Reverse Transcriptase Motifs in the Catalytic Subunit of Telomerase

Joachim Lingner,* Timothy R. Hughes, Andrej Shevchenko, Matthias Mann,† Victoria Lundblad,† Thomas R. Cech†

Telomerase is a ribonucleoprotein enzyme essential for the replication of chromosome termini in most eukaryotes. Telomerase RNA components have been identified from many organisms, but no protein component has been demonstrated to catalyze telomeric DNA extension. Telomerase was purified from *Euplotes aediculatus*, a ciliated protozoan, and one of its proteins was partially sequenced by nanoelectrospray tandem mass spectrometry. Cloning and sequence analysis of the corresponding gene revealed that this 123-kilodalton protein (p123) contains reverse transcriptase motifs. A yeast (*Saccharomyces cerevisiae*) homolog was found and subsequently identified as *EST2* (ever shorter telomeres), deletion of which had independently been shown to produce telomere defects. Introduction of single amino acid substitutions within the reverse transcriptase motifs of Est2 protein led to telomere shortening and senescence in yeast, indicating that these motifs are important for catalysis of telomere elongation in vivo. In vitro telomeric DNA extension occurred with extracts from wild-type yeast but not from *est2* mutants or mutants deficient in telomerase RNA. Thus, the reverse transcriptase protein fold, previously known to be involved in retroviral replication and retrotransposition, is essential for normal chromosome telomere replication in diverse eukaryotes.

Replication of chromosome ends, or telomeres, requires specialized factors that are not essential for replication of internal chromosome sequences. Conventional DNA polymerases cannot fully replicate blunt-ended DNA molecules (*1*) or eukaryotic chromosomes (*2*), which contain 3′-terminal extensions. The key to end replication is telomerase, a ribonucleoprotein (RNP) enzyme that synthesizes the telomeric DNA repeats (*3*). The template for telomeric repeat synthesis is provided by the RNA subunit, which has been identified, cloned, and sequenced in ciliated protozoa (*4*, *5*), yeast (*6*, *7*), and mammals (*8*).

A telomerase RNP was first purified from *Tetrahymena* (*9*). Two protein components, p80 and p95, were specifically associated with the RNA subunit. Human, mouse, and rat homologs of *Tetrahymena* p80 have since been identified and found to be associated with telomerase (*10*). Although this evolutionary conservation suggests that p80 and p95 have important roles in telomere replication, their specific functions remain unclear. Neither protein has been reported to be essential for telomere synthesis, and neither has significant similarity to known polymerases or reverse transcriptases (*11*).

Telomerase RNP has also been purified from *Euplotes aediculatus*, a hypotrichous ciliate only distantly related to *Tetrahymena* (*12*). The hypotrichs present a special opportunity for telomere studies because their macronuclei contain millions of gene-sized DNA molecules. Each cell has about 8×10^7 telomeres (*13*) and about 3×10^5 molecules of telomerase (*12*). Measurements of the specific activity of telomerase throughout the purification indicated that the major activity present in macronuclear extracts was purified (*12*). The active telomerase complex had a molecular mass of ~230 kD, corresponding to a 66-kD RNA subunit and two proteins of 123 kD and ~43 kD (*12*). Photocross-linking experiments implicated the larger protein in specific binding of the telomeric DNA substrate (*14*).

Here we characterize the p123 component of *Euplotes* telomerase and show that it contains sequence hallmarks of reverse transcriptases. Furthermore, it is the homolog of a yeast protein, Est2p, shown previously to function in telomere maintenance. Our genetic and biochemical analyses show that the reverse transcriptase motifs of Est2p are essential for telomeric DNA synthesis in vivo and in vitro. We propose that telomerase, frequently called "a specialized reverse transcriptase," is in fact a reverse tran-

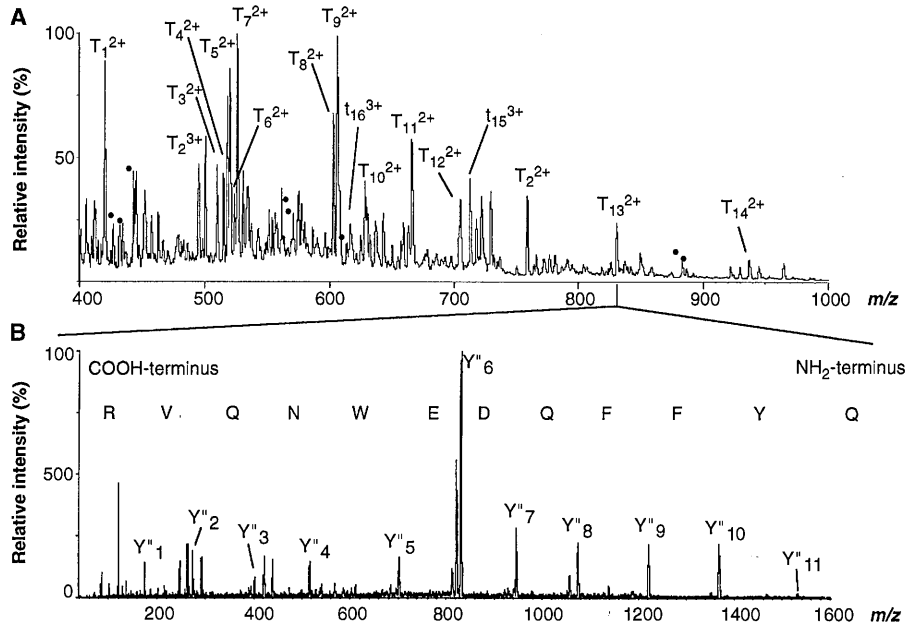


Fig. 1. Sequencing of the p123 subunit of telomerase by nanoelectrospray tandem mass spectrometry. (**A**) Mass spectrum of the unseparated peptide mixture. All peptides that were sequenced completely or partially are marked by the letter T or t, respectively (*15*). The eight peptide ions from which sequence tags were generated are marked by filled circles. Most unlabeled peaks correspond to trypsin autolysis products. (**B**) Tandem mass spectrum of the doubly charged precursor at the mass-to-charge ratio (*m/z*) of 830.4 in (A). Interpretation of the fragment ion mass in (B) and comparison with the esterified form of the peptide allowed the sequence assignment.

J. Lingner and T. R. Cech are at the Howard Hughes Medical Institute, Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO 80309–0215, USA. T. R. Hughes and V. Lundblad are in the Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA. A. Shevchenko and M. Mann are at the European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69012 Heidelberg, Germany.

*Present address: Swiss Institute for Experimental Cancer Research, 1066 Epalinges/VD, Switzerland.
†To whom correspondence should be addressed.

scriptase in terms of its catalytic active site.

Determination of *Euplotes* p123 sequence. The genes encoding the telomerase protein subunits from *E. aediculatus* were isolated by reverse genetics. Telomerase was purified and polypeptides were separated on SDS-polyacrylamide gels. Amino acid sequencing of the trypsin-digested p123 band was accomplished by nanoelectrospray tandem mass spectrometry (*15–17*), a miniaturized form of electrospray (*18*) that allows mass spectrometric interrogation of minute analyte volumes for extended periods of time due to its low flow rate. No chromatography is needed, because the unfractionated peptide mixture obtained after digestion of the protein in a gel slice is separated and sequenced in the spectrometer. For p123, 14 peptides were sequenced de novo (Fig. 1) (*15*).

