Annu. Rev. Biochem. 1992. 61:113–29

TELOMERASES

E. H. Blackburn

Department of Microbiology and Immunology, University of California, San Francisco, California, 94143-0414

KEY WORDS: telomere synthesis, telomerase RNA, *Tetrahymena*, reverse transcriptase

---

CONTENTS

Perspectives 113
Telomeric DNA Sequences 114
The Telomerase Reaction 115
The RNA Moiety of Telomerase 116
The Mechanism of Telomeric DNA Synthesis by Telomerase 120
Processivity of Telomerase in Vivo 124
Primer Specificity of Telomerase 125
Concluding Remarks and Future Directions 128

Perspectives

The DNA at the ends of the linear eukaryotic chromosomes usually consists of tandemly repeated simple sequences. In general, the presence of this simple-sequence telomeric DNA at the chromosomal termini is essential for chromosome stability. Telomeric DNA is unusual in that one strand is synthesized by a ribonucleoprotein enzyme, telomerase, which is distinct from the conventional DNA replication machinery. The telomeric DNA sequence is specified by copying an RNA template sequence within the RNA moiety of telomerase. Telomerases appear to be widespread among eukaryotes, as judged, first, by the evolutionary conservation of telomeric DNA structure and in vivo behavior among eukaryotes, and, more directly, by the identification of telomerase activities from diverse eukaryotes. This review examines the specialized mechanism of telomeric DNA synthesis by telomerase. The structure and function of telomeres and telomeric DNA, and the role of

0066-4154/92/0701-0113$02.00

telomerase in vivo, have been the subjects of previous reviews and are not covered in detail here. The reader is referred to (1–4) for such reviews.

Telomerase is a DNA polymerase that can be classified as a reverse transcriptase, because its mechanism of action involves the copying of an RNA template into DNA. It is an unusual reverse transcriptase, however, because unlike the conventional, purely protein reverse transcriptases found in systems ranging from retroviruses to prokaryotes, it is a ribonucleoprotein which contains its own RNA template as an integral part of the enzyme. The RNA component of telomerases from several species has been identified, sequenced, and a secondary structure model proposed, and the mechanism of telomerase has been investigated both in vitro and in vivo. The protein component(s) of telomerases have to date resisted unambiguous biochemical identification. Thus the relationship of telomerase to conventional reverse transcriptases remains to be elucidated. In particular, whether the RNA component of telomerase plays roles in the action of telomerase besides being the template for telomeric DNA synthesis is a question of considerable interest, from the standpoints of both enzyme (protein or RNA) mechanisms and the evolution of telomerase.

Telomeric DNA Sequences

A given eukaryotic species has a characteristic telomeric DNA sequence, although the same sequence may occur in more than one species. Telomeric repeats can be regular, consisting of perfect tandem repeats of a fixed repeat unit sequence, or irregular, consisting of length or sequence variations of a basic repeat unit (1). However, regardless of whether the repeat units are perfect or irregular, one strand of the tandemly repeated simple telomeric DNA sequences is characterized by containing clusters of G residues, giving the two telomeric DNA strands a composition bias such that there is a G-rich and a C-rich strand. These strands have an invariable orientation with respect to the chromosome end: the G-rich strand is found at the 3' end of each chromosomal DNA strand, and at each end this G-rich strand protrudes 12–16 nucleotides beyond the complementary C-rich strand, at least in the various species in which it has been possible to analyze it (5, 6). Examples of G-rich strand repeat units include the regular sequences TTGGGG, found in the ciliates *Tetrahymena* and *Glaucoma*, TTTTGGGG, in the ciliates *Euplotes* and *Oxytricha*, and TTAGGG, found in humans and other vertebrates as well as some protozoans and molds. Examples of irregular repeats include TG_{1-3} and GGG(G/T)TT, characteristic of the yeast *Saccharomyces cerevisiae* and the ciliate *Paramecium*, respectively (1). (Unless specified otherwise, these and other nucleic acid sequences in this review are written in the usual 5' to 3' direction.)

The Telomerase Reaction

The telomerase of a given species synthesizes the G-rich strand DNA sequence characteristic of its species. Telomerase requires a DNA primer, to which telomeric repeats are added by polymerization in the usual 5' to 3' direction, using the appropriate deoxynucleoside triphosphates (dNTPs) as substrates (7). As described below, the most efficient primers in vitro consist of a few repeats of a single-stranded G-rich telomeric or telomere-like DNA sequence. These correspond to the G-rich protruding strand of natural telomeres (5, 6). Blunt-ended fully duplex DNAs are not used as primers (M. Lee, E. Blackburn, unpublished work). A typical example of an in vitro reaction carried out by the telomerase from *Tetrahymena thermophila*, in which the primer is the DNA oligonucleotide GGGGGGTTGGTT, can be depicted thus:

$$\underline{\text{GGGGTTGGGGTT}} + 4n\ \text{dGTP} + 2n\ \text{TTP} \rightarrow \underline{\text{GGGGTTGGGGTT}}(\text{GGGGTT})_n + 6n\ \text{pyrophosphate.} \quad 1.$$

The primer supplied to the reaction is the underlined sequence. Other than those incorporated into the telomeric sequence, no other deoxynucleoside (or ribonucleoside) triphosphates, or nucleotide cofactors, are required for the reaction (7), and under typical in vitro conditions n can be in the hundreds (8).

