## Architecture of telomerase RNA

# Anamitra Bhattacharyya and Elizabeth H.Blackburn<sup>1</sup>

Department of Microbiology and Immunology, Box 0414, University of California, San Francisco, CA 94143, USA

<sup>1</sup>Corresponding author

Communicated by D.M.J.Lilley

Telomerase, an essential ribonucleoprotein reverse transcriptase, adds telomeric DNA to the ends of eukaryotic chromosomes. We examined the conformational properties of the naked RNA moiety of telomerase from two related ciliates, Tetrahymena thermophila and Glaucoma chattoni. As well as finding evidence for features proposed previously on the basis of phylogenetic comparisons, novel conserved structural properties were revealed. Specifically, although the region around helix III was previously proposed to form a pseudoknot, our results indicate that in the naked RNA this region maintains a level of 'plasticity', probably in an equilibrium favoring one of two helices. In addition, these studies reveal that the templating domain is not entirely single-stranded as previously proposed, but is ordered due to constraints imposed by other parts of the RNA. Finally, our results suggest that the GA bulge in helix IV may introduce a structurally conserved kink. We now propose a 'two-domain' structure for the telomerase RNA based on function: one conformationally flexible domain, which includes the template and the region around helix III, involved with enzymatic function, and a second largely helical domain, including helices I and IV and the proposed kink, which may serve as a scaffold for protein binding.

Key words: pseudoknot/RNA structure/telomerase

## Introduction

Telomerase is the specialized, ribonucleoprotein reverse transcriptase which copies a template sequence within its essential RNA moiety into complementary telomeric DNA repeats at the ends of eukaryotic chromosomes (Greider and Blackburn, 1987, 1989; Yu *et al.*, 1990). Telomerase activity has been shown to be essential for long-term maintenance in the ciliate *Tetrahymena thermophila* (Yu *et al.*, 1990) and in at least one yeast species (M.McEachern and E.H.Blackburn, unpublished results). Some evidence has suggested that telomerase may also be necessary for continued division of cancer cells (Hastie *et al.*, 1990; Counter *et al.*, 1992, 1994). The functioning of another ribonucleoprotein, the ribosome, is disrupted by binding of certain antibiotics (Woodcock *et al.*, 1991; Leviev *et al.*, 1994; reviewed in Noller, 1991). Therefore knowledge

of the structural properties of telomerase RNA will be important for possible therapeutic approaches involving telomerase as a potential target for anti-tumor drug binding.

Comparison of the sequence of the 159 nucleotide telomerase RNA of T.thermophila (Greider and Blackburn, 1989) with telomerase RNAs from five other Tetrahymena species and the related ciliate, Glaucoma chattoni, suggested a conserved secondary structure for these RNAs (Romero and Blackburn, 1991). This phylogenetic model (Romero and Blackburn, 1991), based on the co-variation of residues in putative helical regions of the aligned sequences, proposed the existence of four helical regions (designated I, II, III and IV) and an unstructured domain which included the RNA template. The templating nucleotides, which act as the template for the synthesis of  $G_4T_2$ telomeric DNA repeats (Yu et al., 1990; Yu and Blackburn, 1991; Autexier and Greider, 1994; D.Gilley, M.Lee and E.H.Blackburn, unpublished results) (refer to Figure 1C), were constrained to be single stranded in the phylogenetic model since they are accessible to base pairing by telomeric DNA oligonucleotides (Greider and Blackburn, 1989). It was also proposed independently that the RNA contains a putative pseudoknot structure involving helix III (ten Dam et al., 1991).

Here we present a conformational analysis of the naked telomerase RNA in solution, designed to test directly and refine the phylogenetic secondary structure model. In vitro synthesized telomerase RNA was probed with enzymatic and chemical probes specific for single-stranded or doublestranded regions. While the structures of the naked T.thermophila and G.chattoni telomerase RNAs (35% sequence dissimilarity) were generally consistent with the previously predicted phylogenetic model (Romero and Blackburn, 1991), this analysis also reveals new features of the telomerase RNA: first, the base sequence associated with the helix III region, in isolation from the rest of the telomerase RNA, forms a pseudoknot, but the full-length naked RNA exhibits conformational 'plasticity', in contrast to other parts of the RNA. Second, the residues in the templating domain, previously thought to be singlestranded, appear to be constrained by the folding of other parts of the RNA and thereby ordered. Our results also suggest that the conserved GA bulge in helix IV introduces a structurally conserved kink. These findings suggest that the helix III and templating regions interact with other regions of the telomerase RNA. However, it is not known how the presence of telomerase protein component(s) affects these properties of the RNA. This study provides experimental evidence on the structure of naked telomerase RNA in solution, and provides an essential framework for gaining insight into the function of this RNA moiety, as both a potential scaffold for binding the telomerase protein

#### A.Bhattacharyya and E.H.Blackburn

component(s), and as a contributor to the catalytic properties of the telomerase holoenzyme.

## Results

## Structure probing of the telomerase RNA

The secondary structure of in vitro transcripts of T.thermophila and G.chattoni telomerase RNAs (prepared as described in Materials and methods) were analyzed and compared in order to distinguish common structural features from sequence-specific ones. 3'-End-labeled T.thermophila and G.chattoni RNAs were heat denatured and slow cooled to allow folding into a 'native' conformation using a standard high salt buffer (Stern et al., 1988) in order to mimic the conformation of the RNA as it might exist in the ribonucleoprotein complex, where presumably telomerase protein components would replace the cationic species in shielding electrostatic repulsion between the negatively charged phosphate backbone. The folded telomerase RNAs were treated with each enzyme or chemical probe and positions of the modification/cleavage sites determined by denaturing polyacrylamide gel electrophoresis. A consensus structure was derived by examining the cleavage patterns from three different structural probes, each with different specificities, and comparing these results with the phylogenetic model reported previously (Romero and Blackburn, 1991).

