structure.

p43 is a proteolysis product of a 51 kDa precursor

Recombinant p43 was expressed in *Escherichia coli* (Figure 5B, lane 1), purified and used to immunize rabbits. Anti-p43 antibodies affinity purified from immune serum recognized recombinant p43, which migrates at ~51 kDa (Figure 5A, lane 1), and a ~43 kDa protein in telomerase purified from *E.aediculatus* nuclear extract by heparin chromatography and affinity chromatography (lane 2). In contrast, the antibodies detect a ~51 kDa species in nuclear extract (Figure 5B, lane 3) that co-migrates with recombinant p43 (lane 2). During prolonged storage of nuclear extract at 4°C, the abundance of the 51 kDa species decreases and this protein appears to be chased

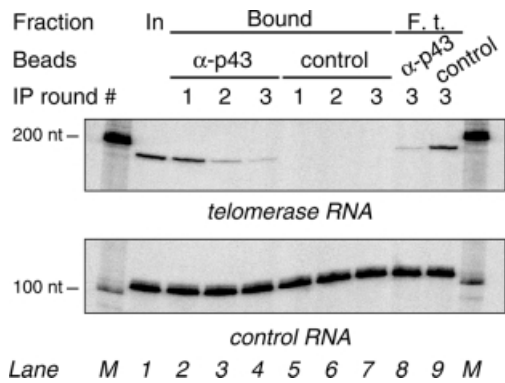


Fig. 6. α-p43 antibody efficiently co-immunoprecipitates telomerase RNA from nuclear extract. *Euplotes aediculatus* nuclear extract was subjected to three successive immunodepletion reactions with either α-p43 antibody beads or control beads. After each round of immunodepletion, supernatants were separated from the beads and added to fresh beads. After addition of a 110mer RNA as a recovery control, samples containing equal amounts of input nuclear extract (In, lane 1), the three α-p43 bead fractions (Bound, lanes 2–4) and control bead fractions (Bound, lanes 5–7) as well as the final flowthroughs (F.t.) of the α-p43 beads (lane 8) and control beads (lane 9) were treated with protease, phenol extracted and analyzed by northern blot hybridization with probes specific for the 189 nt *E.aediculatus* telomerase RNA and the control RNA. Lanes M, uniformly labeled RNAs as size markers.

into the 43 kDa protein (Figure 5C, lanes 3–5). Therefore, we conclude that the 43 kDa protein recognized by our antibodies is identical to p43 from purified telomerase and that this protein is derived from a 51 kDa precursor by partial proteolysis. An analogous degradation artifact has been observed with the human La protein: in purified HeLa cell extracts, α-La antisera recognize a ~40 kDa species that is interpreted to be a breakdown product of the full-length 47 kDa human La protein (Stefano, 1984).

Telomerase RNA is present in a complex with p43 in vivo

To assess what fraction of telomerase RNA is associated with p43 in *vivo*, we bound affinity-purified α-p43 antibodies to protein A-agarose beads and used these for immunoprecipitation studies. Following incubation of nuclear extract with α-p43 beads or control beads prepared from pre-immune IgG, the supernatants were separated from the beads and subjected to two additional rounds of immunodepletion using fresh beads (Figure 6). Input nuclear extract (lane 1), as well as the RNA fractions retained on the beads (lanes 2–7) and those remaining in the final supernatants (lanes 8 and 9), were then treated with protease, phenol extracted and ethanol precipitated. Northern blot analysis with a probe to telomerase RNA showed that the RNA is efficiently retained on the α-p43 beads, but not on the control beads. PhosphorImager quantitation of the results from three independent experiments revealed the extent of immunodepletion: while an average of 95% of input telomerase RNA (lane 1) was captured on the α-p43 beads (lanes 2–4) and 5% remained in the supernatant after three successive rounds of immunodepletion (lane 8), no telomerase RNA was detectable on the control beads (lanes 5–7) and the supernatant contained 89% of input telomerase RNA (lane 9). These data indicate that the