RECOGNITION OF THE PRIMER 3' END SEQUENCE An aspect of the telomerase reaction that was important for determining its mechanism is the recognition of the 3' end of the primer, such that the appropriate next nucleotides are added to complete a telomeric repeat unit. It was found that primers with, for example, a TTGGG, a GGTT, or TTGG 3' end were extended by the *Tetrahymena* telomerase to produce respectively TTGGGGTTGGGGTT.., GGTTGGGGTTGG.., and TTGGGGTTGGGG..; i.e. in each case the next appropriate nucleotides were added to complete some permutation of a perfect TTGGGG repeat that included the 3' end of the primer (8, 9).

Analogous results were obtained subsequently with the telomerase activities isolated from the ciliates *Oxytricha* and *Euplotes* (10, 11), and from human tissue culture cells (12, 13). For example, when supplied with a primer with a GGTT 3' end, the telomerase activities from *Euplotes* or *Oxytricha* produce the reaction product GGTTTT$(\text{GGGGTTTT})_n$. Note that whereas this GGTT 3' primer is extended by the *Tetrahymena* telomerase with nucleotides beginning with four G residues, the *Euplotes* or *Oxytricha* activity extends the same primer first by two T residues, which completes a T_4G_4 repeat unit that includes the primer 3' end sequence. Similarly, the human telomerase activity, which synthesizes AGGGTT repeats, was found to ex-

tend such a primer to produce GGTTAGGGTTAGGG, the first added nucleotides thus completing a TTAGGG telomeric repeat unit (12, 13).

The basis of the 3' end recognition of the primer by telomerase became clear when it was found that the *Tetrahymena* telomerase contained an RNA moiety that had within it the template sequence for telomeric G-rich strand synthesis.

The RNA Moiety of Telomerase

PRIMARY AND SECONDARY STRUCTURE The *Tetrahymena* telomerase has an apparent molecular size between 200 and 500 kDa, as judged by its elution position from a gel filtration column under various salt conditions (9). The first indication that a nucleic acid was an intrinsic part of telomerase and was essential for its enzymic activity came from the finding that partially purified telomerase activity in *Tetrahymena* extracts was sensitive to nuclease activities as well as to protease or heat. From experiments in which the enzyme was pretreated with either micrococcal nuclease or RNase A, the nuclease inactivated, and then telomerase activity assayed, it was deduced that an RNA was required for activity (9). The telomerase activities in *Oxytricha*, *Euplotes*, and human extracts were found to be similarly nuclease sensitive (10–12).

The RNA moiety of the *Tetrahymena* telomerase was initially identified by its co-chromatography with telomerase activity through several column chromatographic fractionations (9). The sequence of this 159-nucleotide RNA species was determined, and was found to include the sequence 3' AACCCCAAC 5' (14). As described below, this sequence was shown to act as the template for TTGGGG repeat synthesis by telomerase. A prediction made from this work was that the telomerase RNAs from other species would each have a template sequence corresponding to its species-specific telomeric repeat sequence. This prediction has been borne out for the telomerase RNAs of the ciliates *Euplotes* (which has T_4G_4 telomeric repeats), *Glaucoma*, and five other *Tetrahymena* species (which have T_2G_4 telomeric repeats) (15, 16). The *Euplotes* telomerase RNA was first identified by its sensitivity to cleavage by RNase H directed by a DNA oligonucleotide containing T_4G_4 repeats (15). The telomerase RNAs of the *Tetrahymena* spp. and the related ciliate *Glaucoma* were initially identified by cross-hybridization of their single-copy genes to the *Tetrahymena thermophila* gene (15). These ciliate RNAs are the only telomerase RNA sequences that have been identified. By analogy and extension of the model for telomere synthesis deduced from the ciliate telomerases (see below), a template sequence has been suggested for the human telomerase RNA (13), although this RNA has not yet been identified.

The *Tetrahymena* telomerase RNA, like a variety of other small RNAs found in ribonucleoprotein complexes, is an RNA polymerase III transcript.

This was deduced from its lack of a 5' cap, the presence in its gene of a run of T residues encompassing the position of the 3' end of the RNA (14), and, most informatively, its sensitivity to α-amanitin in run-off transcription experiments with isolated nuclei (17). All the ciliate telomerase RNA genes have a pair of highly conserved upstream sequences, and no conserved Box A consensus sequence, suggesting that, like U6 RNA, their RNA polymerase III–mediated transcription is regulated by upstream cis-acting elements (15, 16). The regulation of telomerase RNA transcription has not been studied. However, when overexpressed in vivo, the steady-state level of accumulated telomerase RNA was unchanged from the normal situation, even though the transcription rate of the overexpressed genes, as judged by nuclear run-off experiments, was very high. These findings suggested that excess telomerase RNA transcribed but not assembled into the telomerase RNA complex was degraded (17).

