532–533 *Nucleic Acids Research, 1996, Vol. 24, No. 3*

Method for determining RNA 3′ ends and application to human telomerase RNA

Arthur J. Zaug, Joachim Lingner and Thomas R. Cech*

Howard Hughes Medical Institute, Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO 80309-0215, USA

Received November 17, 1995; Accepted November 30, 1995

Mapping the 3′ ends of RNA molecules at nucleotide sequence resolution is a frequently encountered problem. S1 nuclease mapping (1) is a sensitive and often very successful approach for localizing 3′ ends, although it does not usually allow specification of an exact end with confidence. Enzymatic labeling of the 3′-terminus of an RNA followed by RNA sequencing is readily applicable to RNAs that can be obtained in pure form, although it becomes complicated if there is significant 3′-terminal heterogeneity.

Tailing of RNA with poly(A) polymerase from *Escherichia coli* was used >10 years ago in the cloning and sequencing of U7 snRNA, including determination of its 3′ end (2). The availability of very active, pure recombinant yeast poly(A) polymerase (3) facilitates the use of this enzyme for 3′-end determination; the method is applicable to small quantities and requires neither a pure RNA species nor a homogeneous 3′ end. As described by Lingner *et al.* (4), the successive steps are: (i) poly(A)-tailing of the RNA; (ii) cDNA synthesis using reverse transcriptase (RT) and a primer ending in 18 T residues, which hybridizes to the poly(A); (iii) PCR using *Taq* polymerase, the dT_{18} primer and a second DNA primer with the same sequence as a known internal region of the RNA; (iv) cloning of the PCR products; and (v) dideoxynucleotide sequence analysis of multiple cloned DNAs. Because poly(A) polymerase requires a 3′-hydroxyl group on its RNA substrate, this method has an additional advantage over the S1 nuclease method: products of transcriptional termination and most 3′-end processing events are revealed, while degradation products (typically 2′,3′-cyclic phosphate or 3′-phosphate termini) are discriminated against.

In applying the poly(A) polymerase method to HeLa cell RNA using an internal primer specific for the telomerase RNA, we encountered an ambiguity that is of general significance (see also ref. 2). The sequence of the cDNA of the poly(A)-tailed RNA is shown below, aligned with the sequence of the corresponding region of the gene determined by Feng *et al.* (5):

cDNA.....CATGCAAAAA...
Gene.....CATGCAGTTC....

Without additional information, it is impossible to tell if the underlined A was added by poly(A) polymerase (in which case the telomerase RNA terminates in CATGC) or if, alternatively, the telomerase RNA terminates with the underlined A.

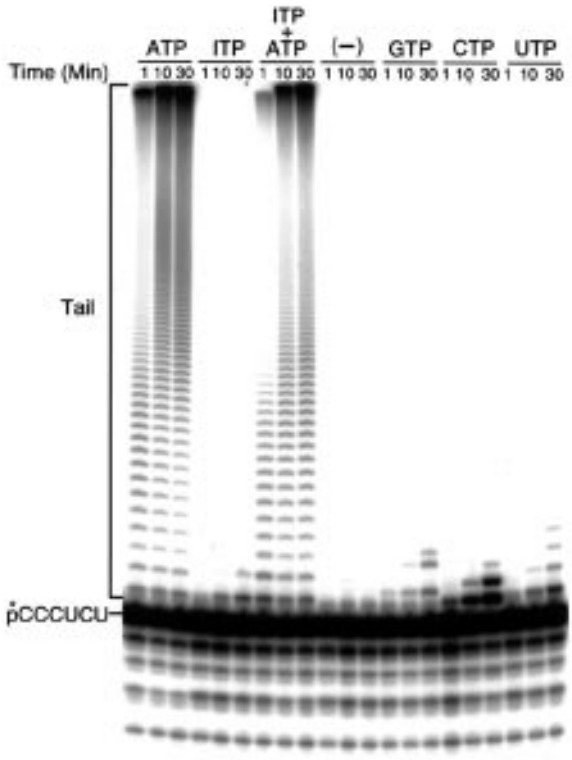

Figure 1. Utilization of different nucleoside triphosphates by poly(A) polymerase. The oligoribonucleotide CCCUCU was chemically synthesized and labeled at its 5′-end with ^{32}P. It was then incubated at 30°C with recombinant yeast poly(A) polymerase (USB, Amershan LIFE SCIENCE Inc.) and 0.5 mM of the indicated nucleoside triphosphate under conditions of Lingner and Keller (3). At the times indicated, portions were removed and reaction was stopped with an equal volume of 95% formamide. Products were analyzed on a 20% polyacrylamide (29:1)–8 M urea gel. (ITP + ATP) lanes: time points following the addition of ATP after a 30 min preincubation with ITP. (–) lanes: no nucleoside triphosphate added.

A solution to this problem was based on the ability of poly(A) polymerase to incorporate other nucleotides at low efficiency. The non-A nucleotide can be used to mark the precise 3′ terminus

* To whom correspondence should be addressed

# of Clones	Telomerase RNA sequence	Tail[a]	Vector sequence
	423[b] 445		
16	CCCAGGACTCGGCTCACACATGC	AAAAAAAAAAAAAAAA	GTCGACCTGCAGG
2	CCCAGGACTCGGCTCACACATGC	GAAAAAAAAAAAAAAA	GTCGACCTGCAGG
1	CCCAGGACTCGGCTCACACATGC	GGAAAAAAAAAAAAAA	GTCGACCTGCAGG
1	CCCAGGACTCGGCTCACACATGC	GGGAAAAAAAAAAAAA	GTCGACCTGCAGG
1	CCCAGGACTCGGCTCACACATGC	AGAAAAGGAAAAAAAA	GTCGACCTGCAGG
1	CCCAGGACTCGGCTCACACATGC	GGAGAAAAAAAAAAAA	GTCGACCTGCAGG