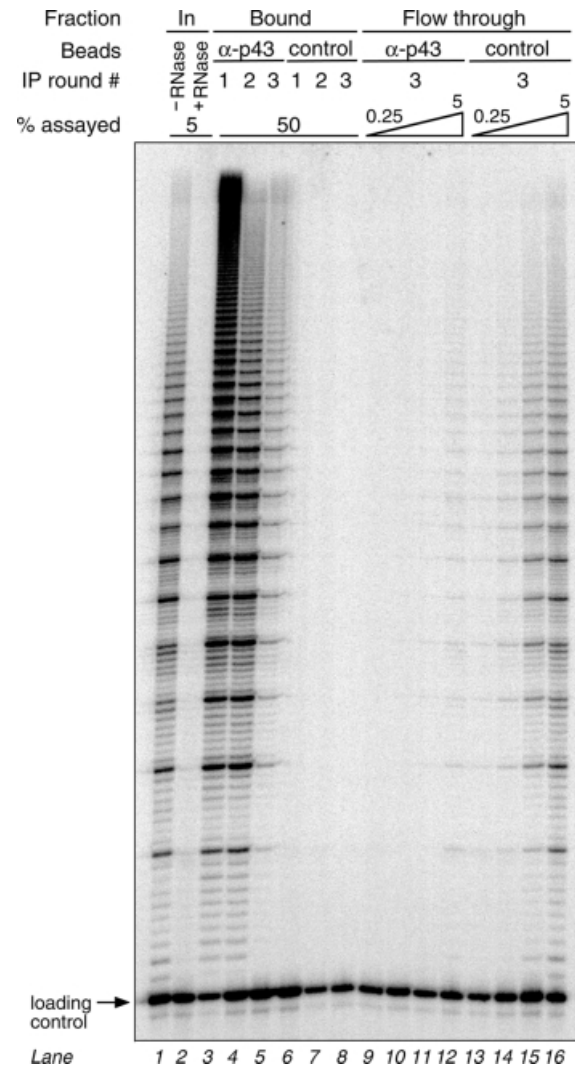


Fig. 7. The vast majority of telomerase activity is associated with p43. *Euplotes aediculatus* nuclear extract was partially purified by glycerol gradient centrifugation and subjected to three consecutive rounds of immunodepletion with α-p43 antibody beads or control antibody beads, as described in the legend to Figure 6. Input, bound and flowthrough fractions were assayed for telomerase activity by measuring extension of a telomeric primer in the presence of a radiolabeled substrate nucleotide, followed by phenol/chloroform extraction and separation of telomerase products in a denaturing polyacrylamide gel. Five percent of input nuclear extract before (lane 1) or after (lane 2) RNase A treatment, 50% of α-p43 antibody beads (lanes 3–5) or control antibody beads (lanes 6–8), as well as 0.25, 0.5, 2.5 and 5% of the final flowthroughs from the α-p43 antibody beads (lanes 9–12) or control antibody beads (lanes 13–16) were assayed. Equal amounts of an end-labeled 23mer oligonucleotide were added to the samples prior to phenol/chloroform extraction to control for recovery and loading.

vast majority of telomerase RNA molecules are bound to p43 or present in a complex with p43 in *vivo*.

Virtually all active telomerase complexes contain p43

To address the question of whether p43 is part of active telomerase complexes, we performed similar immunoprecipitation experiments as described above, using telomerase that had been partially purified from nuclear extracts by glycerol gradient centrifugation. Here, however, we assayed the bound and flowthrough fractions for telomerase activity. As shown in Figure 7, the input telomerase

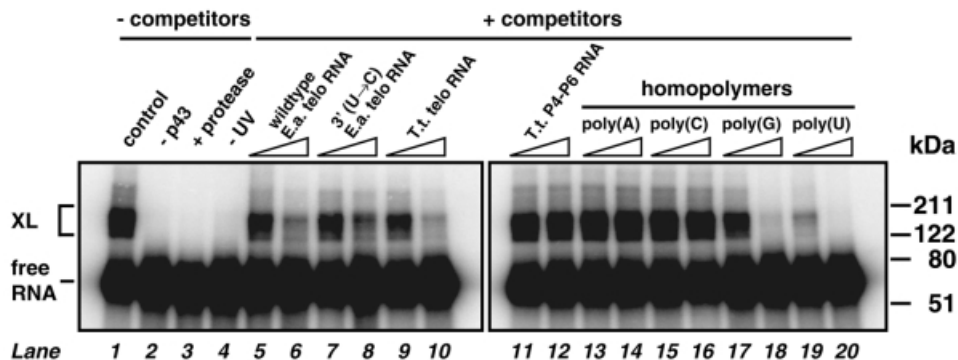


Fig. 8. Purified recombinant p43 binds *in vitro* transcribed *E.aediculatus* telomerase RNA *in vitro*. ^{32}P -labeled telomerase RNA (~ 6.6 nM) carrying the photoaffinity chromophore 5-iodo-U at all U positions was incubated on ice with 330 nM p43 (lanes 1 and 3–20), without p43 (lane 2), or with p43 pre-treated with protease (lane 3), followed by UV irradiation to induce cross-linking (except lane 4). Samples were denatured and analyzed on 4–20% acrylamide gels. Samples run in lanes 5, 7, 9, 11 and 6, 8, 10, 12 included 50-fold and 500-fold molar excess, respectively, of the indicated unlabeled competitor RNAs over labeled *E.aediculatus* telomerase RNA. Samples run in lanes 13, 15, 17, 19 and 14, 16, 18, 20 included 5- and 50-fold excess by mass, respectively, of the indicated homopolymers over labeled *E.aediculatus* telomerase RNA. Protein molecular weight markers are indicated on the right. XL, cross-linked RNA–protein complexes.

activity (lane 1) was almost quantitatively captured on α -p43 antibody beads (lanes 3–5), but not on control beads (lanes 6–8). The activity retained on beads gave the characteristic banding pattern of telomerase observed in the partially purified nuclear extract (compare lanes 1 and 3) and was RNase sensitive as expected (lane 2 and data not shown). Telomerase activity from as little as 0.25% of control flowthrough was readily detected (lane 13), while an equivalent level of activity is observed in 2.5% of the flowthrough from α -p43 beads (lane 11), indicating that immunodepletion was at least 10-fold. Similar results were obtained with unpurified nuclear extract (data not shown). These results indicate that >90% of telomerase activity is associated with p43 and therefore suggest that p43 is an authentic component of the functional *E.aediculatus* telomerase complex.

Purified p43 binds telomerase RNA *in vitro*

p43's La homology suggests that its association with telomerase is mediated at least in part by binding directly to the RNA subunit. This prediction was addressed by *in vitro* experiments in which binding between *in vitro* transcribed photoaffinity-labeled *E.aediculatus* telomerase RNA and purified recombinant p43 was determined by formation of a covalent cross-link upon light excitation (Figure 8). Retarded complexes were observed when ^{32}P -labeled telomerase RNA, in which all uridine positions were substituted with the chromophore 5-iodouridine (Willis *et al.*, 1993), was irradiated with long-wave UV light in the presence of p43 (lane 1). Presumably due to cross-linking at different positions, several covalent species are formed, and their apparent molecular masses (100–135 kDa) are consistent with the notion that they represent telomerase RNA (apparent molecular mass ~ 55 kDa) covalently attached to p43 (51 kDa). The cross-linking efficiency of 21% is within the range typically observed for high-affinity protein–RNA complexes (Meisenheimer and Koch, 1997). As expected, complexes were not detected in the absence of p43 (lane 2), without irradiation (lane 3) or upon proteinase K treatment (lane 4), and were largely competed by an excess of unlabeled *E.aediculatus* telomerase RNA (lanes 5 and 6).