In each sequenced ciliate RNA, the predicted telomeric C-rich putative template sequence is present, as shown in Figure 1 (14–16). The *Euplotes* telomerase RNA is 192 nucleotides in length [recent work has shown that the A residue thought to be upstream of this RNA (16), whose length was previously estimated to be 191 nucleotides, is the 5' nucleotide of this RNA; D. Shippen-Lentz, E. Blackburn, unpublished work]. This RNA contains the 15-nucleotide sequence 3' CCAAAACCCCAAAAC 5', complementary to nearly two full T_4G_4 repeats. However, as described below, only the 12 5' nucleotides of this sequence act as the template (16). The *Tetrahymena* spp. and *Glaucoma* RNAs, all ~160 nucleotides in length, contain a 22-nucleotide sequence that is absolutely conserved among this group of species and which includes the putative template sequence (Figure 1). Apart from the template, and a five-nucleotide sequence 5' to the template which is conserved between *Euplotes* and the distantly related *Tetrahymena* spp. group (see Figure 1), the primary sequences of the telomerase RNAs are highly divergent overall (14–16). This divergence prevents a reliable alignment from being made between the *Euplotes* and *Tetrahymena* RNAs, apart from the template domain. Among the *Tetrahymena* group of RNAs the maximum divergence is ~35%, allowing them to be aligned with each other and compensatory base-pair changes in putative secondary structures to be identified. Such comparison of the sequences of these RNAs allowed a phylogenetically supported secondary structure model to be deduced. Despite their high degree of primary sequence divergence, these RNAs have a strikingly well-conserved secondary structure (15).

THE TEMPLATE FUNCTION OF TELOMERASE RNA The first evidence indicating that the 3' AACCCCAAC 5' sequence in the *Tetrahymena* telomerase RNA acts as the template for synthesis of the complementary

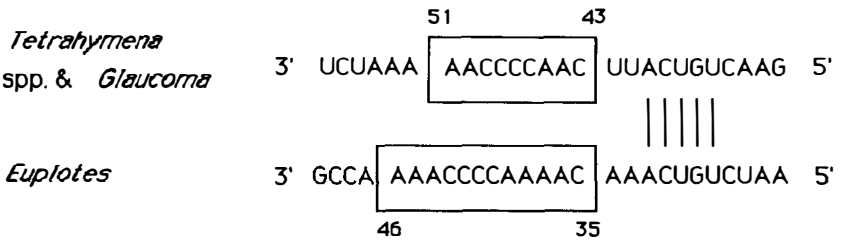

Figure 1 Sequences of the template domain (boxed) and surrounding region of telomerase RNAs from *Tetrahymena* spp. and *Glaucoma* (top) and *Euplotes* (bottom). Nucleotide positions of the template in each RNA sequence are indicated by numbers (14–16).

telomeric TTGGGG repeats came from experiments designed to test the effect on telomerase activity of treatment with RNase H in the presence of DNA oligonucleotides complementary to the putative telomerase RNA sequence (14). Since RNase H will cleave RNA where it is base-paired to a complementary DNA sequence, the loss of telomerase activity in such an experiment indicated the corresponding RNA sequence was required for telomerase. It was found that RNase H cleavage directed by a DNA oligonucleotide whose 3' end was complementary to the 3'AACCCCAAC 5' sequence in the telomerase RNA specifically inactivated telomerase. Furthermore, this oligonucleotide competed with a $(TTGGGG)_4$ primer for telomerase, suggesting it was blocking access of the primer to a site necessary for telomerase activity. A DNA oligonucleotide whose 3' end was complementary to the adjacent sequence, but ending one nucleotide from the 3' end of the putative template sequence, was also extended by GGGGTT repeat addition, suggesting that base-pairing of a DNA oligonucleotide to the telomerase RNA in the vicinity of the template could allow the oligonucleotide to be utilized as a primer. Based on these results, and on the results with 3' primer recognition described above, it was proposed that the 3'AACCCCAAC 5' sequence in the *Tetrahymena* telomerase RNA is the template sequence, and a model was proposed for the telomerase mechanism (14). This model, shown in Figure 2 and described below, was verified in the experiments described next.

A prediction of the model shown in Figure 2 is that altering the template RNA sequence should result in synthesis of telomeric DNA with the corresponding altered sequence. This prediction was confirmed by experiments done in vivo (17). The template sequence of the cloned *Tetrahymena thermophila* telomerase RNA gene was mutated by site-directed mutagenesis to produce three different templates: one with an additional C residue, converting the CCCC sequence to CCCCC, which was predicted to specify the synthesis of G_5T_2 repeats, one with the A at position 44 (see Figure 1) substituted by G (predicting G_4TC repeats), and a third template sequence