Diethylpyrocarbonate (DEP) has been successfully employed previously to study nucleic acid structure (Ehresmann et al., 1987). In general, DEP strongly modifies adenines in a single-stranded environment, while adenines in helical or highly stacked states are protected. We used DEP as a structural probe for the two telomerase RNAs (Figure 1A and B) and the modification patterns are indicated on the phylogenetic models in Figure 1C. In both T.thermophila and G.chattoni telomerase RNAs there was protection of helices II (denoted by circled residues in Figure 1C, e.g. A20, 33 in T.thermophila) and IV (e.g. A113, 115, 124, 140 in T.thermophila), as well as helix I (applicable in *G.chattoni*). However, the protection pattern in and around helix III was more complex and, as discussed below, this region of the RNA appears to form alternate structures. Conversely, adenines were modified in the proposed single-stranded regions, such as the loops of the putative hairpins II (indicated by arrowed residues in Figure 1C, e.g. A28, 29 in T.thermophila), III (e.g. A89, 90 and 91; T.thermophila) and IV (e.g. A133, 136; T.thermophila), and the conserved GA bulge in helix IV (G122). Despite the overall sequence divergence of the two telomerase RNAs, their patterns of chemical reactivities were similar, indicating they were structurally alike. Preliminary chemical probing of the *T.thermophila* RNA *in vitro* with dimethyl sulfate, using primer extension and reverse transcriptase techniques, also produced similar overall results (C.Strahl and E.H.Blackburn, unpublished results).

The T.thermophila and G.chattoni telomerase RNAs were also probed with RNase V1, which cleaves stacked or base paired residues in helical domains (Lowman and Draper, 1986; Wyatt and Walker, 1989). The data from the T.thermophila RNA is presented in Figure 2A, and the results from both species are illustrated in Figure 2B. The previously proposed helical regions I-IV all showed evidence of V1 nuclease sensitivity. V1-sensitive nucleotides were also identified at positions not predicted to have double-stranded character (Romero and Blackburn, 1991), such as in the loops of helices III and IV, suggesting that these loops are structured in some way. Evidence for structure in small RNA loops involving base stacking interactions has been reported (Haasnoot et al., 1986; Tuerk et al., 1988; Cheong et al., 1990; Varani et al., 1991). Unexpectedly, the single-stranded domain between helices II and III was RNase V1 sensitive at numerous positions in both RNAs, including the templating nucleotides (underlined in Figure 2B and D). This was the first suggestion of an ordered structure, possibly involving increased base stacking, in a part of the RNA previously thought to be unstructured.

Structure probing analysis using RNase T1, which cleaves single-stranded guanines (Ehresmann et al., 1987), was performed on the two RNAs. The data from T.thermophila is shown in Figure 2C, and the cleavage results from both RNAs are illustrated in Figure 2D. The strong RNase T1 cleavages in the RNAs in general coincided with guanine bases predicted to possess single-stranded character. For instance, in T.thermophila guanine residues in helices I and II, and G114 and G128 in helix IV, were largely protected from cleavage, while predicted singlestranded regions, such as G26 in the loop of hairpin II, were cleaved. Ribonuclease T1 cleavage was also prominent at G121 in the conserved GA bulge in helix IV. The modulation of RNase T1-induced guanine cleavages (G59, 61 and 64) in the previously proposed singlestranded region between helices II and III in the T.thermophila RNA was consistent with the V1 nuclease data, both probes showing this region to have only partially singlestranded character.

These results were consistent between the two RNA species, and with most aspects of the proposed phylogenetic model (Romero and Blackburn, 1991). However, certain features of the chemical and enzymatic probing

**Fig. 1.** Diethylpyrocarbonate (DEP) modification of 3'-end-labeled *in vitro T.thermophila* and *G.chattoni* telomerase RNAs. (**A**) Autoradiogram of a 10% denaturing urea-polyacryalamide gel showing DEP reactive bases from the whole of the *T.thermophila* (left panel) and *G.chattoni* (right panel) telomerase RNAs. Positions of all adenines from each RNA, identified by performing the chemical modification reaction under denaturing conditions (described in Materials and methods), are indicated to the left of each panel. Roman numerals show the positions of the putative helical regions predicted by the phylogenetic model (Romero and Blackburn, 1991). Chemical modification reactions (lanes marked +) were performed at the temperatures shown, prior to aniline cleavage. A control reaction for each temperature (lanes marked -) was also included which involved incubating the folded RNA, without DEP, under the same conditions of time and temperature, followed by aniline treatment. (**B**) Greater resolution of higher mol. wt products from the same modification reaction. An extra reaction temperature lane for *T.thermophila* (10°C) and *G.chattoni* (40°C) is included. The templating region between helices II and III in each RNA is indicated by a filled triangle. (**C**) Schematic summary of the positions of DEP modifications are indicated on the telomerase phylogenetic models (Romero and Blackburn, 1991). Circled nucleotides represent residues that displayed either very poor, or no cleavage and arrows show residues that were cleaved. Strong and weak cleavages are represented by large and small arrows, respectively. Inset shows the mode of action of this chemical probe.



C T. thermophila

\_\_\_\_\_



#### A.Bhattacharyya and E.H.Blackburn

analyses were unexpected and these were investigated in more detail.

### A pseudoknot associated with helix III

Based on phylogenetic co-variation, ten Dam *et al.* (1991) proposed that a pseudoknot forms between the loop in helix III and the 5' single-stranded region contiguous to it, illustrated in Figure 3A. Pseudoknot motifs have been reported previously in several naturally occurring RNAs (Pleij *et al.*, 1985; McPheeters *et al.*, 1988; Brierley *et al.*, 1989; Powers and Noller, 1991). High resolution NMR analyses of pseudoknotted RNAs (Puglisi *et al.*, 1988, 1990) have suggested that the two pseudoknot helices stack co-axially to form a quasi-continuous A-type RNA helix stabilized by end-on helix—helix stacking interactions.

Chemical probing of helix III in both full-length T.thermophila and G.chattoni RNAs showed it to be more reactive than expected for a stable helix. For example, in T.thermophila, A79, 93 and 94 were modified by DEP (Figure 1C), and G97 and 98 were cleaved by RNase T1 (Figure 2D), even though all these are internal residues in the middle of predicted helix III. Furthermore, RNase V1 cleavages at C72 and U73, located 5' to helix III in T.thermophila (Figure 2B) were consistent with these residues being involved in pairing interactions. These observations on the full-length RNA, together with the proposal that helix III might form a pseudoknot (ten Dam et al., 1991), led us to test whether the helix III region could in fact adopt a pseudoknot conformation. First, to determine whether this part of the T.thermophila telomerase RNA was by itself sufficient to adopt a pseudoknot structure in solution, the secondary structures of short in vitro synthesized oligoribonucleotides were probed with DEP (Figure 3B). The oligoribonucleotide PKab had the wild-type T.thermophila sequence (nucleotides 66-102; Figure 3A) from the helix III region (Figure 3A inset). Three variant oligoribonucleotides, PKa, PKb and PK0, had two mismatches in either stem-b, stem-a, or both, respectively, which were predicted to disrupt pseudoknot formation. These mismatches were introduced as adenine substitutions so that these residues could be probed for DEP reactivity. Each RNA species was folded and treated with DEP under conditions identical to those used for the full-length RNA, to determine whether their patterns of chemical reactivities were consistent with the formation of a pseudoknot in the case of PKab, and its disruption in the variant RNA oligomers.