Using a variety of unlabeled RNAs, these competition studies allowed us to begin elucidating the RNA specificity determinants of p43 binding. *Tetrahymena thermophila* telomerase RNA, featuring a secondary structure similar to its *Euplotes* homolog as well as a 3'-UUU_{OH} terminus but otherwise very little sequence conservation (Lingner *et al.*, 1994), is equally efficient in competing for p43 binding (lanes 9 and 10), while *T.thermophila* P4-P6, a structured RNA of similar size (Murphy and Cech, 1993), has no effect (lanes 11 and 12). A block mutant of *E.aediculatus* telomerase RNA in which the 3' terminal Us were replaced by Cs also competes well (lanes 7 and 8), but reproducibly slightly less efficiently than wild-type telomerase RNA [46% of uncompeted cross-linking efficiency (lane 1) with the wild-type competitor RNA (lane 5) versus 58% with the block mutant (lane 7), and 46 versus 57% in another experiment (data not shown)]. Finally, whereas poly(A) and poly(C) homopolymeric RNAs have virtually no effect on complex formation (lanes 13–16) and poly(G) has an intermediate effect (lanes 17 and 18), poly(U) strongly competes for p43 binding (lanes 19 and 20). In summary, these data indicate that p43 binds *E.aediculatus* telomerase RNA with some measure of specificity *in vitro*, and suggest that this protein recognizes not only runs of Us (as found at the 3' ends of ciliate telomerase RNAs), but also other structural elements of ciliate telomerase RNAs.

Discussion

In previous work, a direct biochemical approach led to the identification of p43 as a putative protein component of telomerase from *E.aediculatus* (Lingner and Cech, 1996). We have now cloned the p43 gene and provided biochemical evidence for p43 as an authentic component of an active telomerase complex. Sequence analysis and RNA binding data suggest that it is a member of the La family of proteins known to function in RNP maturation. Interestingly, this new telomerase component also appears to depend on non-canonical translation for its synthesis.

An unusual frameshift required for p43 synthesis

Characterization of the p43 gene revealed that production of full-length protein is likely to require a programmed

ribosomal frameshift, a translational mechanism that is extremely rare in cellular gene expression (for reviews see Farabaugh, 1996; Gesteland and Atkins, 1996). In other systems, frameshifting is employed to regulate expression levels of either a single protein product (such as *E.coli* peptide release factor 2 and mammalian ornithine decarboxylase antizyme) or of a downstream product that is only expressed when fused, in a frameshift-dependent manner, to an upstream product (such as the *pol* proteins of many retrotransposons and retroviruses). In both cases, frameshifting efficiency is feedback regulated by the cellular concentration of full-length protein and relies, at least in part, on stimulatory mRNA sequences. Frameshifting in p43 might play a similar regulatory function. Further investigation is therefore needed to determine whether the 88 amino acid polypeptide that would be expressed from the p43 gene in the absence of a frameshift is produced, and if so, how its function is related to that of full-length p43.

The net +1 change in reading frame occurs in the highly conserved La motif of p43 (Figures 3 and 4). Since the amino acid sequence predicted from a single +1 translational frameshift shows stronger homology to the La motif than that predicted from two successive –1 frameshifts (data not shown), and since –2 or +2 frameshifts have not been observed except in special test systems, the most plausible rationalization of our observations is a single +1 frameshift at either site 3 or site 4 (Figure 3). How does the sequence context of this +1 frameshift compare with that of other systems? The mechanism of programmed +1 frameshifting events appears to be a translational pause induced by slow decoding of the codon occupying the ribosomal aminoacyl-tRNA binding site (A-site), followed by forward slippage of the peptidyl-tRNA in the peptidyl-tRNA binding site (P-site) (Farabaugh and Bjork, 1999). The first step, ribosome pausing, is dependent on a rare sense codon or a stop codon within a poorly recognized context (Farabaugh, 1996 and references therein), and the second step, tRNA slippage, correlates with the extent to which the slipping A-site tRNA can re-pair with the codon in the +1 register after slippage (Curran, 1993; Sundararajan *et al.*, 1999). Inspection of the frameshift region of the p43 message (Figure 3) suggests a similar mechanism: the codon just upstream of the stop codon in reading frame zero is 5'-AAA-3' within a slippery AU-rich stretch of nucleotides (Figure 3); a tRNA^{Lys}_{uuu} bound to this codon in the P-site may slip by +1 due to slow recognition of the stop codon in the A-site, to base pair with the near-cognate codon 5'-AAU-3'. tRNA slippage at this codon (corresponding to site 4 in Figure 3) would allow formation of two canonical A–U Watson–Crick base pairs, while the U–U interaction at the wobble position may be stabilized by a uridine modification at this site in the tRNA. Post-transcriptional modifications of the wobble uridine are observed in the vast majority of tRNAs^{Lys}_{uuu} analyzed to date (Sprinzl *et al.*, 1998). Clearly, characterization of the actual frameshift mechanism in p43 translation requires further investigation.

Curiously, in the short list of genes employing translational frameshifting in their expression, those involved in telomere metabolism are prominently represented. The non-LTR retrotransposon *HeT-A*, which functions together with another transposable genetic

element (*TART*) to maintain telomeres in the telomerase-independent pathway of *Drosophila* (reviewed in Pardue *et al.*, 1996a), relies on a –1 frameshift for production of its *gag*-like protein (Pardue *et al.*, 1996b). In addition, yeast Est3p, a subunit of yeast telomerase (Hughes *et al.*, 2000), features a +1 programmed frameshift, which mechanistically appears to resemble that of the yeast *Ty* retrotransposons (Morris and Lundblad, 1997). The telomerase component p43 now joins this set of proteins. Despite the absence of obvious sequence or functional homology between these proteins, the frameshift requirement common to p43, Est3p and *HeT-A* may point to a link between these two types of telomere maintenance. In the light of the evolutionary relationship between the TERTs and non-LTR retrotransposon RTs (Eickbush, 1997; Nakamura *et al.*, 1997; Nakamura and Cech, 1998), this idea seems particularly intriguing.