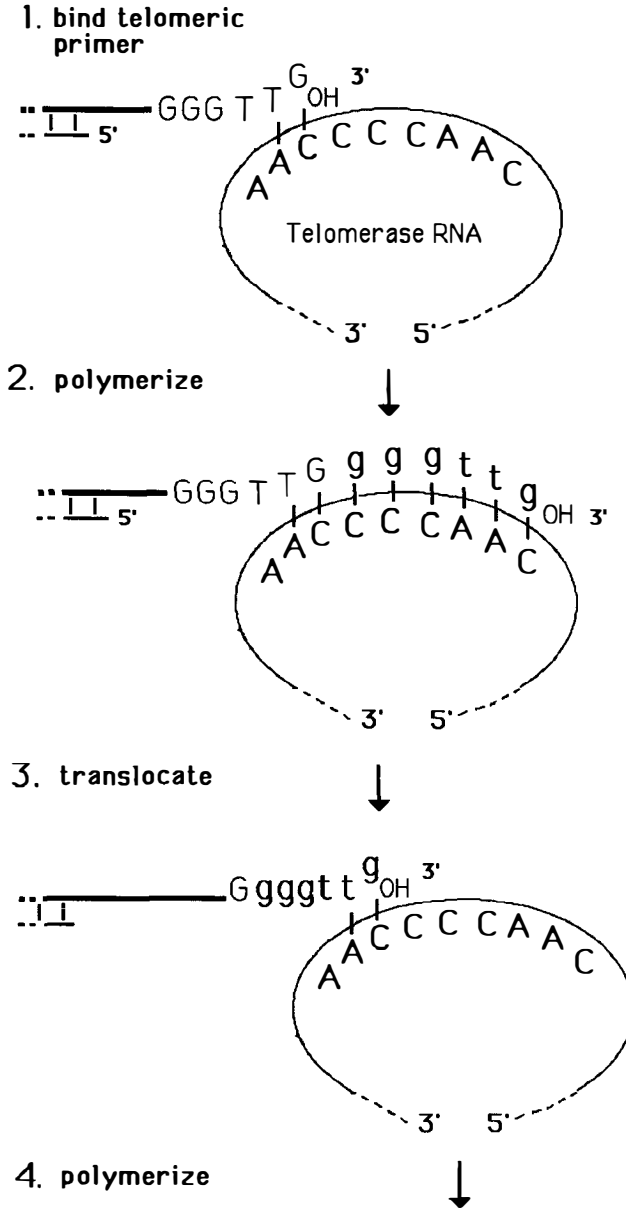

Figure 2 Synthesis of telomeric DNA by the ribonucleoprotein enzyme telomerase from *Tetrahymena*. See text and (14, 16) for explanation.

with the C at position 48 substituted by T (predicting GAG_2T_2 repeats). Each mutated telomerase RNA gene was overexpressed in *Tetrahymena* cells. The first two each resulted in synthesis in vivo of telomeres with the expected mutant telomeric repeat sequences. The C-to-T substitution at position 48, in contrast, resulted in telomere shortening, and none of the expected mutated repeat sequence was detectable in these cells (17). These results are intriguing because they suggest the possibility that this C residue plays a special role in the action of telomerase. This possibility is also interesting in light of a possible base modification of C residues 46, 47, and 48. This was deduced from the resistance of the polynucleotide backbone to cleavage with RNases at these positions, although these positions showed normal sensitivity to alkaline hydrolysis (14).

Independent evidence for the templating function of the RNA moiety of telomerase was obtained by in vitro experiments with the enzyme from *Euplotes* (16). As mentioned above, this enzyme synthesizes T_4G_4 repeats (10). A series of DNA oligonucleotides complementary to the region of the *Euplotes* telomerase RNA that is 3' of the putative templating domain (see Figure 1), and whose 3' ends extended towards, or varying numbers of nucleotides into, the template region, were tested for their ability to prime telomerase (16). In a manner reminiscent of the situation with the *Tetrahymena* enzyme, with which an oligonucleotide complementary to the region 3' of its template was able to prime repeat addition (14), these complementary oligonucleotides competed with a telomeric G-rich DNA primer, suggesting that they interacted with a similar region of the *Euplotes* telomerase. By determining which positions of the primer 3' ends allowed priming of telomeric repeat addition, the functional template domain of the *Euplotes* RNA was demarcated as shown in the boxed sequence in Figure 1 (16).

The Mechanism of Telomeric DNA Synthesis by Telomerase

Figure 2 shows the current model for the mechanism of telomerase (14, 16). The example shown is the *Tetrahymena* telomerase, for which in vivo proof for the template function exists, as described above. However, the 3' boundary of the template region has not been unequivocally deduced by, for example, the methods used to delineate the corresponding border in the *Euplotes* enzyme (14, 16). Otherwise, the model has been confirmed by findings made with the *Tetrahymena* and *Euplotes* enzymes in vitro and in vivo (14, 16, 17), and is supported by results obtained with telomerase activities from *Oxytricha* and human cells (11–13). In this model for synthesis of telomeric DNA:

1. The 3' nucleotides of the terminal chromosomal G-rich overhang in vivo, or of a single-stranded DNA primer oligonucleotide in vitro (thick line;

shown arbitrarily as a TTG 3'end), base-pair with the telomere-complementary sequence in the telomerase RNA.
2. The chromosomal end is extended by polymerization of dGTP and dTTP using the RNA as a template, resulting in the addition of six telomeric nucleotides.
3. The extended DNA terminus unpairs from its RNA template and is repositioned on the 3' portion of the template, becoming available for another round of elongation by telomerase.

Several points about this model merit further consideration.

DNA POLYMERIZATION A round of telomeric repeat synthesis involves a six-nucleotide polymerization cycle in the case of *Tetrahymena* and human telomerases, and an eight-nucleotide cycle in the case of the *Euplotes* and *Oxytricha* enzymes. Inherent in this model is the idea that telomerase has conformational flexibility, since the polymerization active site and the template region of the RNA, which is physically associated with the telomerase protein(s) in the telomerase ribonucleoprotein complex, must move in relation to each other as each template position is copied (16). Thus, each polymerization step is expected to be distinct, as at each position the primer-template will have a particular spatial relation to the telomerase active site. Indeed in vitro, telomerases have characteristic patterns of pausing points along the template, which can be different for different telomerases (9–12, 14, 16). For example, under identical reaction conditions (including dNTP concentration), the telomerases of *Oxytricha* and *Euplotes* predominantly pause respectively at the fourth T and the fourth G (corresponding to position 40 on the template; see Figure 1) during TTTTGGGG repeat synthesis (D. Shippen-Lentz, E. Blackburn, unpublished work).