As shown in Figure 3B and C, PKab RNA displayed

chemical protection of adenines in the putative pseudoknot stem-b (A70, A83 and to a lesser extent the terminal A69) and in stem-a (A79, A80, A93 and A94), while showing modification of the other predicted single-stranded bases (A67, 89, 90, 91, 100, 101). Loop-b in the PKab pseudoknot would have to span the major groove between stems-b and -a. Though it is possible for U74 and C75 to span this groove (perhaps requiring changes in sugar puckering from an A-like C3' endo to a B-like C2' endo conformation), the inclusion of A76 in loop-b (or in slow equilibrium with base pairing to U99) is suggested by the moderate DEP reactivity of A76. DEP modification of PKab as a function of increasing temperature and therefore unfolding of the putative pseudoknot, was also performed. At 15°C (data not shown) and 30°C (Figure 3B and C), similar downward modulations in chemical reactivity in parts of PKab were observed. However, DEP modification at all bases was generally uniform at 60°C, and similar to that at 90°C (data not shown), indicating a concerted melting of the whole PKab structure by 60°C.

For PKa the results were consistent with the prediction that stem-b was disrupted: A70, 71 and 72 were reactive to DEP, and A67 and 69 displayed similar extents of modifications to their counterparts in PKab at 60°C (Figure 3C), as would be expected if they were single-stranded. Significantly, A83 and stem-a residues A79, 80, 93 and 94 were partially protected, although still somewhat more reactive than they were in PKab. We propose that overall stability of stem-a is reduced (although A83 could now pair with U92, extending stem-a) because it can no longer stack onto disrupted stem-b in the pseudoknot. This 'action-at-a-distance' strongly suggests that residues in stem-b and -a are linked structurally. Loop-a residues A89, 90 and 91 were more reactive to DEP in PKa than in PKab, as would be predicted if these residues in the larger loop of PKa were less constrained than in the smaller loop-a of PKab.

In PKb, stem-a was predicted to be severely disrupted by substituting G97 and G98 with adenines. However, A69 and A70 in stem-b showed protection similar to that in PKab, suggesting that stem-b structure was not perturbed by the disruption of stem-a. The remaining adenines showed a more uniform pattern of modification, consistent with these bases adopting a more single-stranded state. In contrast, PK0, disrupted in both the proposed pseudoknot stems, displayed a relatively uniform pattern of DEP modification, suggesting that this RNA is essentially single-stranded. We noted, however, that in PK0 a variation in reactivity between A79, 80 and 83 was observed (Figure

**Fig. 2.** Enzymatic structural probing of the *T.thermophila* and *G.chattoni* telomerase RNAs to identify helical domains. (A) Autoradiogram of 10% denaturing polyacrylamide gel analysis showing shorter RNase V1-sensitive bases from the *T.thermophila* telomerase RNA, and (upper panel) a greater resolution of longer products. The positions of all the guanines were identified with RNase T1 (marked 'T1'), and a hydroxyl cleavage ladder (OH-) provided additional mol. wt markers to allow identification of V1 nuclease-sensitive nucleotides. A no enzyme lane (-E) shows the positions of degradation products. The numbering indicates the guanine positions in the telomerase RNA, and Roman numerals show the positions of the putative helical regions predicted by the phylogenetic model (Romero and Blackburn, 1991). The templating region in each RNA is indicated by a filled triangle. (**B**) Schematic summaries of the positions of RNase V1 cleavages are indicated on the telomerase phylogenetic models (Romero and Blackburn, 1991). Circled nucleotides represent residues that displayed either very poor, or no cleavage, and arrows show bases that were cleaved. Strong and weak cleavages are represented by large and small arrows, respectively. Inset illustrates that the helical trajectory of helix IV is kinked (by an angle, q) to form helices IV-a and -b, by the presence of the conserved GA bulge. (C) Autoradiogram of 10% denaturing polyacrylamide gel showing RNase T1 digestion products from structural probing of the entire *T.thermophila* telomerase RNA. Mol. wt markers were as described above. A no enzyme lane (-E) shows the positions of degradation products. (**D**) Summaries of the RNase T1 cleavage positions from *T.thermophila* and *G.chattoni* RNAs are indicated on the phylogenetic models (Romero and Blackburn, 1991). Positions of RNase-sensitive residues near to the 5'- end of each RNA were determined by running the enzyme-cleaved samples further on denaturing gels so as to resolve them.



3B and C). As no RNA is completely single-stranded and unstructured at the temperature and salt concentrations used, the four extra adenines introduced into this molecule might allow part of it to fold into a structure producing this effect. However, the same pattern was observed even in the absence of any Na<sup>+</sup> and Mg<sup>2+</sup> ions added to the DEP reaction (data not shown), indicating this was not a secondary structure effect. In summary, the structural studies on PKab RNA and its variant derivatives, PKa, PKb and PK0, suggest that the base sequence associated with helix III and the 5' contiguous region has the ability to form a pseudoknot.

The chemical probing experiments on the full-length T.thermophila RNA and the PK series of oligoribonucleotides were carried out under identical reaction conditions and thus could be compared directly. Figure 3D (inset) shows two alternative RNA structures that might form in the region associated with helix III in the full-length T.thermophila RNA, besides the pseudoknot shown in Figure 3A. Helix III-a represents the sequence of PKab shown as the predicted stem-a alone, originally called helix III (Romero and Blackburn, 1991), and III-b is stemb from the pseudoknot, shown here by itself. If PKab is considered to form a pseudoknot, the DEP pattern of the full-length RNA (Figure 3D and discussed in Table I) is not consistent with a full pseudoknot, nor with formation of only helix III-a. The DEP modification pattern of the full-length RNA is most consistent with predominance of helix III-b. Together with the T1 and V1 cleavage data, shown in Figure 3D (inset), this suggests that this region in the full-length telomerase RNA molecule may exist largely as helix III-b. However, A76, predicted to be in the loop of helix III-b, would have to be highly structured to account for its strong protection from DEP. We cannot rule out the possibility that helix III-a is also present, and either helices III-a and -b are in equilibrium (fast in relation to DEP reaction kinetics) favoring helix III-b, or that the full-length RNA molecules exist as two populations, a predominant part of the population folding to form helix III-b, and the other forming III-a, concurrently.