Relationship of p43 to La proteins

Clues to the function of p43 in the telomerase RNP come from its sequence similarity to the La autoantigen. p43 has a convincing La motif at its N-terminus (Figure 4, top) and the typical highly charged C-terminus (Figure 2B), but the presence of a canonical RRM is less certain (Figure 4, bottom). Putative RRMs from a number of proteins, including La proteins, appear to lack critical residues in the RNP1 and RNP2 submotifs, and some of these RRMs, e.g. those of the polypyrimidine tract-binding protein (Piñol-Roma *et al.*, 1989), have thus been hypothesized to represent a new subclass of RRMs (Ghetti *et al.*, 1992). p43 may contain such a highly diverged atypical RRM, whose functionality remains to be proven experimentally.

Database searches have recently identified a family of related La-motif-containing proteins in *S.cerevisiae*, *Caenorhabditis elegans*, *Mus musculus* and *H.sapiens*. These proteins are distinct from the authentic La protein in each organism because their La motifs are not located at the N-terminus and because they are also otherwise unrelated to La proteins (Sobel and Wolin, 1999 and references therein). While most of these proteins await further characterization, the two members of this family that are known in yeast, Sro9p and Slf1p, have been shown to be functionally non-redundant with Lhp1p (the yeast La), and based on their RNA binding *in vitro* and association with polyribosomes *in vivo*, they have been suggested to function in mRNA translation (Sobel and Wolin, 1999). *Euplotes* p43, however, is more closely related to authentic La proteins than to these La-motif-containing proteins, as judged by sequence analysis (data not shown), especially given the N-terminal location of its La motif.

Authentic La proteins have been implicated to function in three pathways: (i) mRNA translation, including cap-independent mRNA translation, translational initiation and mRNA stabilization; (ii) termination and reinitiation of RNA pol III transcription; and (iii) RNP maturation, specifically nuclear retention of pol III transcripts and RNP assembly.

La proteins bind to newly synthesized pol III transcripts by recognizing the hallmark sequence UUU_{OH} that is present at the 3' end of these RNAs (Stefano, 1984), which include precursors to tRNAs, 5S ribosomal RNA, U6 snRNA, signal recognition particle (SRP) RNA, as well as

the poorly understood cytoplasmic Y RNAs. La has also been shown to bind to internal sequences and/or structures (Chang *et al.*, 1994; Grimm *et al.*, 1997) *in vitro*. Telomerase RNAs of ciliates are also pol III transcripts, and in particular, the *E.aediculatus* telomerase RNA has been shown to end in a stretch of 3–4 uridylyte residues (Lingner *et al.*, 1994). In our *in vitro* binding competition experiments (Figure 8), homopolymeric RNAs showed marked differences in their efficiency to compete with telomerase RNA for binding to p43, with $U > G \gg A \approx C$. This trend is similar to that of human La protein, where the hierarchy of binding stabilities was determined to be $U > G > A > C$ (Stefano, 1984). Thus, both structural homology and RNA binding experiments support p43 being the *Euplotes* La protein. However, we also find that p43 recognizes internal features of telomerase RNA that are not present in highly structured P4-P6 RNA and tRNA. This leaves open the possibility that p43 shares some characteristics of La proteins, but has a more specialized, perhaps telomerase-specific, function.

Possible roles for p43 in telomerase function

Involvement of La in nuclear retention of pol III transcripts is indicated because deletion of the 3'-terminal oligouridylyte stretch of *Xenopus* U6 snRNA results in partial loss of nuclear retention of this RNA (Boelens *et al.*, 1995), and an analogous mutation in human Y RNA stimulates export of this RNA to the cytoplasm (Simons *et al.*, 1996). The La binding sites of pre-tRNAs are removed prior to nuclear export (reviewed in Deutscher, 1990). The cytoplasmic fractions of Y RNAs (Simons *et al.*, 1994) and 5S rRNA (Guddat *et al.*, 1990) are not bound by La, apparently due to nucleolytic removal of their 3'-terminal UUU_{OH}. These observations have led to the hypothesis that loss of La binding is a requirement for nuclear export of these RNAs. This idea was later confirmed by the finding that an *in vivo* selection of a combinatorial library of RNAs for their ability to be imported into the nucleus yielded a class of RNAs dependent on La binding for both nuclear retention and nuclear import (Grimm *et al.*, 1997).

Telomerase is an enzyme that ultimately functions in the nucleus. Consistent with the observation that, at least in *Tetrahymena*, telomerase RNA present in active telomerase retains its 3'-terminal oligouridylyte stretch (Greider and Blackburn, 1989), the *Euplotes* telomerase RNP may utilize association with p43 to secure its nuclear localization. If p43 is the *Euplotes* La protein, the function of La as a nuclear retention signal for RNAs is now extended to include a mature active RNP.

Studies on the nuclear retention of *Xenopus* U6 snRNA led Boelens *et al.* (1995) to propose that La might primarily act as a molecular chaperone in RNP formation, with nuclear retention of U6 snRNA resulting from the requirement of La binding for efficient RNP formation in the nucleus (Boelens *et al.*, 1995). Indeed, the *S.cerevisiae* La protein Lhp1p was shown to stabilize both newly synthesized U6 snRNA and, more recently, U4 snRNA, to facilitate the assembly of these RNAs into a functional snRNP (Pannone *et al.*, 1998; Xue *et al.*, 2000). In addition, Lhp1p and the *S.pombe* La protein Sla1p appear to stabilize pre-tRNAs in conformations that allow them to enter their proper maturation pathway (van Horn *et al.*,

1997; Yoo and Wolin, 1997). These results suggest that La may perform the function of a general molecular chaperone for a variety of small structural RNAs. Therefore, p43 might play a similar role in the *Euplotes* telomerase RNP. Since p43 is associated with >95% of the telomerase RNA molecules in the cell (Figure 6), this protein may facilitate binding of the catalytic subunit to the RNA by promoting productive conformations of the RNA subunit. The presence of p43 in the active telomerase complex (Figure 7) suggests that continual or periodic stabilization of the RNA by p43 may be required for telomerase function even after RNP assembly. Similarly, active human telomerase assembled *in vitro* appears to remain associated with the molecular chaperones p23 and Hsp90 (Holt *et al.*, 1999).