Two of the peptide sequences were used to design degenerate polymerase chain reaction (PCR) primers (arrows in Fig. 2) to amplify a portion of the macronuclear gene encoding p123. A genomic library was prepared from macronuclear DNA and screened with this fragment to isolate the full-length gene (*19*). The p123 gene was found to be encoded by a 3279–base pair (bp) macronuclear chromosome containing an uninterrupted 1031–amino acid open reading frame. In a Southern (DNA) blot experiment the PCR fragment hybridized to a single macronuclear chromosome of ~3.3 kb (*20*). The open reading frame predicts a protein of 122,562 daltons, corresponding to the size estimated by SDS–polyacrylamide gel electrophoresis of purified protein [120 kD (*12*)]. More than 150 amino acids identified in the purified polypeptide by mass spectrometry could be assigned in the open reading frame (Fig. 2). This includes all 14 peptides that were completely sequenced. The tandem mass spectra of 10 additional peptides also matched the gene sequence through partial sequences or peptide sequence tags (*21*).

Reverse transcriptase motifs in *Euplotes* p123 and its yeast homolog Est2p. In a BLAST search of protein databases, *Euplotes* p123 was found to be most similar to *Saccharomyces cerevisiae* Est2p ($P = 7 \times 10^{-7}$) and to a group II intron–encoded reverse transcriptase from the cyanobacterium *Calothrix* ($P = 2 \times 10^{-4}$) (*22*, *23*). Yeast Est2p has a predicted molecular

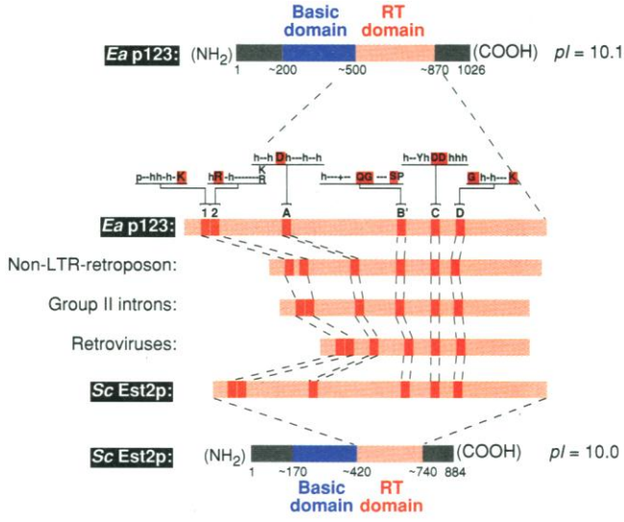


Fig. 3. Block diagrams of p123 and Est2p and comparison of the reverse transcriptase (RT) domains with those of other reverse transcriptases. The spacing of sequence motifs (red) is diagnostic for each reverse transcriptase family (*27*). In the consensus sequence, abbreviations are as in Fig. 2. The isoelectric point (*pI*) is the pH at which the protein has no net charge.

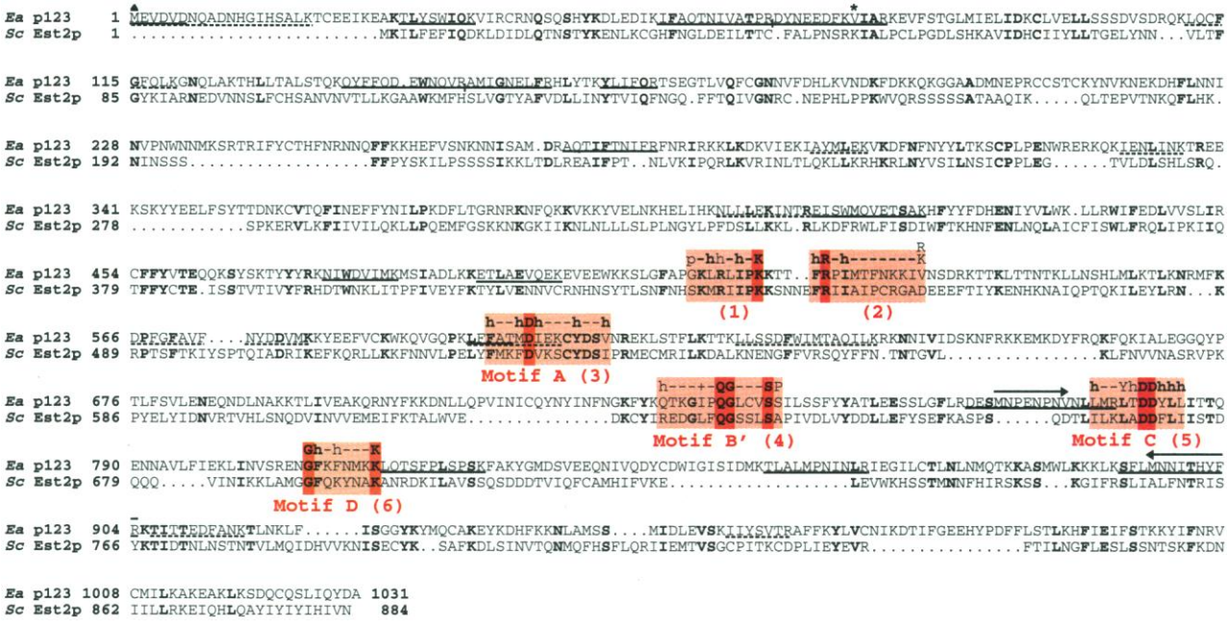
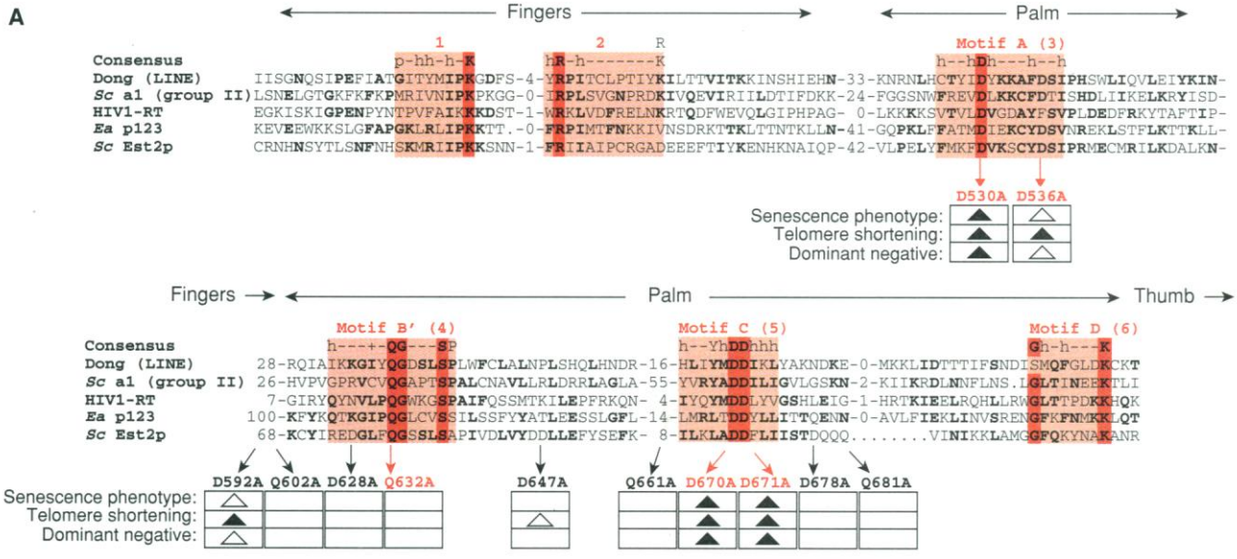