The base sequence in and around helix III was predicted phylogenetically to adopt a pseudoknot (ten Dam *et al.*, 1991), and the oligoribonucleotide PKab, consisting of just the sequence from nucleotides 66 to 101 in *T.thermophila* in the absence of the rest of the telomerase RNA, does form one. The available evidence argues that the remainder of the full-length *T.thermophila* telomerase RNA, when free in solution, prevents the helix III region from adopting a full pseudoknot under the conditions employed.

## The templating domain: single-stranded but ordered

The previous model based on phylogenetic analysis suggested that the region between helices II and III (nucleotides 38-75 in T.thermophila) was single-stranded (Romero and Blackburn, 1991). However, the ability of part of this sequence to participate in a pseudoknot or helix III-b indicated that this region was more structured. In addition, much of the rest of this region showed reactivities to DEP, and to RNase V1 (Figure 2B) and RNase T1 that were consistent with a structure that is not entirely singlestranded. For example, in both RNAs the conserved stretch of five adenines (A50-54 in T.thermophila) at the 3'-end of the templating domain was poorly modified by DEP compared with adenines in other single-stranded regions (see Figure 1). This diminished DEP reactivity was observed even at a higher temperature (40°C). The extent of RNase T1 cleavage also differed between G59, 61 and 64 (see Figure 2D). Structural order in this region might have resulted from a tertiary folding effect or alternatively could be intrinsic to the particular base sequence. To distinguish between these possibilities, we prepared two template region oligoribonucleotides with the sequences of the region between helices II and III-a (nucleotides 37-75 in T.thermophila and 38-75 in G.chattoni), and determined their patterns of sensitivity to DEP and RNase V1, under conditions identical to those used to probe the full-length RNA. A pattern of DEP and V1 nuclease sensitivity similar to that in the whole RNAs would argue that the RNA sequence per se in the oligonucleotide is sufficient to adopt the structure observed in the full-length RNA. Conversely, a different pattern of reactivities would suggest that other telomerase RNA regions outside the templating domain affect its conformation. Figure 4 shows the result of the DEP probing of the T.thermophila template region oligoribonucleotide. In contrast to the full-length RNA (Figure 1A), adenines A40, 44, 45, 50-54, 58, 65 and 67 were all generally uniformly reactive (Figure 4). A similar result was observed for the corresponding positions of the G.chattoni sequence oligoribonucleotide (data not shown). Furthermore, for both T.thermophila and G.chattoni RNAs the positions of RNase V1 cleavages in the short oligonucleotides were different compared to their full-length telomerase RNA counterparts (data not shown). We conclude that while the templating region is likely to have in general single-stranded character, it is constrained into an ordered conformation by folding of the rest of the telomerase RNA.

To address further the folding properties in the templat-

Fig. 3. Conformational flexibility of helix III and 5' contiguous region from the *T.thermophila* telomerase RNA. (A) Schematic diagram showing the nucleotides from the phylogenetic model that participate in the proposed pseudoknot interaction. Inset shows the pseudoknot in the co-axially stacked helical form, with positions of base changes in the mutant RNA oligomers (predicted to disrupt pseudoknot formation) indicated. To facilitate comparison, the numbering of adenines in the oligomers is the same as in the full-length T.thermophila RNA. (B) DEP probing analysis of the conformational properties of oligoribonucleotides with the wild-type and mutant T.thermophila helix III region. Oligomers were folded and treated with DEP; the cleavage products were separated by 20% denaturing polyacrylamide gel electrophoresis. DEP reactions on PKab were performed under both unfolded (marked 'denaturing') and folded conditions. The former was predicted to be representative of the DEP pattern from the PKab molecule in a single-stranded state, and also to indicate positions of adenines. The results of the chemical modification reactions on folded PKa, PKb and PKO are shown. PKO was predicted to be disrupted for pseudoknot formation. Adenine positions are numbered. (C) Phosphorimager quantitation of <sup>32</sup>P-radioactive bands from adenines in each of the PK RNA oligomers shown in part B. The quantitation for the DEP reactivity of PKab at 60°C was obtained from a separate experiment (gel data not shown). The schematic beside each plot (PKab at 60°C, excepted) represents the predicted folded form of each oligomer, with numbers showing adenine positions, and putative stems (S) and loops (L) indicated. Stems and loops predicted to be disrupted are marked (e.g. 'Sb', 'Lb'). The dotted line between A83 and U93 in the schematic of PKa represents a base pairing interaction. (D) Comparisons of DEP modification patterns were performed quantitatively by laser densitometry of autoradiograms of the helix III region of the full-length T.thermophila RNA and the pseudoknot-forming PKab oligoribonucleotide. Both molecules were DEP treated under the same reaction conditions. Inset (right panel): schematic showing two possible RNA hairpin conformers of the helix III region from the full-length T.thermophila telomerase RNA (discussed in Table I). Adenine positions in each RNA structure are indicated. Arrows represent RNase V1 cleaved bases in the full-length RNA. Strong and weak cleavages are represented by large and small arrows, respectively. Guanines shown in plain or highlighted were uncleaved or cleaved, respectively, by RNase T1 in the full-length T.thermophila telomerase RNA.

ing region, preliminary gel mobility shift experiments were performed (described in Materials and methods), using the *in vitro* synthesized 159 nt *T.thermophila* telomerase RNA and two radiolabeled 15 nt antisense DNA oligonucleotides, one complementary to the templating region (nucleotides 47–61) and the other to part of helix II (nucleotides 26–40). A clear retarded species was observed for the first but not for the second oligonucleotide (data not shown). This suggests that the sequence comple-

.

mentary to the second oligonucleotide is sequestered in intramolecular helix II and hence that, to a first approximation, any folding of the template region is likely to involve fewer pairing interactions than in helix II.