While *S.cerevisiae* telomerase RNA has the 5'-trimethylguanosine cap and Sm site that indicate a similarity of the maturation of this telomerase complex to that of snRNPs (Seto *et al.*, 1999), human telomerase RNA has a 3' domain characteristic of small nucleolar RNAs (snoRNAs), suggesting a nucleolar stage (Mitchell *et al.*, 1999). The presence of an La motif protein in *E.aediculatus* telomerase points to yet another mechanism for telomerase RNP biogenesis. It remains to be determined whether the other telomerase RNPs that contain a pol III transcript are also associated with a La-related protein.

Materials and methods

Purification of *Euplotes* telomerase

For cloning of its protein components, telomerase was purified from *E.aediculatus* nuclear extracts as described (Lingner and Cech, 1996) using heparin chromatography and affinity chromatography followed by glycerol gradient centrifugation. Throughout, telomerase was quantitated by measuring its RNA component and activity. To prepare partially purified telomerase for immunoprecipitation experiments, nuclear extract was subjected to glycerol gradient centrifugation (Hammond *et al.*, 1997).

Cloning of the p43 gene

Polypeptides from purified telomerase were separated by SDS-PAGE. The bands corresponding to the p43 doublet were excised from the Coomassie Blue-stained gel and in-gel digested with trypsin as described (Shevchenko *et al.*, 1996). Tryptic peptides from the unfractionated digest were sequenced *de novo* by nanoelectrospray tandem mass spectrometry as described (Wilm *et al.*, 1996; Shevchenko *et al.*, 1997) on a triple quadrupole mass spectrometer API III (PE SCIEX, Ontario, Canada). Peptides were in turn isolated by the first quadrupole of the mass spectrometer, fragmented in the collision cell and spectra of their product ions were acquired. Peptide sequences were determined by considering the mass differences between adjacent fragment ions containing the peptide C-terminus (y -ions). y -ions were identified by comparison of the tandem mass spectra of native peptides and peptides esterified with methanol (data not shown).

Full sequences were obtained for the following tryptic peptides: T₁, DGAELLK; T₂, FLCEVCK; T₃, SSLDELKAEK; T₄, FVDSSVTGLNNK; T₅, QEEELSQSCFLK; T₆, LLSLPLSLVK. L signifies leucine or isoleucine, two isobaric amino acids that cannot be distinguished by our methods. The following discrepancies were found between the sequence determined by mass spectrometry and that deduced from the DNA sequence: the aspartate in peptide T₁ was an asparagine, the order of the adjacent amino acids valine and cysteine in peptide T₂ was reversed, and the glutamine at the N-terminus of peptide T₅ was a lysine. Peptide T₆ could not be assigned to p43 and may originate from a contaminating protein or a substoichiometric telomerase component. The following tryptic peptides were sequenced partially and retrospectively matched to the full-length sequence of p43: t_1 , SPSLTIPK; t_2 , EHYGENYLILK; t_3 , SLLKEHYGENYLILK; t_4 , TAEELIAFTIR; t_5 , EALDCDRPSYIQYDDKETAIVIR; t_6 , SFVNKYQYQSLIEIDEATAQK.

Degenerate oligonucleotides were designed as follows. Primer 5'-NNNGTNACHGGHATHAAYAA-3' corresponds to a part of peptide T₄ (forward direction) and primer 5'-DGCDGTYTCYTRGTRCRRRTA-3' corresponds to a part of peptide t₅ (reverse direction), where R = A/G, Y = C/T, H = A/C/T, D = A/G/T and N = A/C/G/T (equimolar mixtures in all cases). A PCR reaction with these primers and genomic DNA as the template yielded a 0.3 kb fragment. Direct sequencing of this PCR product revealed its correspondence to a part of the p43 gene since it could encode for peptide T₅. The PCR product was random primed and used as a probe to screen an *E.aediculatus* genomic DNA library. This library was constructed by subcloning mung bean nuclease-treated *E.aediculatus* genomic DNA into the *Sma*I site of plasmid pPCR-Script SK(+) (Stratagene) as described (Lingner et al., 1997a). The macro-nuclear gene encoding p43 was identified by colony hybridization and sequenced. The DDBJ/EMBL/GenBank accession No. for the p43 gene sequence is AF307939.

Expression and purification of recombinant p43

His₆-tagged p43 was overexpressed in *E.coli* BL21(DE3) cells from plasmid pET15b p43 by induction with isopropyl-β-D-thiogalactopyranoside (IPTG) at 0.5 mM and growth at 37°C. Under these conditions, p43 is sequestered in inclusion bodies. Recombinant protein was solubilized in buffer containing 8 M urea and purified on Ni²⁺-NTA resin (Qiagen) under denaturing conditions, followed by preparative SDS-PAGE. The excised gel slice was used to immunize rabbits (Macromolecular Resources, Colorado State University, Fort Collins, CO). Soluble recombinant p43 was produced in *E.coli* BL21(DE3) cells by induction with IPTG at 0.1 mM and growth at 25°C, and purified under native conditions on Ni²⁺-NTA agarose. For photo-cross-linking experiments, p43 was further purified to >95% homogeneity (by Coomassie gel) on heparin-agarose and poly(U)-Sepharose as described (Yoo and Wolin, 1997).

Immunoprecipitations

Immunoprecipitations were performed by incubating 20 μl of nuclear extract or partially purified telomerase with 10 μl (packed volume) of α-p43 or control antibody beads in a total volume of 200 μl in nuclear extract buffer [400 mM potassium glutamate, 50 mM Tris-acetate pH 8.0, 10 mM MgCl₂, 10% glycerol, 0.1% NP-40, 1 mM dithiothreitol (DTT)] for 2 h at 4°C with agitation. Supernatants were separated from beads by centrifugation and subjected to two additional rounds of immunodepletion with fresh antibody beads. Beads were washed extensively in nuclear extract buffer prior to further analysis by northern blotting or telomerase assay.