Fig. 2. Sequence alignment of *Euplotes* (*Ea*) p123 and yeast [*S. cerevisiae* (*Sc*)] Est2p (*50*). Identical amino acids are noted in boldface. The PCR primers used to amplify a portion of the gene are indicated by the arrows. Assigned reverse transcriptase motifs [designated by letters (*26*) or alternatively by numbers in parentheses (*27*)] are shown in orange, with the most highly conserved amino acids in red. In the consensus sequences of the motifs, h designates a hydrophobic amino acid, p a polar amino acid, and + a positively charged amino acid. The underlined sequences in p123 are the 14 peptides completely sequenced by nanoelectrospray tandem mass spectrometry. The dashed lines below the p123 sequence indicate another 10 peptides whose tandem mass spectra matched the sequence. One of the peptides contained an acetylated methionine (solid triangle) at its NH_2-terminus, indicating that it was the NH_2-terminal peptide of the protein. The nucleotide sequence of the *Euplotes* p123 gene has been deposited in GenBank (accession number U95964).

mass of 103 kD and, like p123, is very basic (Fig. 3). Although the overall sequence identity of *Euplotes* p123 and yeast Est2p is only 20% (Fig. 2), sequence similarity (correspondence of acidic, basic, hydrophobic, and hydrophilic amino acids) can be detected over the entire length of the two proteins.

The *EST2* (ever shorter telomeres) gene was one of four complementation groups identified by screening yeast mutants for reduction in telomere length and senescence (*24*, *25*). Epistasis analysis had indicated that the four *EST* genes function in the same pathway as *TLC1*, the gene encoding the telomerase RNA subunit (*6*), suggesting that the *EST* genes encode either components of the telomerase or positive regulators of its activity. The homology of yeast Est2p with *Euplotes* p123, the latter isolated because of its physical asso-



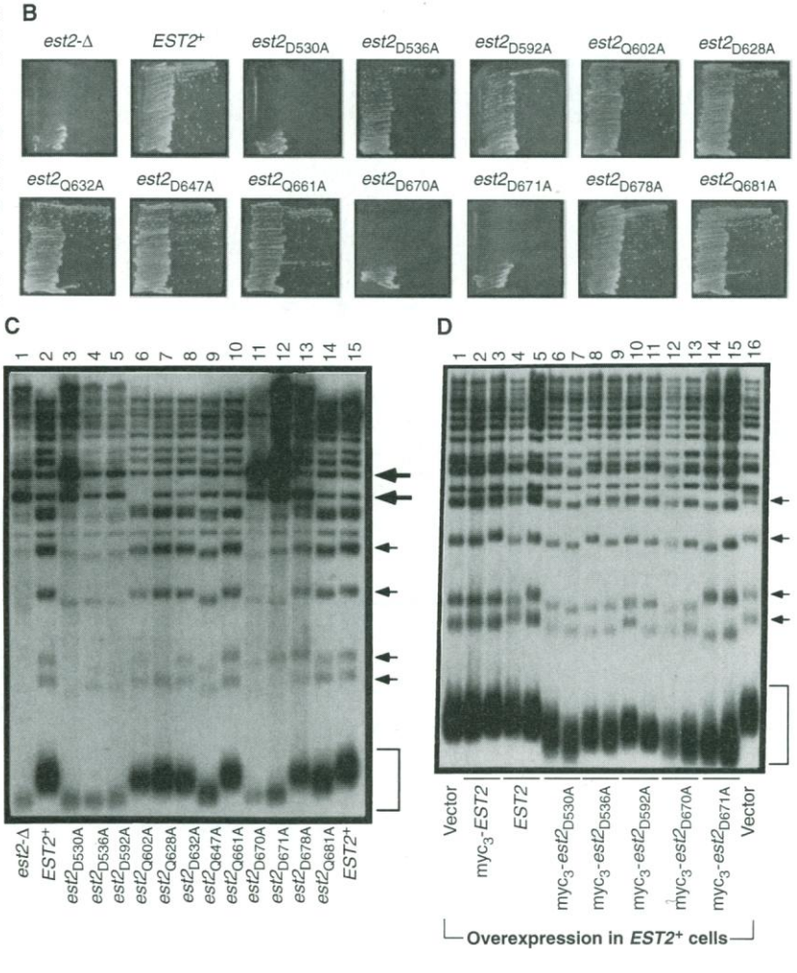


Fig. 4. In vivo functional analysis of the reverse transcriptase motifs in Est2p. (**A**) The 12 amino acids changed to alanines within the reverse transcriptase domain of Est2p are indicated by downward arrows (red, telomerase-conserved residues; black, nonconserved residues). The phenotypic effects of the mutations are indicated by solid triangles (strong mutant phenotype) and open triangles (weak mutant phenotype). The sequence alignment includes members of three other reverse transcriptase families (*27*). Boldface residues indicate identity of at least two sequences in the alignment. See (*50*) for amino acid abbreviations. (**B**) Senescence phenotype of *est2* mutants shown by spreading single colonies on plates (*51*). Photographs were taken after ~75 generations of growth. (**C**) Telomere length of *est2* mutants. Southern blot of genomic yeast DNA, hybridized with a telomere-specific probe (*24*). Single-copy plasmids carrying the wild-type *EST2* gene (lanes 2 and 15), the indicated *est2* mutant genes (lanes 3 to 14), or empty vector (lane 1) were transformed into an *est2*-Δ strain. Genomic DNA was prepared after ~75 generations of growth, at a time of maximal senescence for an *est2* null strain. The bracket and four small arrows indicate telomeric bands, and the two larger arrows indicate the subtelomeric repeat fragments that are amplified late in the growth of *est2* mutant strains (*24*, *28*). Five independent transformants of each missense mutant were assayed, one of which is shown. (**D**) Dominant-negative effect, resulting from overexpression of certain Est2p mutants. A Southern blot of genomic yeast DNA, prepared from a wild-type *EST2*$^+$ strain transformed with a high-copy plasmid expressing wild-type or the indicated mutant *est2* genes, was developed as in (C). In each case, the *EST2* promoter was replaced with the constitutive promoter of the alcohol dehydrogenase (*ADH*) gene. Lanes 1 and 16, empty vector. Two transformants are shown for each mutation after ~50 generations of growth. Additional growth resulted in further telomere shortening, although this additional length reduction is not sufficient to confer a senescence phenotype (*52*).

ciation with telomerase RNA and its copurification with telomerase activity, supported the proposal that both proteins are intrinsic subunits of their respective telomerases.