### Helix IV and the GA bulge

The structural probing experiments on the full-length *in vitro* telomerase RNAs support the existence of helix IV (Figures 1 and 2). The GA dinucleotide bulge in helix



Adenine position	Base modifications <sup>a</sup>				Structure favored
	Helix IIIa (data from PKb)	Helix IIIb (data from PKb)	Helix IIIab (data from PKab RNA)	159 nucleotide RNA	
A67	+	+	+	+	_
A69	+	-	_	+	а
A70	+	-	_	+	а
A76	_	+	+	-	а
A79	_	+	_	+	b
A80	_	+	_	+	b
A83	_	_	_	-	-
A89	+	+	+	+	-
A90	+	+	+	+	-
A91	+	+	+	+	-
A93	_	+	_	+	b
A94	_	+	_	+	b
A100	+	+	+	+	_
A101	+	+	+	+	-

Table I. Comparison of observed chemical reactivity associated with putative RNA conformers of the helix III region of the *T.thermophila* telomerase RNA

The three RNA conformers (shown in Figure 3A and D) were analyzed to see how their observed DEP base modification (determined from the chemical reactivities of PKa, PKb and PKab) correlated with the observed chemical reactivity pattern observed from the full-length (159 nucleotide) RNA, to determine what structure was formed in the full-length *in vitro T.thermophila* telomerase RNA.

<sup>a</sup>Adenines unreactive (or poorly) reactive to DEP, are indicated by '-'; those in more single-stranded states are marked by '+' indicating modification by the chemical reagent. The last column indicates (when it is possible to differentiate between structures) which of the RNA conformers, namely, helix IIIa ('a'), IIIb ('b') or IIIab (ab; pseudoknot) is favored by modification at the particular adenine position.



Fig. 4. Diethylpyrocarbonate modification time course of the T.thermophila template oligoribonucleotide. Autoradiogram of 20% denaturing polyacrylamide gel showing time course of DEP modification of the folded, 3'-end-labeled T.thermophila template domain oligomer. Two in vitro transcribed 39 nucleotide long RNAs were generated (Milligan et al., 1987) which corresponded to the templating domains of the T.thermophila and G.chattoni telomerase RNAs. The sequence of the T.thermophila RNA oligomer contained the base sequence G37-C75 from the full-length RNA. The sequence of the G.chattoni oligoribonucleotide corresponded to G38-C75. Folding and modification conditions of oligomer RNAs were identical to the full-length telomerase RNAs. Positions of the adenines were identified by performing the chemical modification at 90°C (Peattie and Gilbert, 1980) and are indicated to the left of the panel. To be consistent, the numbering of adenines in the oligomer is the same as in the full-length T.thermophila RNA.

IV and the two G-C base pairs on either side of it are phylogenetically conserved in the telomerase RNAs of 12 tetrahymenine group ciliated protozoa (Romero and Blackburn, 1991; D.P.Romero, personal communication). It is known that bulged nucleotides can distort the helical axis of DNA and RNA helices (Bhattacharyya and Lilley, 1989; Hsieh and Griffith, 1989; Bhattacharyya et al., 1990). Moreover, Weeks and Crothers (1993) have argued that a bulged adenine can be reactive to DEP while still being stacked into a helix, due to widening of the major groove by the induced kink. Therefore, one interpretation of the structure probing data is that the dinucleotide bulge creates a kink, widening the major groove (Weeks and Crothers, 1993) of helix IV. Hence the adenine N7 of A122 (in T.thermophila and G.chattoni) is accessible and modified (Figure 1C). Fluorescence resonance energy transfer experiments on DNA and RNA helices indicate that bulges can introduce an appreciable kink (~60° for an A3 bulge; D.M.J.Lilley, personal communication). As a corollary to the helical distortion introduced by the GA bulge, the minor groove on the other side of the helix becomes narrower. Rich and co-workers (Auron et al., 1982) postulated previously that RNase V1 recognizes the minor groove of an RNA helix. Hence, the kink at the GA bulge in helix IV may reduce the extent of nuclease accessibility, by steric hindrance (on the concave side of the kink). Indeed lack of RNase V1 cleavage is observed at G146 and G147 residues opposite the bulge in both telomerase RNAs in this region (Figure 2B), and these residues remain uncleaved even after extensive V1 nuclease digestion (data not shown). We also have evidence using previously established gel electrophoretic methods (Bhattacharyya et al., 1990) that this particular GA bulge kinks an RNA helix containing part of the helix IV region of the T.thermophila RNA in vitro (A.Bhattacharyya and E.H.Blackburn, unpublished data). The structural probing properties of the GA bulge and neighboring G-C pairs were similar in the two telomerase RNAs. Thus the predicted kink introduced by the GA bulge in helix IV is conserved, and helix IV can be partitioned into helices IVa and b on either side of the kink (Figure 2B inset).

## Discussion

Understanding the three-dimensional structure formed by the telomerase RNA molecule, both free and in the ribonucleoprotein holoenzyme, is crucial for analyzing its specific recognition by telomerase proteins and the role of the RNA in telomeric DNA synthesis. Here we have shown that the naked telomerase RNA adopts an overall structure in solution consistent with the existence of four helical regions. The structure probing studies on the T.thermophila and G.chattoni telomerase RNAs converge on the same secondary structure and are generally consistent with the phylogenetic model (Romero and Blackburn, 1991). Thus, the base sequence of the RNA itself is likely to be an important determinant of the secondary structure of the telomerase RNA within the native ribonucleoprotein. This study provides information about the solution structure of the telomerase RNA and its conformational repertoire. The ability of the in vitro transcripts to adopt the predicted secondary structure indicates that base modifications are not essential for this structure. This is further supported by in vitro reconstitution of enzymatically active telomerase ribonucleoprotein with in vitro synthesized telomerase RNA (Autexier and Greider, 1994).

We propose that folding of the telomerase RNA leads to the creation of two broad functional domains, one largely helical (including helices I and IV), perhaps providing a scaffold for protein binding, and the second containing the templating region (together with protein) implicated more with the enzymatic functions. Evidence is presented for three new refinements to the model suggested previously: first, although the region around helix III has the capacity to form a pseudoknot by itself, this region is probably in an equilibrium favoring one of two helices in the full-length free RNA. Second, the template domain is ordered. Third, the conserved GA bulge most likely introduces a kink into helix IV.