Photo-cross-linking assays

Binding reactions contained 2 × 10⁵ c.p.m. (~6.6 nM) of iodouridine-substituted RNA, 330 nM purified recombinant p43 (by Bradford assay; see Figure 5B for protein purity), 100 ng/μl yeast tRNA and 100 ng/μl bovine serum albumin in 250 mM sucrose, 50 mM NaCl, 25 mM KCl, 5 mM MgCl₂, 3 mM DTT, 50 mM Tris-HCl pH 7.5, in the presence of unlabeled *in vitro* transcribed competitor RNAs or homoribopolymers (Sigma) where indicated. Telomerase is fully active in this buffer (data not shown). After incubation on ice for 10 min, RNA-protein complexes were cross-linked by irradiation for 15 min in a Stratalinker (Stratagene) fitted with 312 nm bulbs, boiled in SDS-protein loading buffer, run in 4–20% reducing Tris/glycine gels and quantitated using a PhosphorImager.

Supplementary data

Supplementary data for this paper are available at *The EMBO Journal* Online.

Acknowledgements

We thank Tracy Bryan, Dan Harrington and Anita Seto for help with experiments, Sandy Wolin, Jim Goodrich, Ravinder Singh, Phil Farabaugh and John Atkins for helpful discussions, and Tracy Bryan and Kathy Friedman for critically reading the manuscript. This work was supported by the Howard Hughes Medical Institute and NIH grant GM28039. We thank the W.M.Keck foundation for support of cellular and molecular structure research on the Boulder campus.

References

- Bandziulis,R.J., Swanson,M.S. and Dreyfuss,G. (1989) RNA-binding proteins as developmental regulators. *Genes Dev.*, **3**, 431–437.
- Beattie,T.L., Zhou,W., Robinson,M.O. and Harrington,L. (1998) Reconstitution of human telomerase activity *in vitro*. *Curr. Biol.*, **8**, 177–180.
- Birney,E., Kumar,S. and Krainer,A.R. (1993) Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Res.*, **21**, 5803–5816.
- Blasco,M.A., Lee,H.-W., Hande,M.P., Samper,E., Lansdorf,P.M., DePinho,R.E. and Greider,C.W. (1997) Telomere shortening and tumor formation by mouse cells lacking telomerase RNA. *Cell*, **91**, 25–34.
- Boelens,W.C., Palacios,I. and Mattaj,I.W. (1995) Nuclear retention of RNA as a mechanism for localization. *RNA*, **1**, 273–283.
- Bryan,T.M. and Cech,T.R. (1999) Telomerase and the maintenance of chromosome ends. *Curr. Opin. Cell Biol.*, **11**, 318–324.
- Bryan,T.M., Goodrich,K.J. and Cech,T.R. (2000) A mutant of *Tetrahymena* telomerase reverse transcriptase with increased processivity. *J. Biol. Chem.*, **275**, 24199–24207.
- Burd,C.G. and Dreyfuss,G. (1994) Conserved structures and diversity of functions of RNA-binding proteins. *Science*, **265**, 615–621.
- Chambers,J.C., Kenan,D., Martin,B.J. and Keene,J.D. (1988) Genomic structure and amino acid sequence domains of the human La autoantigen. *J. Biol. Chem.*, **263**, 18043–18051.
- Chang,Y.-N., Kenan,D.J., Keene,J.D., Gatignol,A. and Jeang,K.-T. (1994) Direct interactions between autoantigen La and human immunodeficiency virus leader RNA. *J. Virol.*, **68**, 7008–7020.
- Chapon,C., Cech,T.R. and Zaug,A.J. (1997) Polyadenylation of telomerase RNA in budding yeast. *RNA*, **3**, 1337–1351.
- Chen,J.-L., Blasco,M.A. and Greider,C.W. (2000) Secondary structure of vertebrate telomerase RNA. *Cell*, **100**, 503–514.
- Collins,K. and Gandhi,L. (1998) The reverse transcriptase component of the *Tetrahymena* telomerase ribonucleoprotein complex. *Proc. Natl Acad. Sci. USA*, **95**, 8485–8490.
- Collins,K., Kobayashi,R. and Greider,C.W. (1995) Purification of *Tetrahymena* telomerase and cloning of genes encoding the two protein components of the enzyme. *Cell*, **81**, 677–686.
- Curran,J.F. (1993) Analysis of effects of tRNA:message stability on frameshift frequency at the *Escherichia coli* RF2 programmed frameshift site. *Nucleic Acids Res.*, **21**, 1837–1843.
- Deutscher,M.P. (1990) Ribonucleases, tRNA nucleotidyltransferase and the 3' processing of tRNA. *Prog. Nucleic Acid Res. Mol. Biol.*, **39**, 209–240.
- Eickbush,T.H. (1997) Telomerase and retrotransposons: which came first? *Science*, **277**, 911–912.
- Evans,S.K. and Lundblad,V. (1999) Est1 and Cdc13 as comediators of telomerase access. *Science*, **286**, 117–120.
- Farabaugh,P.J. (1996) Programmed translational frameshifting. *Annu. Rev. Genet.*, **30**, 507–528.
- Farabaugh,P.J. and Bjork,G.R. (1999) How translational accuracy influences reading frame maintenance. *EMBO J.*, **18**, 1427–1434.
- Gandhi,L. and Collins,K. (1998) Interaction of recombinant *Tetrahymena* telomerase proteins p80 and p95 with telomerase RNA and telomeric DNA substrates. *Genes Dev.*, **12**, 721–733.
- Gesteland,R.F. and Atkins,J.F. (1996) Recoding: dynamic reprogramming of translation. *Annu. Rev. Biochem.*, **65**, 741–768.
- Ghetti,A., Piñol-Roma,S., Michael,W.M., Morandi,C. and Dreyfuss,G. (1992) hnRNP I, the polypyrimidine tract-binding protein: distinct nuclear localization and association with hnRNAs. *Nucleic Acids Res.*, **20**, 3671–3678.
- Greider,C.W. and Blackburn,E.H. (1987) The telomere terminal transferase of *Tetrahymena* is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell*, **51**, 887–898.
- Greider,C.W. and Blackburn,E.H. (1989) A telomeric sequence in the RNA of *Tetrahymena* telomerase required for telomere repeat synthesis. *Nature*, **337**, 331–337.
- Grimm,C., Lund,E. and Dahlberg,J.E. (1997) *In vivo* selection of RNAs that localize in the nucleus. *EMBO J.*, **16**, 793–806.
- Guddat,U., Bakken,A.H. and Pieler,T. (1990) Protein-mediated nuclear export of RNA: 5S rRNA containing small RNPs in *Xenopus* oocytes. *Cell*, **60**, 619–628.
- Hammond,P.W., Lively,T.N. and Cech,T.R. (1997) The anchor site of telomerase from *Euplotes aediculatus* revealed by photo-cross-linking