Euplotes p123 contains reverse transcriptase motifs, and the alignment reveals the presence of these motifs in a similar region of Est2p (Fig. 3). The primary sequences of reverse transcriptases are highly divergent: Only a few amino acids are absolutely conserved within separate short motifs (*26, 27*), but these motifs are believed to form a common tertiary fold. Both p123 and Est2p contain these key conserved amino acids, most notably the three invariant aspartates in motifs A and C, which are thought to be directly involved in catalysis (Fig. 2). Conserved motifs are spaced differently in the two major branches of reverse transcriptases, those encoded by retroviruses and long terminal repeat (LTR) retroposons and those encoded by non-LTR retroposons and group II introns (*27*). The spacing of sequence motifs in p123 and Est2p resembles that in the latter branch. However, the interval between motifs A and B′ in p123 and Est2p is unusually large (Fig. 3), suggesting that these two polypeptides may be members of a previously unknown subcategory.

Requirement of the reverse transcriptase motifs for Est2p function in vivo. The presence of reverse transcriptase motifs in both p123 and Est2p suggests that this region may define the catalytic active site of telomerase. To test the importance of these motifs for Est2p function, we used site-directed mutagenesis to change conserved and nonconserved aspartic acid (D) and glutamine (Q) residues in and around motifs A, B′, and C to alanine (A) (Fig. 4A). Each mutant, present on a single-copy *ARS CEN* plasmid, was tested for in vivo function in a complementation assay. Plasmids were transformed into the *est2*-Δ strain (Δ designates deletion), in parallel with either the empty vector or an *EST2*$^+$ plasmid. Transformants were assessed for the senescence phenotype (Fig. 4B) and for chromosome telomere length (Fig. 4C).

Consistent with the prediction that the reverse transcriptase motifs are required for Est2p function, mutation of any of the three conserved aspartates in motifs A and C prevented normal telomerase activity. Transformants expressing these mutant proteins became senescent and had shortened telomeric tracts, phenotypes indistinguishable from those of the null mutant (Fig. 4, B and C). Furthermore, a bypass pathway for telomere maintenance (*28*) was evident in these three mutant strains. Activation of this alternative pathway occurs as the result of a global amplification and rearrangement of both telomeric G-rich repeats and subtelomeric regions, and has only been observed in *est* and *tlc1* mutant strains with a severe telomere shortening phenotype (*24, 29*). A feature of this pathway is the amplification of two subtelomeric bands (Fig. 4C); these diagnostic restriction fragments were substantially amplified only in the *est2* null mutant and the three proposed active site mutants.

Mutations of amino acids other than the three most conserved aspartates had less severe or no phenotypic effects. The residue Asp^{536} of motif A is conserved between Est2p and p123, and the D536A mutation (Asp mutated to Ala at position 536) caused substantial telomere shortening and a modest senescence phenotype. Of the conserved residues tested, Gln^{632} of motif B′ was the only one that was functionally insensitive to replacement with alanine. However, this glutamine is not strictly conserved in reverse transcriptases (*27*), and when it is changed to alanine in human immunodeficiency virus–1 (HIV-1) reverse transcriptase, polymerase activity in vitro is reduced but not completely eliminated (*30*). In contrast to the phenotypes seen upon mutation of the semiconserved amino acids, mutation of six of the seven nonconserved amino acids tested showed little or no alteration of Est2p function.

Two observations indicate that stable Est2 protein was produced in the five *est* mutants with a diminished capacity to complement the *est2*-Δ strain. First, Myc_3-epitope-tagged versions of each mutant protein were visualized immunologically after immunoprecipitation (*31*). Second, overexpression of each of the five mutant alleles in a wild-type yeast strain with a functional chromosomal *EST2*$^+$ gene resulted in telomere shortening (Fig. 4D), whereas overexpression of the wild-type *EST2* gene had little effect. The dominant-negative phenotype shows that each mutant protein is being made and suggests that excess mutant Est2p can titrate components away from the wild-type telomerase complex.

Requirement of Est2p for telomerase activity in vitro. If Est2p is the catalytic protein subunit of telomerase, then telomerase activity should be abolished in *est2* mutant extracts. An in vitro assay was developed with extracts fractionated by glycerol gradient centrifugation (*32*). Telomerase-containing fractions were identified by detection of the RNA subunit on Northern blots (Fig. 5). Yeast telomerase sedimented as a 19*S* to 20*S* particle, substantially faster than the sedimentation of the deproteinized telomerase RNA (~17*S*). Telomerase-containing glycerol gradient fractions were pooled, concentrated, and tested for the ability to elongate a single-stranded telomeric oligonucleotide (Fig. 6). An activity was detected in wild-type extracts that had the characteristics of telomerase. It was dependent on the presence of oligonucleotide substrate and fractionated extract (Fig. 6A,

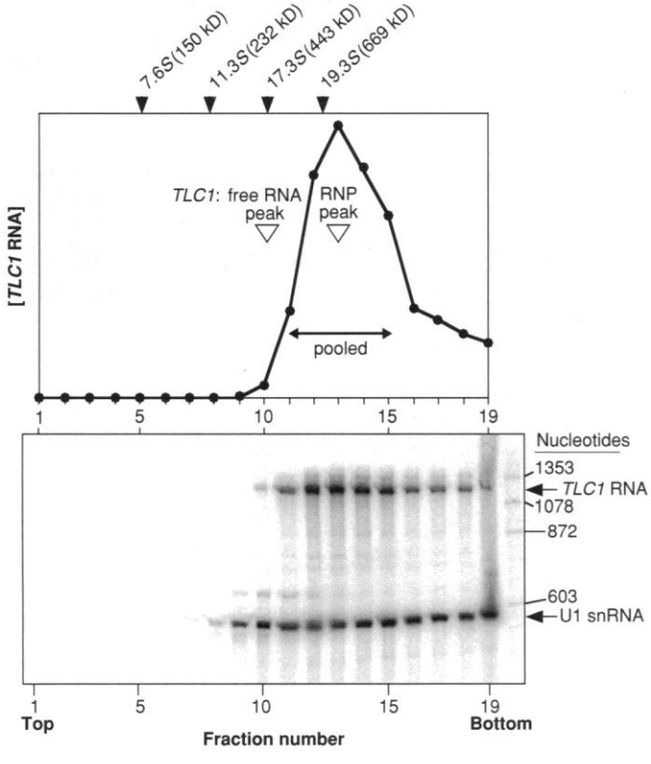


Fig. 5. Sedimentation of telomerase. Yeast extract was fractionated on a glycerol gradient (*32*), and telomerase RNA was detected by Northern blotting (bottom) and its concentration quantified on a PhosphoImager (top). Detection of U1 snRNP served as an internal control. Fractions pooled for activity assays are indicated. Telomerase RNP sedimented as a 19*S* to 20*S* particle, whereas deproteinized telomerase RNA sedimented at ~17*S*. The sedimentation value was determined relative to marker proteins that were run in parallel gradients and that had sedimentation coefficients of 7.6*S* (alcohol dehydrogenase), 11.3*S* (catalase), 17.3*S* (apoferritin), and 19.3*S* (thyroglobulin).

lanes 1 to 3). Addition of T and G residues occurred in an ordered manner consistent with the expected alignment of substrate and RNA template (Fig. 6A, lanes 5 and 6). The activity was sensitive to low concentrations of ribonuclease (RNase) A and was not stimulated by adenosine triphosphate (ATP) (Fig. 6B). These characteristics, in addition to the observed single round of extension of primer (Fig. 6A), are similar to those of the telomerase activity described by Blackburn and co-workers (*33*). A different activity described as telomerase by Lue and Wang (*34*) gives rise to long products and is stimulated by ATP. This latter activity was not detectable in our telomerase-containing glycerol gradient fractions.