Results obtained with a mutated telomerase RNA gene of *Tetrahymena* in vivo support the idea that the spatial relationship of the template with the rest of the telomerase ribonucleoprotein is important (18). The *Tetrahymena* telomerase normally synthesizes perfect T_2G_4 hexanucleotide repeats. When the template sequence was expanded with an extra C residue as described above, in addition to synthesis of telomeres with the expected G_5T_2 repeats, under certain circumstances irregular repeats, with the sequence $G_{6-8}T_2$, were synthesized in vivo. From this result it was proposed that expanding the template promoted slippage of the active site on the run of C residues (18).

TRANSLOCATION One of the most striking features of the telomerase reaction is that it involves not only copying of an internal template, but also an efficient translocation event which occurs after the last [5' most] residue of the template has been copied into DNA (see Figure 2). The efficiency of the translocation step has been deduced from studies showing that telomerase is

processive, at least in vitro (8, 10, 19) (M. Lee, E. Blackburn, unpublished results) (but see below for a discussion of the in vivo situation). In vitro, telomerase will initiate synthesis on a telomeric sequence DNA primer and, in the presence of an excess of the same primer, or a high concentration of a challenging primer added to the reaction, continue to elongate the first primer for up to hundreds of nucleotides before dissociation (19) (M. Lee, E. Blackburn, unpublished results). Because of this high processivity, the rate of elongation of an individual primer molecule before it dissociates can be determined by measuring the length of the elongation products of short reactions fractionated on DNA sequencing gels. Depending on reaction conditions, this growth rate, which is determined by the rates of the translocation as well as all six polymerization steps, has been determined to be ~30 to 70 nucleotides polymerized per minute for the *Tetrahymena* enzyme (19) (M. Lee, E. Blackburn, unpublished results). Similar rates have been estimated for the *Euplotes* activity (D. Shippen-Lentz, E. Blackburn, unpublished results). Despite this processivity, under certain conditions the translocation step is marked by a significant pause in both the *Tetrahymena* and *Euplotes* telomerase elongation reactions (14, 16). Detailed analyses of the kinetics and products of the *Tetrahymena* telomerase reaction with various DNA primers suggest when dissociation of the elongating primer does occur, it is most likely to be at the translocation step (M. Lee, E. Blackburn, unpublished results). This finding is reminiscent of the findings with *Escherichia coli* RNA polymerase, in which termination (transcript-template dissociation) is mechanistically closely linked to pausing (20).

The translocation position clearly does not correspond to the 5' end of the telomerase RNA—i.e. it is not determined by "run-off" from the RNA template. In the *Tetrahymena* telomerase RNA the translocation position (position 43) is followed by two U residues (14), so it could be argued that this telomerase, which normally only utilizes dGTP and TTP, cannot incorporate dATP efficiently, and hence polymerization is prevented from continuing along this RNA sequence. [It should be noted, however, that when the RNA template was mutated to include an rG residue, this telomerase used dCTP to synthesize G_4TC repeats in vivo (17).] However, the corresponding 5' end of the C- and A-containing template in the *Euplotes* telomerase RNA (position 35) is followed by three more upstream A residues, then a C residue, but none of these are copied (16). Thus there must exist in telomerase mechanisms to (*a*) prevent copying of nucleotides upstream of the template domain, and (*b*) promote both dissociation of the newly elongated primer from the template and its repositioning on the 3' portion of the template, preparatory to another round of template copying. These mechanisms could be coupled or distinct.

Models of how the 5' end of the template sequence is defined fall into two

classes. The first involves simple steric hindrance: it can be proposed that the telomerase RNA sequence 5' to the template is blocked by tightly bound telomerase protein, so that further polymerization is prevented. Consistent with this idea, experiments determining the accessibility of oligonucleotides complementary to the *Tetrahymena* telomerase RNA to various regions of this RNA indicated that only the region including the template was accessible (14). The rest of the RNA is presumably buried by protein in the ribonucleoprotein complex. But simple steric hindrance does not explain the efficiency with which the elongated primer is dissociated from the template in order to be repositioned. Implicit in the model for the mechanism of telomerase shown in Figure 2, and in the above discussion, is the idea that a DNA-RNA helix builds up along the template. As shown in Figure 2, up to nine Watson-Crick base-pairs of an RNA-DNA helix could exist at the end of a cycle of templated synthesis on the RNA template of the *Tetrahymena* telomerase. With the *Euplotes* telomerase, from results obtained in vitro using primers complementary to the region 3' of the template and extending into the template, it was deduced that a minimum of 11 base-pairs must be dissociated after the initial round of elongation of such a primer, and eight base-pairs must be dissociated in each subsequent elongation round, as each GGGGTTTT repeat is added (16). Is such a long helix in fact made? If so, unwinding it is expected to require energy. The possibility that a DNA-RNA helicase activity is either intrinsic to, or associated with, telomerase needs to be investigated. Although ATP is not required for telomerase activity, the possibility has not been excluded that the same dNTPs incorporated into the product are also used by telomerase in an energy-requiring step.