- to single- and double-stranded DNA primers. *Mol. Cell. Biol.*, **17**, 296–308.
- Harrington,L., McPhail,T., Mar,V., Zhou,W., Oulton,R., Amgen EST Program, Bass,M.B., Arruda,I. and Robinson,M.O. (1997) A mammalian telomerase-associated protein. *Science*, **275**, 973–977.
- Hinkley,C.S., Blasco,M.A., Funk,W.D., Feng,J., Villeponteau,B., Greider,C.W. and Herr,W. (1998) The mouse telomerase RNA 5'-end lies just upstream of the telomerase template sequence. *Nucleic Acids Res.*, **26**, 532–536.
- Hoffman,D.C., Anderson,R.C., DuBois,M.L. and Prescott,D.M. (1995) Macronuclear gene-sized molecules of hypotrichs. *Nucleic Acids Res.*, **23**, 1279–1283.
- Holt,S.E. *et al.* (1999) Functional requirement of p23 and Hsp90 in telomerase complexes. *Genes Dev.*, **13**, 817–826.
- Hughes,T.R., Evans,S.K., Weilbaecher,R.G. and Lundblad,V. (2000) The Est3 protein is a subunit of yeast telomerase. *Curr. Biol.*, **10**, 809–812.
- Kenan,D.J., Query,C.C. and Keene,J.D. (1991) RNA recognition: towards identifying determinants of specificity. *Trends Biochem. Sci.*, **16**, 214–220.
- Klobutcher,L.A., Gyax,S.E., Podoloff,J.D., Vermeesch,J.R., Price,C.M., Tebeau,C.M. and Jahn,C.L. (1998) Conserved DNA sequences adjacent to chromosome fragmentation and telomere addition sites in *Euplotes crassus*. *Nucleic Acids Res.*, **26**, 4230–4240.
- Lendvay,T.S., Morris,D.K., Sah,J., Balasubramanian,B. and Lundblad,V. (1996) Senescence mutants of *Saccharomyces cerevisiae* with a defect in telomere replication identify three additional *EST* genes. *Genetics*, **144**, 1399–1412.
- Licht,J.D. and Collins,K. (1999) Telomerase RNA function in recombinant *Tetrahymena* telomerase. *Genes Dev.*, **13**, 1116–1125.
- Lingner,J. and Cech,T.R. (1996) Purification of telomerase from *Euplotes aediculatus*: Requirement of a primer 3' overhang. *Proc. Natl Acad. Sci. USA*, **93**, 10712–10717.
- Lingner,J., Hendrick,L.L. and Cech,T.R. (1994) Telomerase RNAs of different ciliates have a common secondary structure and a permuted template. *Genes Dev.*, **8**, 1984–1998.
- Lingner,J., Hughes,T.R., Shevchenko,A., Mann,M., Lundblad,V. and Cech,T.R. (1997a) Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science*, **276**, 561–567.
- Lingner,J., Cech,T.R., Hughes,T.R. and Lundblad,V. (1997b) Three Ever Shorter Telomere (EST) genes are dispensable for *in vitro* yeast telomerase activity. *Proc. Natl Acad. Sci. USA*, **94**, 11190–11195.
- Lundblad,V. and Szostak,J.W. (1989) A mutant with a defect in telomere elongation leads to senescence in yeast. *Cell*, **57**, 633–643.
- Meisenheimer,K.M. and Koch,T.H. (1997) Photocross-linking of nucleic acids to associated proteins. *Crit. Rev. Biochem. Mol. Biol.*, **32**, 101–140.
- Mitchell,J.R., Cheng,J. and Collins,K. (1999) A box H/ACA small nucleolar RNA-like domain at the human telomerase RNA 3' end. *Mol. Cell. Biol.*, **19**, 567–576.
- Morris,D.K. and Lundblad,V. (1997) Programmed translational frameshifting in a gene required for yeast telomere replication. *Curr. Biol.*, **7**, 969–976.
- Murphy,F.L. and Cech,T.R. (1993) An independently folding domain of RNA tertiary structure within the *Tetrahymena* ribozyme. *Biochemistry*, **32**, 5291–5300.
- Nakamura,T.M. and Cech,T.R. (1998) Reversing time: origin of telomerase. *Cell*, **92**, 587–590.
- Nakamura,T.M., Morin,G.B., Chapman,K.B., Weinrich,S.L., Andrews,W.H., Lingner,J., Harley,C.B. and Cech,T.R. (1997) Telomerase catalytic subunit homologs from fission yeast and human. *Science*, **277**, 955–959.
- Nakayama,J., Saito,M., Nakamura,H., Matsuura,A. and Ishikawa,F. (1997) TLP1: a gene encoding a protein component of mammalian telomerase is a novel member of WD repeats family. *Cell*, **88**, 1–20.
- Olovnikov,A.M. (1973) A theory of marginotomy. *J. Theor. Biol.*, **41**, 181–190.
- Pannone,B.K., Xue,D. and Wolin,S.L. (1998) A role for the yeast La protein in U6 snRNP assembly: evidence that the La protein is a molecular chaperone for RNA polymerase III transcripts. *EMBO J.*, **17**, 7442–7453.
- Pardue,M.L., Danilevskaya,O.N., Lowenhaupt,K., Slot,F. and Traverse,K.L. (1996a) *Drosophila* telomeres: new views on chromosome evolution. *Trends Genet.*, **12**, 48–52.