A telomerase RNA template mutation that alters the specificity of nucleotide incorporation to produce a Hae III restriction site (6) provides an additional test for the authenticity of the in vitro telomerase assay. An extract of this *TLC1-1*(*HaeIII*) mutant, fractionated on a glycerol gradient, gave the predicted extension of the telomeric oligonucleotide only in the presence of deoxycytidine triphosphate (dCTP) (Fig. 6C, lanes 6 to 8), a nucleotide that has no effect on extension by a wild-type extract (Fig. 6C, lanes 2 and 3). This nucleotide specificity change supports the dependence of the assay on the *TLC1* RNA. Because the *TLC1-1*(*HaeIII*) strain also undergoes senescence (*29*), this result also provides confidence that telomerase activity can still be detected in senescing cells, as long as they are not subcultured too extensively.

We then assayed fractionated extracts from *est2*-Δ and *tlc1*-Δ strains for telomerase activity (Fig. 6D). As expected, no activity was detectable in *tlc1*-Δ yeast, which has the gene for telomerase RNA deleted. In the *est2*-Δ strain, telomerase RNA was still assembled into an RNP, as assessed by glycerol gradient centrifugation and Northern blotting (*32*), but telomerase activity was completely absent. This indicates that Est2p is essential for telomerase activity. As described above, the absence of activity is not simply a secondary consequence of senescence. We also measured telomerase activity in extracts from *est2*-Δ and strains expressing two of the proposed active site mutants in the presence of the chain-terminating analog ddGTP (Fig. 6E). According to the proposed primer-template alignment, extension should terminate after addition of two nucleotides. A practical advantage is the higher signal-to-noise ratio obtained when all products are concentrated in one or two bands. Again, activity was dependent on functional *TLC1* and *EST2* genes.

Telomerase structure. The presence of a reverse transcriptase domain in the catalytic subunit of telomerase provides a framework for exploring the structure and mechanism of this enzyme. Reverse transcriptases have been studied in great detail, and the three-dimensional structure of HIV-1 reverse transcriptase has been solved (*35*). The structure can be compared with a right hand with fingers, palm, and thumb, with the active site residing in the palm (*36*). A model for telomerase structure based on that of HIV-1 reverse transcriptase (HIV-1 RT) is shown in Fig. 7 with the telomerase RNA and a telomeric DNA substrate superimposed.

The catalytic subunit of telomerase has several features that distinguish it from other reverse transcriptases. Telomerase uses only a small portion of its RNA subunit as a template. The borders of this template must somehow be recognized. Furthermore, during processive synthesis of telomeric repeats the substrate translocates from one end of the template to the other by an as yet unknown mechanism. The large gap between motifs A and B′ of telomerase p123 and Est2p indicates an unusual finger domain structure. In HIV-1 RT this domain may be involved in template strand binding (*35*, *36*); whether and how it contributes to the unusual reaction mechanism of the telomerase RNP remain to be investigated. Finally, the telomerase protein is stably associated with its RNA subunit, as shown by our isolation of the *Euplotes* p123-RNA complex and by coimmunoprecipitation of the yeast RNA subunit with Est2p (*31*).