In an alternative type of model, as polymerization proceeds along the template, steric strain within the RNP is built up. This strain could prevent further polymerization once the 5' limit of the template is reached. The energy for building up such strain could come from the polymerization of the dNTPs and the formation of the DNA-RNA helix as the template is copied. It was further proposed that as a strained RNA structure builds up, the strain could be relieved by dissociation of the DNA-RNA helix when the translocation point is reached (15). In the secondary structure model for telomerase RNA, the template is located in an RNA loop whose ends are confined by base-pairing (15). It was proposed that the strain from building a DNA-RNA helix could result from distortion of this constrained loop (15). However, cleavage of the telomerase RNA sequence downstream from, but not including, the template by oligonucleotide-directed RNase H cleavage did not inactivate telomerase (14). This cleaved sequence is located within the template loop in the secondary structure model (15). Its cleavage would therefore be expected to prevent build-up of steric strain and, hence, translocation, if translocation depended on such strain. This result is still consistent with a

model in which steric strain is built up because the RNA is constrained by other interactions, such as RNA-protein interactions, with the rest of the telomerase particle.

Processivity of Telomerase In Vivo

As discussed above, telomerase is a highly processive enzyme in vitro. However, the sequence analysis of telomeres bearing two telomeric repeat unit variants strongly suggests that this is not the case in vivo. A *Tetrahymena* telomerase RNA gene with a mutated template specifying mutant G_5T_2 repeats (17, 18) was expressed in *Tetrahymena* cells along with the wild-type gene, which specifies G_4T_2 repeat synthesis. Cloned telomeres from these cells were found to be made up of a highly interspersed mixture of both mutant G_5T_2 and wild-type G_4T_2 repeats (18). Previous results had indicated that the two repeat sequences result from synthesis by different telomerase ribonucleoprotein molecules carrying RNAs with different template sequences (17). Hence the interspersion of the different repeat units provides information about the processivity of telomerase in vivo. The number of tandem repeats of the same sequence was most frequently one to four, suggesting that telomerase acts in a mostly distributive fashion in vivo (18). Thus, unlike the situation in vitro, telomerase in vivo may synthesize as little as one telomeric repeat following binding to a primer, and then dissociate so that the next repeat can be added by another telomerase molecule with a different template RNA. *Tetrahymena* telomeres in vivo appear to be largely if not completely complexed with telomere-binding proteins (21), and these could compete efficiently with telomerase for the newly elongated telomeric terminus after each round of template copying. In support of this idea, in vitro experiments show that competition occurs between the purified telomere-binding protein of *Oxytricha*, which binds the G-rich overhanging end of the telomere (22–24), and the *Tetrahymena* telomerase (D. Shippen-Lentz, E. Blackburn, C. Price, unpublished results). It has been proposed that the extended G-rich strand synthesized by telomerase is copied by lagging-strand synthesis in vivo (25). Polymerase-primase could therefore also compete for this G-rich strand and curtail telomerase elongation. Alternatively, the difference in processivity in vitro and in vivo could reflect a difference in the intrinsic reaction mode of the enzyme under the two different kinds of conditions.

In some species, the wild-type telomeres are mixtures of two repeat unit sequences of identical length, but with two alternative nucleotides in a given position in the repeat. In both *Paramecium* [G_4T_2 and G_3T_3 repeats (1)] and *Plasmodium* [AG_3T_3 and AG_3T_2C repeats; (1)] the distribution of different repeats appears random, and a histogram of the interspersion of different

repeat types in the published telomeric sequences of these organisms is qualitatively very similar to that obtained in the *Tetrahymena* cells transformed with a mutated telomerase RNA gene (18). Results obtained in *Tetrahymena* suggest that *Paramecium* and *Plasmodium* contain two telomerase RNAs, with two alternative nucleotides at one position in the template sequence. Specifically, in *Plasmodium*, one telomerase RNA would have an rG and the other an rA at the position where either a dC or a dT is incorporated in the two repeat types, AG_3T_2C or AG_3T_3. The *Paramecium* telomerase RNAs would have either an rC or rA in the template position specifying the alternative dG or dT in the G_4T_2 and G_3T_3 repeats.

Primer Specificity of Telomerase

PRIMER SPECIFICITY IN VITRO Telomeric DNA sequences are characterized by a highly conserved G-richness, in the form of clusters of G residues, on the strand that is extended by telomerase (1). This conservation could result from (as yet undefined) specific structural requirements for the functioning of telomeric DNA, which are uniquely met by clusters of G or C residues on one or the other strand. One such function could be some aspect of telomere synthesis by telomerase, since this mechanism appears to be highly conserved (1–3). The only hints of a mechanistic reason for the necessity of telomerase to synthesize G-rich DNA are the observations that mutating the C residue at position 48 in the template of the *Tetrahymena* telomerase RNA appears to inhibit telomere synthesis in vivo (17), and this particular C residue was one of the three C residues which are apparently base-modified (14). These findings suggest the possibility that these template C residues play a special base-specific role in the telomerase reaction, for which another base cannot substitute, and that the synthesis of G-rich telomeric DNA is simply the consequence of the requirement for these C residues in the template. However, currently no definitive data exist to refute or support this possibility.