Based on previous work on bulged nucleotides (Hsieh Griffith, 1989; Bhattacharyya et al., and 1990; Bhattacharyya and Lilley, 1989), we anticipate that the conserved GA bulge in helix IV distorts the helical trajectory in a structurally conserved manner, and may serve as a recognition motif for protein binding. Nucleic acids bearing specific bends and kinks have previously been reported to bind proteins in a structure-dependent manner (Frederick et al., 1984; Bhattacharyya et al., 1991). For example, the HIV Tat peptide recognizes and binds the TAR bulge with specificity largely derived from the structure of the RNA (Puglisi et al., 1992, 1993). The HIV TAR RNA bulge has been shown to kink an RNA helix when free in solution (Riordan et al., 1992). As helical elements I, II and IV in the telomerase RNA are in close proximity to each other, they together may create a stable, base paired domain, providing motifs for protein binding. It is worth noting that the more divergent 191 nucleotide telomerase RNA moiety from the hypotrichous ciliate, Euplotes crassus (Shippen-Lentz and Blackburn, 1990), also possesses similar helical domains when visually folded (D.Shippen, personal communication), implicating these regions of the molecule in important functions in the ribonucleoprotein.

In contrast to the more defined structure adopted by helices I. II and IV, the properties of the region in and around helix III suggest it is more conformationally flexible. In T.thermophila and the other ciliate species in this group, this region of the RNA (nucleotides 67-101 in T.thermophila) can theoretically form one of three possible conformations (shown in Figure 3A and D): two alternative stem-loops or a pseudoknot. Here we have shown that, in isolation from the rest of the RNA, this base sequence by itself forms a pseudoknot; however, it does not do so when part of the full-length, naked RNA in solution. Instead, the data presented indicate that, under the conditions employed, the full-length T.thermophila RNA folds into one (or both) of stem-loops III-a or -b, but not the pseudoknot. As helix III-b is not expected to be favored in the full-length RNA based simply on the number of possible base pairs in its stem compared with helix III-a, the predominance of helix III-b must therefore be due to structural constraints imposed by the folding of other regions of the naked RNA. It is unlikely that the full-length RNA that we have studied is trapped in an unfavorable structure, as the phylogenetic and structure probing data are completely consistent in all other parts of the RNA, for both T.thermophila and G.chattoni. Therefore we suggest that the helix III region of the RNA is structurally dynamic compared to other parts. The finding that folding interactions with other parts of a fulllength RNA specifically prevent pseudoknot formation has not been reported previously. Evidence of conformationally dynamic behavior exists in other RNAs. For example, LeCuyer and Crothers (1993) have reported that the spliced leader (SL) RNA sequence from Leptomonas collosum can switch between two alternate and competing structural forms on a fast (<1 s) time scale. Moreover, neither a phylogenetic nor an RNA folding algorithm predicted this situation in the SL RNA (LeCuyer and Crothers, 1993).

We suggest a two-state model for the helix III region of the telomerase RNA in the ribonucleoprotein, in which the conformation of the RNA changes between pseudoknot and hairpin states, with the former being stabilized by protein interactions. Conformational change may result, for example, upon telomeric DNA primer binding, and thus be a transient structural change. This may be relevant to the enzymatic action of telomerase in that conformational changes in the RNA (and ribonucleoprotein) may be coupled to the DNA synthesis functions (associated with the templating region of the RNA) which take place in close spatial proximity to the helix III region. Crothers and colleagues have hypothesized that the rapid structural 'switching' between the two conformation states in the L.collosum SL RNA may be initiated by formation of a pseudoknot (LeCuyer and Crothers, 1994). It was suggested that the 'switching' may be coupled to a splicingrelated helicase function (LeCuyer and Crothers, 1994). Though no telomerase proteins have yet been purified, it will be important to determine if they, or DNA primer binding, stabilize the pseudoknot form, or otherwise alter the dynamic behavior.

Here we have presented evidence for an ordered

structure in the templating nucleotides and residues 3' to it. Several kinds of evidence from studies of the telomerase reaction have suggested that this region is important in enzymatic activity. The templating sequence of the telomerase RNA is required for the correct enzymatic addition of DNA sequences onto telomeric primers (Greider and Blackburn, 1989; Yu et al., 1990). Antisense oligonucleotides complementary to the region 3' to the templating nucleotides, when used as telomeric primers, specifically alter the catalytic properties  $(k_{cat})$  of the T.thermophila telomerase (Lee et al., 1994). A 'second' or 'lagging product' site on telomerase has been implicated in primer interactions with telomerase besides base pairing of the 3'-end of the primer to the template (Morin, 1991; Collins and Greider, 1993; Lee and Blackburn, 1993; Lee et al., 1994). These interactions can cause dramatic, sequence-specific primer effects on  $k_{cat}$ of the T.thermophila telomerase (Lee and Blackburn, 1993). The lagging product site is named by analogy to the mechanism of action of Escherichia coli RNA polymerase, in which there are also translocations of the catalytic site relative to the template (Das, 1993). It has been proposed that in telomerase, as in RNA polymerase, during rounds of elongation the newly synthesized product is displaced from the DNA template into the lagging product site in the enzyme, which also facilitates unwinding of the RNA-DNA helix (Lee et al., 1994). Therefore this niche may be formed by both RNA and protein components in close spatial proximity to the template, and include the RNA region 3' to the template.

Phylogenetic, computer (Romero and Blackburn, 1991; ten Dam *et al.*, 1991) and now structural methods all provide evidence for a conserved architecture of the telomerase RNAs of *T.thermophila* and related ciliates. We also hypothesize that there are two functional domains in the RNA, one primarily for protein binding and the other associated with enzymatic function. We anticipate that a thorough mutational analysis of the *T.thermophila* telomerase RNA will provide more information about the functional roles of this telomerase RNA structure in the enzyme mechanism.

## Material and methods

### Oligonucleotide synthesis

Deoxyoligonucleotides were synthesized using  $\beta$ -cyanoethyl phosphoramidite chemistry (Beaucage and Caruthers, 1981; Sinha *et al.*, 1984). Deprotected oligonucleotides were purified by denaturing polyacrylamide gel electrophoresis and electroelution.