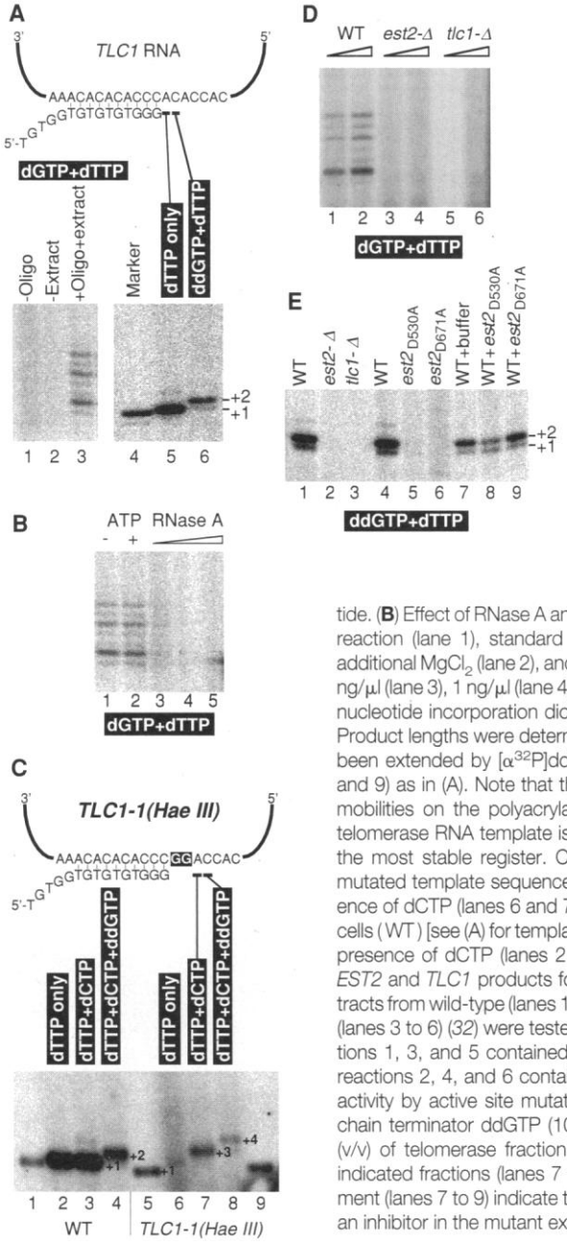


Fig. 6. In vitro functional analysis of reverse transcriptase motifs in Est2p. Telomerase was partially purified by glycerol gradient centrifugation and assayed for the ability to extend a telomeric DNA substrate (*32*). In the assay [α^{32}P]dTTP was included to visualize products elongated by 1, 2, 3, or 4 nucleotides (+1, +2, +3, or +4). (**A**) The telomerase RNA template region maximally base-paired to the DNA substrate is indicated schematically. Product lengths were determined relative to the same DNA substrate extended by one nucleotide at its 3′ end by reaction with [α^{33}P]ddTTP and terminal deoxynucleotidyl transferase (lane 4). Up to seven nucleotides were added in the presence of dGTP and dTTP (lane 3), one nucleotide in the presence of only dTTP (lane 5), and two nucleotides in the presence of dTTP and the chain-terminating analog ddGTP (lane 6). Oligo, oligonucleotide. (**B**) Effect of RNase A and ATP on telomerase activity. Standard reaction (lane 1), standard reaction plus 1 mM ATP and 1 mM additional $MgCl_2$ (lane 2), and standard reaction plus RNase A at 0.1 ng/μl (lane 3), 1 ng/μl (lane 4), and 10 ng/μl (lane 5). (**C**) Specificity of nucleotide incorporation dictated by the RNA template sequence. Product lengths were determined relative to DNA markers that had been extended by [α^{32}P]ddTTP (lane 1) or [α^{32}P]ddCTP (lanes 5 and 9) as in (A). Note that these two markers had slightly different mobilities on the polyacrylamide gel. The mutant *TLC1-1(HaeIII)* telomerase RNA template is indicated with the substrate bound in the most stable register. Consistent with this alignment and the mutated template sequence, efficient extension required the presence of dCTP (lanes 6 and 7). Telomerase in extract from *TLC1-wt* cells (WT) [see (A) for template sequence] was not influenced by the presence of dCTP (lanes 2 and 3). (**D**) Requirement of functional *EST2* and *TLC1* products for telomerase activity. Fractionated extracts from wild-type (lanes 1 and 2) and the indicated mutant strains (lanes 3 to 6) (*32*) were tested at two extract concentrations. Reactions 1, 3, and 5 contained 10% (v/v) of telomerase fraction, and reactions 2, 4, and 6 contained 20%. (**E**) Alleviation of telomerase activity by active site mutations in Est2p. All assays included the chain terminator ddGTP (100 μM). The reactions contained 10% (v/v) of telomerase fraction (lanes 1 to 6) or 5% of each of the indicated fractions (lanes 7 to 9). The results of the mixing experiment (lanes 7 to 9) indicate that the absence of activity is not due to an inhibitor in the mutant extracts.

This last feature distinguishes telomerase from the retroviral and LTR retroposon reverse transcriptases, but is similar to some mitochondrial and group II intron–encoded reverse transcriptases that also form complexes with their RNA templates (*37*).

Reverse transcriptase essential for chromosome replication in diverse eukaryotes. Reverse transcriptases have not previously been considered essential for normal cell physiology. Initially discovered as retroviral enzymes that catalyze the defining RNA-to-DNA step of retroviral replication (*38*), they were later found to mediate the transposition of DNA elements within eukaryotic genomes through an RNA intermediate (*39*). Reverse transcriptases are also present in some prokaryotes (*40*) and in *Neurospora* mitochondria (*41*), where they replicate genetic elements that are nonessential to their "host." Our discovery that a structurally related enzyme is essential for chromosome replication and cell division provides another example of the opportunism of nature: once a useful protein motif is stumbled upon, natural selection promotes its exploitation in diverse ways.

The evolutionary relationship between telomerase and the other reverse transcriptases is intriguing. It is well established that retroviruses acquired oncogenes such as v-*src*, v-*abl*, v-*ras*, and v-*fos* from cellular genomes. According to Temin's protovirus hypothesis, retroviruses also acquired their reverse transcriptase gene from normal cells, where the enzyme presumably contributed to some normal cellular process (*42*). Could this cellular source have been the telomerase p123/*EST2* gene, which mutated so that the protein product used an exogenous rather than an intrinsic RNA template? Alternatively, telomerase and the reverse transcriptases encoded by retrotransposons and retroviruses may all be descendants of an ancestral protein that emerged from an "RNA world" (*43*).

Telomere replication in the fruit fly *Drosophila* has been mysterious because this organism does not have short repeated telomeric sequences and presumably no telomerase. Rather, the non-LTR retroposons *HeT-A* and *TART* cap the chromosome ends (*44*). The *TART* reverse transcriptase is closely related to p123 and Est2p, which suggests that the *Drosophila* telomere replication machinery may in fact not be so different from that of other eukaryotes (*45*).

We have no satisfactory explanation for the lack of correspondence between the *Euplotes* and yeast p123/Est2p proteins and the *Tetrahymena* p80 or p95 protein (*9*). The small protein subunit of *Euplotes* telomerase (p43) also shows no similarity to the *Tetrahymena* proteins (*46*), and the complete yeast genome sequence does not reveal obvious p80 and p95 homologs. There are three possible explanations: (i) *Tetrahymena* may have a different telomerase in which p80 and p95 provide the active site (that is, telomerase was invented more than once in evolution). (ii) *Tetrahymena* may have two telomerases, one containing p80 and p95 and one (unisolated) containing a p123/Est2p homolog (for example, one telomerase for de novo telomere formation during macronuclear development and one for telomere replication). (iii) The *Tetrahymena* p80-p95-RNA complex may not be an active enzyme but may require a p123/Est2p subunit that was underrepresented upon purification of the particle.

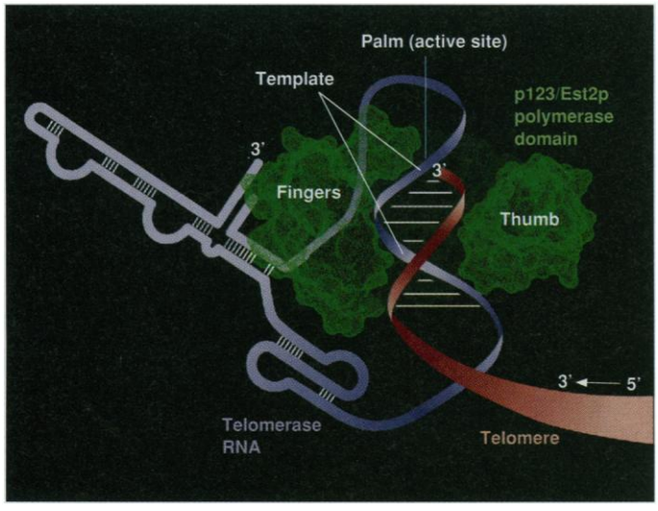


Fig. 7. Model of telomerase as an RNA–reverse transcriptase complex. The p123/Est2p subunit (green) is based on the right hand model of HIV-1 RT (*35*); thumb and fingers extend toward the reader. Motifs A, B′, C, and D are in the palm, and the active site aspartates are near the 3′ end of the telomeric DNA substrate (red). The RNA subunit (purple) has its template region in the palm; the location of the remainder of the RNA is unknown and is shown schematically in its secondary structure representation (*5*). Additional protein subunits may be associated (not shown). The telomeric DNA substrate is shown base-paired but not intertwined with the RNA subunit. The extent of base pairing and the sites of interaction of the nucleic acids with the protein are not known.

Mass spectrometric methods have recently become very successful for the identification of proteins whose genes are already partially or completely contained in sequence databases (*47*). The sequencing of more than 150 amino acids of the p123 telomerase subunit at protein amounts too low for chemical methods shows that mass spectrometry is now also valuable for sequencing previously unidentified proteins.

Telomerase activation accompanies the immortalization of cultured mammalian cells and is also a common property of human tumor cells (*48*). Thus, telomerase is considered to be a potential target for the development of tumor-specific drugs. Certain reverse transcriptase inhibitors developed as anti-HIV drugs have already been tested against telomerase with some success (*49*). The finding that the telomerase active site is related to that of known reverse transcriptases is expected to stimulate such efforts.