Figure 2. Sequences of cDNA clones of human telomerase RNA that identify the RNA 3′ terminus. HeLa cell RNA was subjected to polyacrylamide gel electrophoresis, and a wide region of the gel containing telomerase RNA was excised. The resulting RNA was treated with poly(A) polymerase as described in Figure 1, with 0.5 mM ITP for 20 min followed by 0.5 mM ATP for an additional 10 min. Then cDNA was synthesized with AMV reverse transcriptase (Life Sciences) and primer dT_{18} complementary to the newly synthesized poly(A) tail, first at 37°C for 20 min and then 50°C for 15 min. [Primer dT_{18} is dGT**GTCGAC**(T_{18}); nucleotides in bold give a *Sal*I restriction endonuclease site.] The mixture was heated at 95°C for 3 min to denature the reverse transcriptase. dNTPs were removed by gel filtration on a G25 spin column (Boehringer). Amplification by PCR using primers dT_{18} and HTelo-294R, which anneals to a known sequence within the telomerase RNA, resulted in a double-stranded DNA product of the appropriate size (data not shown). [Primer HTelo-294R is dCGC**GGATCC**GGAGGCAC-CCACTGCCACCC; nucleotides in bold give a *Bam*HI site.] The restriction endonuclease sites were added during the PCR amplification step for ease of cloning. Following gel purification and cleavage with *Sal*I and *Bam*HI restriction enzymes, the fragment was cloned into pUC19 and sequenced with a Sequenase kit (USB, Amershan LIFE SCIENCE Inc.).
[a]In general tails contained 18 As, occasionally more. Inefficient utilization of ITP during tailing followed by RT–PCR was expected to give a small number of Gs in the tail region of the cDNA, as seen in six of the 22 clones.
[b]Position in telomerase RNA sequence (5).

of an unknown RNA, while the As provide the primer binding site for reverse transcription. Studies with a synthetic oligoribonucleotide (Fig. 1) showed that the polymerase adds inosine inefficiently, and that this incorporation does not inhibit subsequent extension with As. Other NTPs also serve as substrates, CTP being used more efficiently than the others (although we have only tested ITP in the application described below). Repeating this experiment with a model RNA ending in As (CUUAAAAA) gave extension with the various NTPs similar to that seen in Figure 1 (data not shown). Although careful substrate specificity studies remain to be done, it appears that the method will mark the precise 3′ ends of a variety of RNAs, without a large preference for those ending in A or not ending in A.

Application of this modified technique to the human RNA preparation led to identification of a single 3′ end for the telomerase RNA (Fig. 2). The RNA was incubated with ITP plus poly(A) polymerase prior to the addition of ATP, followed by RT–PCR amplification. While most clones showed an A-tail immediately after position 445 and thus were inconclusive regarding whether A446 was part of the original RNA, two clones had a single G at position 446 preceding the A tail, one clone had two Gs, and one had three Gs. (It is expected that inosine has template properties like G during reverse transcription.) We conclude that the most abundant 3′ end of this telomerase RNA is at C445, with no evidence for any heterogeneity in the 22 clones sequenced. Our Northern hybridization analysis (data not shown) gave a size of telomerase RNA of 450 nt, relative to denatured ϕX174 *Hae*III restriction fragments which provide reliable molecular weight standards for RNA, consistent with the 3′ end occurring at position 445. This 3′ end is in the vicinity of that estimated earlier (5).

Thus, the human telomerase RNA does not terminate heterogeneously within a run of U residues, as do the corresponding RNAs of ciliates (4,6). The ciliate RNAs are probably transcribed by RNA polymerase III (7). Alpha amanitin sensitivity suggests that the human telomerase RNA is instead transcribed by pol II, but is not polyadenylated (8). We now add that it is not even oligoadenylated. The homogeneity of the human telomerase RNA 3′ end strongly suggests that it is generated by RNA processing. If so, it will be interesting to determine whether the processing machinery is shared with that responsible for maturation of snRNAs or other poly(A)$^-$ RNAs.

ACKNOWLEDGEMENTS

We thank Dr Bryant Villeponteau and his co-workers at Geron Corporation for providing a plasmid containing the telomerase RNA gene and its sequence prior to publication. A.J.Z. is a Senior Associate and T.R.C. is an Investigator of the Howard Hughes Medical Institute. J.L. is supported by a fellowship from the Swiss National Science Foundation.

REFERENCES

1 Berk, A.J. and Sharp, P.A. (1977) *Cell*, **12**, 721–732.
2 Strub, K., Galli, G., Busslinger, M. and Birnstiel, M.L. (1984) *EMBO J.*, **3**, 2801–2807.
3 Lingner, J. and Keller, W. (1993) *Nucleic Acids Res.*, **21**, 2917–2920.
4 Lingner, J., Hendrick, L.L. and Cech, T.R. (1994) *Genes Dev.*, **8**, 1984–1998.
5 Feng, J., Funk, W.D., Wang, S.-S., Weinrich, S.L., Avilion, A.A., Chiu, C.-P., Adams, R.R., Chang, E., Allsopp, R.C., Yu, J., Le, S., West, M.D., Harley, C.B., Andrews, W.H., Greider, C.W. and Villeponteau, B. (1995) *Science*, **269**, 1236–1241.
6 Greider, C.W. and Blackburn, E.H. (1989) *Nature*, **337**, 331–337.
7 Yu, G.L., Bradley, J.D., Attardi, L.D. and Blackburn, E.H. (1990) *Nature*, **344**, 126–132.
8 Avilion, A.A. (1995) Ph.D. thesis, State University of New York, Stony Brook, as cited in ref. 5.