- Pardue,M.L., Danilevskaya,O.N., Lowenhaupt,K., Wong,J. and Erby,K. (1996b) The *gag* coding region of the *Drosophila* telomeric retrotransposon, *HeT-A*, has an internal frame shift and a length polymorphic region. *J. Mol. Evol.*, **43**, 572–583.
- Piñol-Roma,S., Swanson,M.S., Gall,J.G. and Dreyfuss,G. (1989) A novel heterogeneous nuclear RNP protein with a unique distribution on nascent transcripts. *J. Cell Biol.*, **109**, 2575–2587.
- Seto,A.G., Zaug,A.J., Sobel,S.G., Wolin,S.L. and Cech,T.R. (1999) *Saccharomyces cerevisiae* telomerase is an Sm small nuclear ribonucleoprotein particle. *Nature*, **401**, 177–180.
- Shevchenko,A., Wilm,M., Vorm,O. and Mann,M. (1996) Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Anal. Chem.*, **68**, 850–858.
- Shevchenko,A., Wilm,M. and Mann,M. (1997) Peptide sequencing by mass spectrometry for homology searches and cloning of genes. *J. Protein Chem.*, **16**, 481–490.
- Shippen-Lentz,D. and Blackburn,E.H. (1990) Functional evidence for an RNA template in telomerase. *Science*, **247**, 546–552.
- Simons,F.H., Pruijn,G.J. and van Venrooij,W.J. (1994) Analysis of the intracellular localization and assembly of Ro ribonucleoprotein particles by microinjection into *Xenopus laevis* oocytes. *J. Cell Biol.*, **125**, 981–988.
- Simons,F.H., Rutjes,S.A., van Venrooij,W.J. and Pruijn,G.J. (1996) The interactions with Ro60 and La differentially affect nuclear export of hY1 RNA. *RNA*, **2**, 264–273.
- Singer,M.S. and Gottschling,D.E. (1994) *TLC1*: template RNA component of *Saccharomyces cerevisiae* telomerase. *Science*, **266**, 404–409.
- Sobel,S.G. and Wolin,S.L. (1999) Two yeast La motif-containing proteins are RNA-binding proteins that associate with polyribosomes. *Mol. Biol. Cell*, **10**, 3849–3862.
- Sprinzl,M., Horn,C., Brown,M., Ioudovitch,A. and Steinberg,S. (1998) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.*, **26**, 148–153.
- Stefano,J.E. (1984) Purified lupus antigen La recognizes an oligouridylylate stretch common to the 3' termini of RNA polymerase III transcripts. *Cell*, **36**, 145–154.
- Sundararajan,A., Michaud,W.A., Qian,Q., Stahl,G. and Farabaugh,P.J. (1999) Near-cognate peptidyl-tRNAs promote +1 programmed translational frameshifting in yeast. *Mol. Cell*, **4**, 1005–1015.
- van Horn,D.J., Yoo,C.J., Xue,D., Shi,H. and Wolin,S.L. (1997) The La protein in *Schizosaccharomyces pombe*: a conserved yet dispensable phosphoprotein that functions in tRNA maturation. *RNA*, **3**, 1434–1443.
- Virta-Pearlman,V., Morris,D.K. and Lundblad,V. (1996) Est1 has the properties of a single-stranded telomere end-binding protein. *Genes Dev.*, **10**, 3094–3104.
- Watson,J.D. (1972) Origin of concatemeric T7 DNA. *Nature New Biol.*, **239**, 197–201.
- Weinrich,S.L. *et al.* (1997) Reconstitution of human telomerase with the template RNA component hTR and the catalytic protein subunit hTERT. *Nature Genet.*, **17**, 498–502.
- Willis,M.C., Hicke,B.J., Uhlenbeck,O.C., Cech,T.R. and Koch,T.H. (1993) Photocrosslinking of 5-iodouracil-substituted RNA and DNA to proteins. *Science*, **262**, 1255–1257.
- Wilm,M., Shevchenko,A., Houthaeve,T., Breit,S., Schweigerer,L., Fotsis,T. and Mann,M. (1996) Femtomole sequencing of proteins from polyacrylamide gels by nanoelectrospray mass spectrometry. *Nature*, **379**, 466–469.
- Xue,D., Rubinson,D.A., Pannone,B.K., Yoo,C.J. and Wolin,S.L. (2000) U snRNP assembly in yeast involves the La protein. *EMBO J.*, **19**, 1650–1660.
- Yoo,C.J. and Wolin,S.L. (1994) La proteins from *Drosophila melanogaster* and *Saccharomyces cerevisiae*: a yeast homolog of the La autoantigen is dispensable for growth. *Mol. Cell. Biol.*, **14**, 5412–5424.
- Yoo,C.J. and Wolin,S.L. (1997) The yeast La protein is required for the 3' endonucleolytic cleavage that matures tRNA precursors. *Cell*, **89**, 393–402.
- Yu,G.-L., Bradley,J.D., Attardi,L.D. and Blackburn,E.H. (1990) *In vivo* alteration of telomere sequences and senescence caused by mutated *Tetrahymena* telomerase RNAs. *Nature*, **344**, 126–132.
- Zhou,J., Hidaka,K. and Fletcher,B. (2000) The Est1 subunit of yeast telomerase binds the Tlc1 telomerase RNA. *Mol. Cell. Biol.*, **20**, 1947–1955.

Received April 20, 2000; revised September 19, 2000;
accepted September 20, 2000