REFERENCES AND NOTES

1. J. D. Watson, *Nature* **239**, 197 (1972); J. Olovnikov, *J. Theor. Biol.* **41**, 181 (1973).
2. J. Lingner, J. P. Cooper, T. R. Cech, *Science* **269**, 1533 (1995); for a similar model regarding the macronuclear DNA of hypotrichous ciliates, see A. M. Zahler and D. M. Prescott, *Nucleic Acids Res.* **16**, 6953 (1988).
3. C. W. Greider and E. H. Blackburn, *Cell* **51**, 887 (1987).
4. ———, *Nature* **337**, 331 (1989); D. Shippen-Lenz and E. H. Blackburn, *Science* **247**, 546 (1990); D. P. Romero and E. H. Blackburn, *Cell* **67**, 343 (1991).
5. J. Lingner, L. L. Hendrick, T. R. Cech, *Genes Dev.* **8**, 1984 (1994).
6. M. S. Singer and D. E. Gottschling, *Science* **266**, 404 (1994).
7. M. J. McEachern and E. H. Blackburn, *Nature* **376**, 403 (1995).
8. J. Feng *et al.*, *Science* **269**, 1236 (1995); M. A. Blasco, W. Funk, B. Villeponteau, C. W. Greider, *ibid.*, p. 1267.
9. K. Collins, R. Kobayashi, C. W. Greider, *Cell* **81**, 677 (1995).
10. L. Harrington *et al.*, *Science* **275**, 973 (1997); J. Nakayama, M. Saito, H. Nakamura, A. Matsuura, F. Ishikawa, *Cell* **88**, 875 (1997).
11. J. Lingner and T. Nakamura, unpublished data. The BLAST search of protein databases was performed at the National Center for Biotechnology Information with GCG (Program Manual for the Wisconsin Package, Version 8.1, Genetics Computer Group, Madison, WI).
12. J. Lingner and T. R. Cech, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10712 (1996).
13. D. M. Prescott, *Microbiol. Rev.* **58**, 233 (1994).
14. P. Hammond, T. Lively, T. R. Cech, *Mol. Cell. Biol.* **17**, 296 (1997).
15. Amino acid analysis of 10% of the protein indicated a low picomole amount, insufficient for internal sequencing by Edman degradation. Protein bands were excised from the Coomassie-stained gel and digested in the gel (*16*). Reconstituted peptide mixture was desalted on a 50-nl microcartridge and eluted directly into the nanoelectrospray needle. Peptides were sequenced on a tandem mass spectrometer API III (PE-Sciex, Ontario, Canada) equipped with a nanoelectrospray ion source (*17*). Mass spectra (Q1 scans) were recorded with 0.1-dalton stepwidth and unit mass resolution. For the tandem mass spectrometric investigation, Q1 was set to transmit a 2-dalton mass window. The tan-