### Cloning

Both the telomerase RNA genes from T.thermophila and G.chattoni were cloned into a pUC-based plasmid under the control of a T7 RNA polymerase promoter. The clone containing the T.thermophila gene (pDT7fg) was provided by D.P.Romero. A plasmid containing the G.chattoni telomerase RNA gene was constructed using PCR methods as follows. Two PCR primers were used: the upstream primer when annealed to a genomic subclone of the G.chattoni gene (pGC0.7L, kindly provided by D.P.Romero) did so in a manner which produced a 5'unhybridized tail, containing the T7 promoter and a single G residue at the +1 position, while the 3'-end of the primer hybridized to the immediate 5'-coding region of the gene. The downstream primer when similarly annealed to the same plasmid, possessed a 5'-single-stranded tail containing a DraI restriction site. The resulting amplified PCR fragment was cloned with blunt ends into the unique SmaI site of pUC18. The resulting clone, pGCT7, was sequenced (Sanger et al., 1977) on both strands of the cloned DNA to verify correct sequence. The DraI site facilitated a T7 run-off transcript to be made (after restriction cleavage) which would leave the same 3'-end as the wild type RNA. The full-length *in vitro* RNA from *T.thermophila* and *G.chattoni* differed only from the *in vivo* RNA sequence in having a 5'-initiating guanine residue instead of an adenine.

#### Transcription/labeling/folding of the RNA

Plasmids containing either the T.thermophila (pDT7fg) or the G.chattoni (pGCT7) telomerase RNA gene were digested to completion with FokI or DraI, respectively. Digestion was verified by agarose gel electrophoresis. Linearized plasmids were then used in in vitro run-off transcription reactions to generate the appropriate telomerase RNA molecules. The *in vitro* transcripts were 3'-end-labeled with  $[\alpha$ -<sup>32</sup>P]pCp (Amersham) and T4 RNA ligase (USB), and the full-length, uniquely end-labeled transcripts were purified on 0.4 mm thickness 6% denaturing polyacrylamide gels. Labeled RNAs were purified: visualized in the gel by autoradiography, excised, eluted overnight in water and ethanol precipitated to concentrate. The short oligoribonucleotides employed in this study were generated by in vitro run-off transcription of partially single-stranded synthetic DNA templates according to Milligan et al. (1987). All RNA molecules were folded in a high salt buffer containing 50 mM Tris-HCl pH 7.2, 150 mM NaCl, 20 mM MgCl<sub>2</sub> (Stern et al., 1988). For all RNAs used for DEP reactions, molecules were folded in a buffer exactly the same except that 50 mM HEPES pH 8.2 replaced the Tris-HCl, since the latter inactivates DEP.

#### Nuclease cleavage and chemical modification

Both RNase T1 (USB) and V1 (Pharmacia) digestions were performed in the Tris folding buffer (previously described) using enzyme concentrations and reaction times found empirically to produce single-hit kinetics. RNases T1 and V1 were used at dilutions of  $10^{-3}$  and  $10^{-2}$ , respectively, of the stock enzyme concentration, at a temperature, unless otherwise stated, of 30°C, for times between 1 and 20 min. The partial nuclease digestions were performed on folded telomerase RNAs. The positions of all the guanines (to provide mol. wt markers) were identified by performing RNase T1 digestion under denaturing conditions (i.e. at 55°C). Reactions were quenched either by ethanol precipitation or by plunging sample onto dry ice, and addition of 10 M urea loading dye prior to gel loading. Diethylpyrocarbonate (DEP; Sigma) reactions were performed as described in Peattie and Gilbert (1980), although reaction times and temperatures were changed to those described in the text. DEP modifies the N7 group of adenines (> guanines) forming a carbethoxylated chemical base adduct, the phosphate backbone is cleaved at the position of the adduct by treatment with aniline. By performing modification reactions under limiting conditions on the end-labeled, folded telomerase RNAs, and fractionating the cleaved products by denaturing gel electrophoresis, it is possible to infer whether adenines are in either single-stranded (DEP accessible) or in helical (protected from DEP modification) states. Sequence DEP reactions were generally performed at 90°C for 1 min, while standard DEP reactions under 'native' RNA folding conditions were carried out at 30°C, unless otherwise stated, for times between 10 and 70 min. Oligoribonucleotides were folded and treated with DEP under identical conditions to that used for the full-length T.thermophila RNA to facilitate comparison of results. DEP reactions carried out in no exogeneously added ions (when mentioned), were performed in HEPES buffer only. Products of nuclease digestion or chemical modification reactions were analyzed on 0.4 mm thickness 10% denaturing polyacrylamide gels or, in the case of the oligoribonucleotide RNAs studied, on 20% denaturing gels. Gels were dried down onto Whatman 3MM paper and autoradiographed at -80°C with Kodak XAR intensifying screens. Quantitation of radioactive gels was performed using Molecular Dynamics phosphor screens and Phosphorimager. Quantitation of autoradiographic film was carried out using an LKB laser densitometer and integrator.

#### Gel mobility shift assays

Gel shift experiments were performed using 10% native polyacrylamide gels (50:1 acrylamide:bis ratio), with buffer recirculation (buffer contained 90 mM Tris-borate pH 8.3, 50 mM NaCl), at a constant temperature of 10°C for 2 h. Binding was carried out under limiting concentrations of folded *T.thermophila* telomerase RNA, excess <sup>32</sup>Plabeled 15 nucleotide long antisense DNA oligonucleotide (10 000 c.p.m./binding reaction) in the presence of excess tRNA at 30°C for 60 min prior to gel loading. The sequences of antisense oligonucleotides 1 and 2 were 5'-dGTGATCTAAAAACCC, and 5'-dACTGTCAA-GATAATG, respectively. A clear retarded species was observed for oligonucleotide 1, indicating that it could bind well to its complementary sequence in the folded RNA. Oligonucleotide 2 gel shifted very poorly.

- - - - -

## Acknowledgements

We thank Alan Frankel, Jody Puglisi and members of the Blackburn laboratory, especially Karen Kirk, Geoffrey Kapler and Anat Krauskopf for critical discussion and comments on the manuscript. We are also grateful to Dan Romero for providing plasmids, Dorothy Shippen, David Lilley and Dan Romero for communicating results prior to publication. This work was supported by a Lucille P. Markey Charitable Trust Visiting Postdoctoral Fellowship to A.B. and a National Institutes of Health Grant GM 26259 to E.H.B.