A feature of telomerases that appears to be conserved is their efficiency of utilization of G-rich DNA as a primer in vitro. This property was proposed to underlie the recognition of G-rich but heterologous telomeric sequences that occurs in vivo [reviewed in (1–3)]. Again, this recognition could be a secondary result of the other possible reasons for G-rich telomeric DNA mentioned above, or it could reflect a more fundamental aspect of the telomerase enzyme. Therefore it is of interest to define which aspects of the telomerase reaction are influenced by G-richness of the primer. Recognition of G residues at the 3' end sequence of the primer can be understood based on their base-pairing the RNA template sequence (Figure 2). However, extensive analysis also indicates that additional sequence features of primers, besides

the 3' end sequence, determine their efficiency of utilization by telomerase (7–9, 16); M. Lee, E. Blackburn, unpublished results). Interestingly, although in vitro G-rich DNA sequences of sufficient length or concentration can form respectively intra- or intermolecular structures stabilized by "G-quartets" (26–29), the G-quartet form of DNA is strongly disfavored as a primer by the *Oxytricha* telomerase (30), and apparently also by the *Tetrahymena* enzyme (31; M. Lee, E. Blackburn, unpublished results).

The initial characterization of the telomerase activity of *Tetrahymena* (7, 9) suggested that G-rich primers were utilized by telomerase much more efficiently than non-G-rich primers of comparable length. For the *Tetrahymena* and *Euplotes* activities this was true even for long primers lacking G-richness except for a ..GGG or a ..GG 3' end (7, 9, 16), which would enable them to base-pair appropriately with the RNA template sequence. However, the results were obtained with crude telomerase preparations contaminated with nuclease and other unknown DNA-binding activities, making it difficult to attribute differences in primer utilization unambiguously to telomerase. These comparisons were made using high (micromolar) primer concentrations, and differences in binding versus k_{cat} could not be distinguished. With the further purification of telomerases from *Tetrahymena* and human cells, some of these problems have been minimized and more valid comparisons between primers can be made. Partial purification has led to increased telomerase activities, and allowed priming activity of short but template-complementary oligonucleotides such as the hexanucleotide TTGGGG, previously thought to be unable to act as a primer, to be detected (M. Lee, E. Blackburn, unpublished results). However, kinetic analysis shows that much higher primer concentrations are required for comparable extents of product formation than with longer G-rich primers (M. Lee, E. Blackburn, unpublished results). These results are consistent with a higher affinity for the latter primers. Using extensively purified *Tetrahymena* telomerase, this has been confirmed directly in primer competition assays (M. Lee, E. Blackburn, unpublished results). For single-stranded primers with the same length and 3' end sequence, and which are unable to form G-quartet structures under the assay conditions used, primers that have a G-rich sequence 5' to the template-complementary region at their 3' end are recognized much more efficiently by the *Tetrahymena* telomerase than non-G-rich primers (M. Lee, E. Blackburn, unpublished results). Whether this is true for the telomerases of other species is currently being debated, although the weight of current evidence is not inconsistent with this conclusion for these telomerases as well.

Results with different telomerases have often been difficult to compare because experiments have been carried out in different ways. However, using both the *Tetrahymena* and (unpurified) *Oxytricha* activities, the activity of various DNA primers has been compared in "abortive initiation" assays, in

which polymerization is confined to one or a few nucleotides by including only one dNTP. With the *Oxytricha* enzyme, the k_{cat}/K_m ratio (which measures substrate specificity of the enzyme) of T_4G_4 is higher than that for $(T_4G_4)_2$ (9). From this result the authors concluded that only base-pairing with the template was important for primer utilization, and G-richness did not increase primer utilization. However, as the k_{cat} for the abortive initiation reaction can be dominated by dissociation of the elongated primer product, these results are not informative about the more normal processive reaction in which large numbers of nucleotides are added. Contrasting results were obtained with the extensively purified *Tetrahymena* enzyme (M. Lee, E. Blackburn, unpublished results). With this telomerase, in the abortive initiation reaction, k_{cat}/K_m is higher for $G_4T_2G_4$ than for T_6G_4 or $TA_3T_2G_4$. More significantly, with the extensively purified *Tetrahymena* enzyme, in the processive reactions k_{cat}/k_m is higher for $G_4T_2G_4$ than for T_6G_4, T_2G_4, or $TA_3T_2G_4$. Hence in all these experiments the *Tetrahymena* enzyme exhibited a clear specificity for a G-rich oligonucleotide, compared with other primers of the same length and with the same (or more extensive) base-pairing possible with the template.

In a study with partially purified human telomerase activity, a somewhat different conclusion was reached: that interaction with the template was more important than the G-rich nature of the rest of the primer in determining its priming activity (13). However, in contrast to the *Tetrahymena* telomerase studies, these assays were carried out in very long reactions. Results with the *Tetrahymena* enzyme have shown that in long reactions, measurement of the extent of the reaction is complicated by dissociation and re-initiation events (M. Lee, E. Blackburn, unpublished results). Therefore the conclusions made with the different enzymes as to primer specificity have not yet been made under comparable conditions, so the question of species variation in the relative importance of different aspects of primer recognition remains open.