dem mass spectra were acquired with 0.2-dalton stepwidth. The mass resolution was sufficient to assign unambiguously an amino acid series to the identifiable COOH-terminal ion series. The procedure was repeated with the esterified peptide mixture as described (*16*). The following peptide sequences (*50*) were assigned: T1, YLLFQR; T2, DYNEEDFQVL(VK,AR); T3, (L,N)WDVLMK; T4, SFLMNNLTHYFR; T5, TLYSWLQK; T6, ETLAEVQEK; T7, AMLGNELFR; T8, LQTSFPLSPSK; T9, (A,Q)TLFTNLFR; T10, (L,T)ALMPNLNLR; T11, LFAQTNLVATPR; T12,LESWMQVETSAK; T13, QYFFQDEWNQVR; and T14, (E,D)SMNPENPNVNLLMR. The partial sequence EVDVD was obtained in the NH_2-terminal peptide t15 (MEVDVDNQADNHGIHSALK) and the sequence LFFAT in the peptide t16 (LFFATMDIEK). There are two discrepancies between the sequence determined by mass spectrometry and that deduced from the DNA sequence. The glutamine in peptide T2 was a lysine, and the order of the first two amino acids in T12 was reversed. In the above sequences, L signifies isoleucine or leucine, two isobaric amino acids that cannot be distinguished by our methods, and the parentheses indicate that the order of the first two amino acids could not be distinguished.
16. M. Wilm *et al.*, *Nature* **379**, 466 (1996).
17. M. S. Wilm and M. Mann, *Int. J. Mass Spectrom. Ion Proc.* **136**, 167 (1994); *Anal. Chem.* **68**, 1 (1996).
18. J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, C. M. Whitehouse, *Science* **246**, 64 (1989).
19. DNA was isolated from *E. aediculatus*, treated with mung bean nuclease, and subcloned into the Sma I site of pCR-Script SK (+) (Stratagene). The macronuclear gene encoding p123 was identified by colony hybridization.
20. J. Lingner and T. R. Cech, unpublished data.
21. M. Mann and M. S. Wilm, *Anal. Chem.* **66**, 4390 (1994).
22. *P* is the probability that the specified segment pair score would be observed in a search of the same size with a random sequence.
23. J.-L. Ferat and F. Michel, *Nature* **364**, 358 (1993).
24. T. S. Lendvay, D. K. Morris, J. Sah, B. Balasubramanian, V. Lundblad, *Genetics* **144**, 1399 (1996).
25. *EST1*, which has been described [V. Lundblad and J. W. Szostak, *Cell* **57**, 633 (1989)], is not essential for core telomerase activity (*33*). *EST4* is identical to *CDC13* [B. Garvik, M. Carson, L. Hartwell, *Mol. Cell. Biol.* **15**, 6128 (1995)]. Biochemical analysis has shown that Est1p and Cdc13p are single-stranded telomeric DNA binding proteins that are proposed to facilitate access of telomerase to chromosome ends [C. I. Nugent, T. R. Hughes, N. F. Lue, V. Lundblad, *Science* **274**, 249 (1996); J. J. Lin and V. A. Zakian, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13760 (1996); V. Virta-Pearlman, D. K. Morris, V. Lundblad, *Genes Dev.* **10**, 3094 (1996)].
26. O. Poch, I. Sauvaget, M. Delarue, N. Tordo, *EMBO J.* **8**, 3867 (1989).
27. Y. Xiong and T. H. Eickbush, *ibid.* **9**, 3353 (1990).
28. V. Lundblad and E. H. Blackburn, *Cell* **73**, 347 (1993).
29. C. Nugent and V. Lundblad, unpublished data.
30. S. G. Sarafianos, V. N. Pandey, N. Kaushik, M. J. Modak, *Biochemistry* **34**, 7207 (1995).
31. Est2p was overexpressed under control of the *ADH* promoter as a Myc_3-tagged fully functional derivative. Extracts were immunoprecipitated with an antibody to c-Myc. Wild-type and mutant Est2 proteins were precipitated with the antibody and detected on an immunoblot. Telomerase RNA was coimmunoprecipitated with Est2p only in strains harboring the Myc_3-tagged gene and was detected on a Northern (RNA) blot. U1 snRNA was not substantially precipitated but was detected in the crude extract.
32. Yeast was grown in yeast extract, peptone, and dextrose (YPD) medium to an optical density (600 nm) of 1.8, and extracts were prepared as described [A. Ansari and B. Schwer, *EMBO J.* **14**, 4001 (1995)] with the following modifications: Cell pellets were resuspended in two volumes of extract buffer [20 mM tris-acetate (pH 7.5), 300 mM potassium glutamate, 1.1 mM $MgCl_2$, 0.1 mM EDTA, 5% glycerol, 1 mM dithiothreitol (DTT), and 0.5 mM phenylmethylsulfonyl fluoride]. Extracts were concentrated threefold on centricon 30 (Amicon) and fractionated by glycerol gradient centrifugation in extract buffer containing 15 to 40% glycerol (*12*). Centrifugation was performed in a SW55Ti rotor (Beckman) at 304,000g_{max} for 12 hours at 4°C. Nineteen fractions were collected from the top of the gradient and assayed for the telomerase RNA subunit *TLC1* by Northern blotting. In *est2-Δ* and *est2* mutant strains, telomerase did not show a major shift in its sedimentation coefficient relative to the wild type, but the large size of the RNA and the limited resolution of the glycerol gradients (Fig. 5) would have prevented a small shift from being observed. Telomerase-containing fractions were pooled and concentrated fivefold on centricon 30. Telomerase assays (10 μl) contained 20 mM tris-acetate (pH 7.5), 1 mM $MgCl_2$, 30 mM potassium glutamate, 1 mM DTT, 1 μM oligonucleotide substrate (5′-TGTGGTGTGTGTGGG-3′), 100 μM deoxy- or dideoxyguanosine triphosphate (dGTP or ddGTP), 15% (v/v) [α-^{32}P]dTTP (800 Ci/mmol; 10 μCi/μl), and 10% (v/v) telomerase fraction. After 1 hour at 30°C, the reactions were stopped, digested with proteinase K, ethanol precipitated, and separated as in (*5*). Fractions were derived from the following strains: Fig. 6, A and B: TVL268 [*Mat***a** *ura3-52 ade2-101 lys2-801 leu2-Δ1 his3-Δ200/CF (TRP1 SUP11)*]; Fig. 6C, lanes 2 to 4: TVL268; Fig. 6C, lanes 6 to 8: TVL268 *TLC1-1(HaeIII)*; Fig. 6D, lanes 1 and 2, and Fig. 6E, lane 1: AVL78 (*Mat***a** *leu2 trp1 ura3-52 prb prc pep4-1*); Fig. 6D, lanes 3 and 4, and Fig. 6E, lane 2: AVL78 *est2-Δ1::URA3*; Fig. 6D, lanes 5 and 6, and Fig. 6E, lane 3: AVL78 *tlc1-Δ::LEU2*; Fig. 6E, lanes 4, 7, 8, and 9: TVL268; Fig. 6E, lanes 5 and 8: TVL268 *est2-D530A*; and Fig. 6E, lanes 6 and 9: TVL268 *est2-D671A*. The *est2-Δ1* and *tlc1-Δ* deletion strains were constructed by one-step gene disruption, whereas the *est2-D530A*, *est2-D671A*, and *TLC1-1(HaeIII)* mutant alleles were constructed by replacement of the *EST2-wt* and *TLC1-wt* wild-type alleles, respectively. In each case, extracts were prepared for analysis after as few generations of growth as possible.
33. M. Cohn and E. H. Blackburn, *Science* **269**, 396 (1995); J. Prescott and E. H. Blackburn, *Genes. Dev.* **11**, 528 (1997).
34. N. F. Lue and J. C. Wang, *J. Biol. Chem.* **270**, 21453 (1995).
35. L. A. Kohlstaedt, J. Wang, J. M. Friedman, P. A. Rice, T. A. Steitz, *Science* **256**, 1783 (1992); E. Arnold *et al.*, *Nature* **357**, 85 (1992); A. Jacobo-Molina *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6320 (1993).
36. D. L. Ollis, P. Brick, T. A. Steitz, *Nature* **313**, 765 (1985). For comparison of different polymerase structures, see C. M. Joyce and T. A. Steitz, *Annu. Rev. Biochem.* **63**, 777 (1994), and R. Sousa, *Trends Biochem. Sci.* **21**, 186 (1996).
37. J. C. Kennell, H. Wang, A. M. Lambowitz, *Mol. Cell. Biol.* **14**, 3094 (1994); S. Zimmerly *et al.*, *Cell* **83**, 529 (1995).
38. D. Baltimore, *Nature* **226**, 1209 (1970); H. M. Temin and S. Mizutani, *ibid.*, p. 1211.
39. J. D. Boeke and V. G. Corces, *Annu. Rev. Micro.* **43**, 403 (1989).
40. M. Inouye and S. Inouye, *ibid.* **45**, 163 (1991).
41. H. Wang and A. M. Lambowitz, *Cell* **75**, 1071 (1993).
42. H. M. Temin, *Annu. Rev. Genet.* **8**, 155 (1974).
43. J. E. Darnell and W. F. Doolittle, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 1271 (1986). See also N. Maizels and A. Weiner, in *The RNA World*, R. F. Gesteland and J. F. Atkins, Eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1993), pp. 577–602.
44. H. Biessmann *et al.*, *Cell* **61**, 663 (1990); R. W. Levis, R. Ganesan, K. Houtchens, L. A. Tolar, F. Sheen, *ibid.* **75**, 1 (1993).
45. M. L. Pardue, O. N. Danilevskaya, K. Lowenhaupt, F. Slot, K. L. Traverse, *Trends Genet.* **12**, 48 (1996).
46. J. Lingner, A. Shevchenko, M. Mann, T. R. Cech, unpublished data.
47. See, for example, A. Shevchenko *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 14440 (1996). Reviewed in S. D. Patterson and R. Aebersold, *Electrophoresis* **16**, 791 (1995).
48. C. M. Counter *et al.*, *EMBO J.* **11**, 1921 (1992); T. de Lange, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 2882 (1994).
49. C. Strahl and E. H. Blackburn, *Nucleic Acids Res.* **22**, 893 (1994); *Mol. Cell. Biol.* **16**, 53 (1996).
50. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
51. Twenty single colonies from five independent transformations of an *est2-Δ* strain were assayed for each mutation. The results, including the slight reduction in colony size observed for $est2_{D536A}$ and $est2_{D592A}$, were highly reproducible. In addition, the senescence phenotype of these two mutants (as well as that of the more severely affected $est2_{D530A}$, $est2_{D670A}$, and $est2_{D671A}$ mutants) was enhanced in an *est2 rad52* strain background (*29*). The absence of *RAD52* gene function has been shown to eliminate a backup pathway for telomere maintenance, thereby enhancing the phenotype of *est* mutant strains (*24*, *28*).
52. S. Evans and V. Lundblad, unpublished data.
53. We thank C. Grosshans, K. Goodrich, and N. Walcott for technical assistance; T. Nakamura, C. Chapon, A. Zaug, and J. Hansen for many helpful discussions; and D. Underwood and E. Jabri for help with Fig. 7. Supported in part by the Swiss National Science Foundation (J.L.), NIH grant AG11728 and Geron Corporation (V.L.), and a grant from the German Technology Ministry (Bundesministerium für Biologische Forschung) (M.M.).

4 March 1997; accepted 2 April 1997