## References

- Auron, P.E., Weber, L.D. and Rich, A. (1982) *Biochemistry*, **21**, 4700–4706. Autexier, C. and Greider, C.W. (1994) *Genes Dev.*, in press.
- Beaucage, S.L. and Caruthers, M.H. (1981) Tetrahedron Lett., 22, 1859– 1862.
- Bhattacharyya, A. and Lilley, D.M.J. (1989) Nucleic Acids Res., 17, 6821-6840.
- Bhattacharyya, A., Murchie, A.I.H. and Lilley, D.M.J. (1990) Nature, 343, 484–487.
- Bhattacharyya, A., Murchie, A.I.H., von Kitzing, E., Diekmann, S. and Lilley, D.M.J. (1991) J. Mol. Biol., 221, 1191–1207.
- Brierley, I., Dignard, P. and Inglis, S.C. (1989) Cell, 57, 537-547.
- Cheong, C., Varani, G. and Tinoco, I., Jr (1990) Nature, 346, 680-682.
- Collins, K. and Greider, C.W. (1993) Genes Dev., 7, 1364-1376.
- Counter, C.M., Avilion, A.A., LeFeuvre, C.E., Stewart, N.G., Greider, C.W., Harley, C.B. and Bacchetti, S. (1992) *EMBO J.*, **11**, 1921–1929.
- Counter, C.M., Hirte, H.W., Bacchetti, S. and Harley, C.B. (1994) Proc. Natl Acad. Sci. USA, 91, 2900–2904.
- Das, A. (1993) Annu. Rev. Biochem., 62, 893.
- Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J.-P. and Ehresmann, B. (1987) Nucleic Acids Res., 19, 9109–9128.
- Frederick, C.A., Grable, J., Melia, M., Samudzi, C., Jen-Jacobson, L., Wang, B.-C., Green, P., Boyer, H.W. and Rosenberg, J.M. (1984) *Nature*, **309**, 327–330.
- Greider, C. and Blackburn, E.H. (1987) Cell, 51, 887-898.
- Greider, C. and Blackburn, E.H. (1989) Nature, 337, 331-337.
- Haasnoot, C.A.G., Hilbers, C.W., van der Marel, G.A., van Boom, J.H., Singh, U.C., Pattbiraman, N. and Kollman, P.A. (1986) J. Biomol. Struct. Dynam., **3**, 843–857.
- Hastie, N.D., Dempster, M., Dunlop, M.G., Thompson, A.M., Green, D.K. and Allshire, R.C. (1990) *Nature*, 346, 866–868.
- Hsieh,C.-H. and Griffith,J.D. (1989) Proc. Natl Acad. Sci. USA, 86, 4833-4837.
- LeCuyer, K.A. and Crothers, D.M. (1993) Biochemistry, 32, 5301-5311.
- LeCuyer, K.A. and Crothers, D.M. (1994) Proc. Natl Acad. Sci. USA, 91, 3373–3377.
- Lee, M.S. and Blackburn, E.H. (1993) Mol. Cell. Biol., 13, 6586-6599.
- Lee, M.S., Gallagher, R.C., Bradley, J. and Blackburn, E.H. (1994) Cold Spring Harbor Symp. Quant. Biol., 58, in press.
- Leviev, I.G., Rodriguez-Fonseca, C., Phan, H., Garrett, R.A., Heilak, G., Noller, H.F. and Mankin, A.S. (1994) *EMBO J.*, **13**, 1682–1686.
- Lowman, H.B. and Draper, D.E. (1986) J. Biol. Chem., 261, 5396-5403.
- McPheeters, D.S., Stormo, G.D. and Gold, L. (1988) J. Mol. Biol., 201, 517–535.
- Milligan, J.F. et al. (1987) Nucleic Acids Res., 15, 8783.
- Morin,G.B. (1991) Nature, 353, 454-456.
- Noller, H.F. (1991) Nature, 353, 302-303.
- Peattie, D.A. and Gilbert, W. (1980) Proc. Natl Acad. Sci., 77, 4679-4689.
- Pleij,C.W.A., Rietveld,K. and Bosch,L. (1985) Nucleic Acids Res., 13, 1717–1731.
- Powers, T. and Noller, H.F. (1991) EMBO J., 10, 2203-2214.
- Puglisi, J.D., Wyatt, J.R. and Tinoco, I., Jr (1988) Nature, 331, 283-286.
- Puglisi,J.D., Wyatt,J.R. and Tinoco,I.,Jr (1990) J. Mol. Biol., 214, 437-453.
- Puglisi, J.D., Tan, R., Calnan, B.J., Frankel, A.D. and Williamson, J.R. (1992) Science, 257, 76–80.
- Puglisi, J.D., Chen, L., Frankel, A.D. and Williamson, J.R. (1993) Proc. Natl Acad. Sci. USA, 90, 3680–3684.
- Riordan, F.A., Bhattacharyya, A., McAteer, S. and Lilley, D.M.J. (1992) J. Mol. Biol., 226, 305–310.
- Romero, D.P. and Blackburn, E.H. (1991) Cell, 67, 343-353.

- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl Acad. Sci. USA, 74, 5463–5467.
- Shippen-Lentz, D. and Blackburn, E.H. (1990) Science, 247, 546-552.
- Sinha, N.D., Biernat, J., McManus, J. and Koster, H. (1984) Nucleic Acids Res., 12, 4539–4557.
- Stern, S., Moazed, D. and Noller, H. (1988) Methods Enzymol., 164, 481-489.
- ten Dam, E., van Belkum, A. and Pleij, C.W.A. (1991) Nucleic Acids Res., 19, 6951.
- Tuerk, C. et al. (1988) Proc. Natl Acad. Sci. USA, 85, 1364–1368.
- Varani, G., Cheong, C. and Tinoco, I., Jr (1991) Biochemistry, 30, 3280-3289.
- Weeks, K.M. and Crothers, D.M. (1993) Science, 261, 1574-1577.
- Woodcock, J., Moazed, D., Cannon, M., Davies, J. and Noller, H.F. (1991) *EMBO J.*, **10**, 3099–3103.
- Wyatt, J.R. and Walker, T. (1989) Nucleic Acids Res., 17, 7833-7842.
- Yu,G.-L. and Blackburn,E.H. (1991) Cell, 67, 823-832.
- Yu,G.-L., Bradley,J.D., Attardi,L.D. and Blackburn,E.H. (1990) Nature, 344, 126–132.
- Received on August 1, 1994; revised on September 12, 1994