HEALING OF BROKEN CHROMOSOMES BY TELOMERASE IN VIVO All the available results with telomerases from different species are consistent in showing that some combination of template pairing and G-richness, and possibly other, as yet poorly defined, structural features of the primer, are important for its recognition by telomerase in vitro. Oligonucleotides consisting of the sequence of the position to which telomeric repeats are added to a healed human chromosome have been tested in vitro and shown to act as primers for the human and *Tetrahymena* telomerases (13, 31). However, this sequence had a G-rich portion, and in this respect was not unlike telomeric DNA sequences. In contrast, much indirect evidence indicates that under some circumstances in vivo, telomeric repeats are added to broken chromosome ends lacking any semblance of telomeric sequences (2). Direct in vivo

evidence showing that this addition is accomplished by telomerase has been obtained in *Tetrahymena* (18). These experiments showed that, at least in one developmental stage, in vivo telomerase adds telomeric repeats onto nontelomeric sequences lacking G-rich sequences or the ability to pair at their 3' end with the telomerase RNA. This conclusion came from experiments in which a *Tetrahymena* telomerase RNA gene with a mutated template specifying mutant G_5T_2 repeats was overexpressed in *Tetrahymena* cells, as described above. The mutant *Tetrahymena* telomerase RNA was used to monitor telomerase action in vivo, during the particular developmental stage in which a developmentally programmed chromosome fragmentation process occurs in this organism. Telomerase was shown to heal the nontelomeric ends generated by the chromosome fragmentation process, by adding the first telomeric repeats de novo onto these ends (18).

Concluding Remarks and Future Directions

Telomerase is a ribonucleoprotein polymerase whose specialized mechanism involves copying an internal RNA template sequence into DNA. Many interesting questions about this mechanism remain unanswered. Whether the RNA moiety of telomerase collaborates with the protein moiety of the enzyme in catalysis or substrate binding, in addition to the template interaction with the DNA primer, or has catalytic activity in the absence of protein, or simply supplies the template sequence, is not yet clear. The protein component(s) of the telomerase ribonucleoprotein have not been biochemically identified. Reconstitution of telomerase in vitro therefore awaits this identification. The mechanisms that accomplish the efficient translocation and resulting high processivity characteristic of the in vitro reaction of telomerase are yet to be determined.

ACKNOWLEDGMENTS

I thank Margaret Lee for critical reading of the manuscript. Support was provided by grants GM26259 and 32565 from the National Institutes of Health to E.H.B.

Literature Cited

1. Blackburn, E. H. 1990. *Science* 249: 489–90
2. Blackburn, E. H. 1991. *Nature* 350: 569–73
3. Blackburn, E. H., Szostak, J. W. 1984. *Annu. Rev. Biochem.* 53:163–94
4. Zakian, V. A. 1989. *Annu. Rev. Genet.* 23:579–604
5. Klobutcher, L. A., Swanton, M. T., Donini, P., Prescott, D. M. 1981. *Proc. Natl. Acad. Sci. USA* 78:3015–19
6. Henderson, E. R., Blackburn, E. H. 1989. *Mol. Cell. Biol.* 9:345–48
7. Greider, C. W., Blackburn, E. H. 1985. *Cell* 43:405–13
8. Blackburn, E. H., Greider, C. W., Henderson, E., Lee, M. S., Shampay, J., Shippen-Lentz, D. 1989. *Genome* 31: 553–60
9. Greider, C. W., Blackburn, E. H. 1987. *Cell* 51:887–98
10. Shippen-Lentz, D., Blackburn, E. H. 1989. *Mol. Cell. Biol.* 9:2761–64
11. Zahler, A. M., Prescott, D. M. 1988. *Nucleic Acids Res.* 16:6953–72
12. Morin, G. B. 1989. *Cell* 59:521–29

13. Morin, G. 1991. *Nature* 353:454–56
14. Greider, C. W., Blackburn, E. H. 1989. *Nature* 337:331–37
15. Romero, D., Blackburn, E. 1991. *Cell* 67:343–53
16. Shippen-Lentz, D., Blackburn, E. H. 1990. *Science* 247:546–52
17. Yu, G.-L., Bradley, J. D., Attardi, L. D., Blackburn, E. H. 1990. *Nature* 344:126–32
18. Yu, G.-L., Blackburn, E. H. 1991. *Cell* 67:823–32
19. Grieder, C. W. 1991. *Mol. Cell. Biol.* 11:4572–80
20. Fisher, R. F., Yanofsky, C. 1983. *J. Biol. Chem.* 258:8146–50
21. Budarf, M., Blackburn, E. H. 1986. *J. Biol. Chem.* 261:363–69
22. Gottschling, D. E., Zakian, V. A. 1986. *Cell* 47:195–205
23. Price, C. M., Cech, T. R. 1989. *Biochemistry* 28:769–74
24. Raghuraman, M. K., Dunn, C. J., Hicke, B. J., Cech, T. R. 1989. *Nucleic Acids Res.* 17:4235–53
25. Shampay, J., Szostak, J. W., Blackburn, E. H. 1984. *Nature* 310:154–57
26. Henderson, E. R., Moore, M., Malcolm, B. A. 1990. *Biochemistry* 29: 732–37
27. Sen, D., Gilbert, W. 1990. *Nature* 344:410–14
28. Sundquist, W. I., Klug, A. 1989. *Nature* 342:825–29
29. Williamson, J. R., Raghuraman, M. K., Cech, T. R. 1990. *Cell* 59:871–80
30. Zahler, A., Williamson, J. R., Cech, T. R., Prescott, D. M. 1991. *Nature* 350:718–20
31. Harrington, L. A., Greider, C. W. 1991. *Nature* 353:451–54