Cell, Vol. 59, 871–880, December 1, 1989, Copyright © 1989 by Cell Press

Monovalent Cation–Induced Structure of Telomeric DNA: The G-Quartet Model

James R. Williamson,* M. K. Raghuraman,† and Thomas R. Cech*†
*Department of Chemistry and Biochemistry
Howard Hughes Medical Institute
†Department of Molecular, Cellular and Developmental Biology
University of Colorado
Boulder, Colorado 80309

Summary

We have investigated the structures formed by oligonucleotides composed of two or four repeats of the telomeric sequences from Oxytricha and Tetrahymena. The Oxytricha four-repeat molecule ($d(T_4G_4)_4$ = Oxy-4) forms structures with increased electrophoretic mobility in nondenaturing gels containing Na^+, K^+, or Cs^+, but not in gels containing Li^+ or no added salt. Formation of the folded structure results in protection of a set of dG's from methylation by dimethyl sulfate. Efficient UV-induced cross-links are observed in Oxy-4 and the related sequence from Tetrahymena ($d(T_2G_4)_4$ = Tet-4), and join thymidine residues in different repeats. Models proposed to account for these data involve G-quartets, hydrogen-bonded structures formed from four guanosine residues in a square-planar array. We propose that the G-quartet structure must be dealt with in vivo by the telomere replication machinery.

Introduction

Telomeres, the structures at the ends of linear chromosomes, are involved in the maintenance and replication of the chromosomal ends, and appear to help mediate the organization of chromosomes within the nucleus. Telomeres are composed of stretches of repeated DNA sequences as well as associated proteins. The sequences of the telomeric DNA repeats from a wide variety of organisms adhere to the consensus sequence $d(T/A)_{1-4}dG_{1-8}$ (Blackburn and Szostak, 1984). Examination of a subset of the known telomeric sequences reveals a more specific consensus:

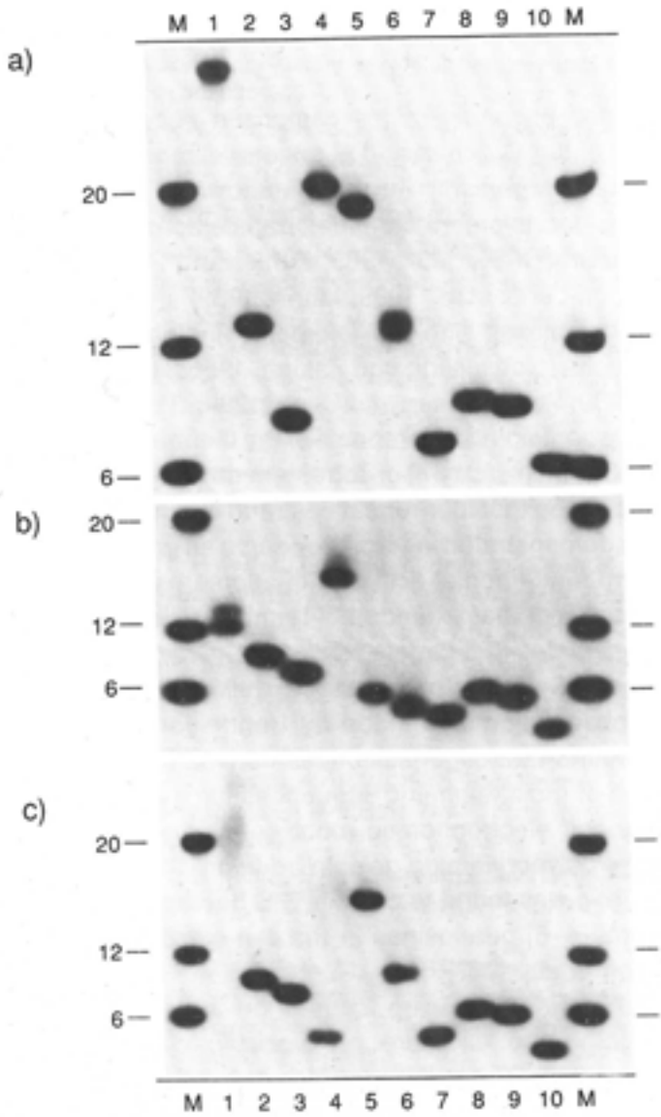
d(TTGGGG)	Tetrahymena	(Blackburn and Gall, 1978)
d(TTAGGG)	human	(Moyzis et al., 1988)
d(TTTAGGG)	Arabidopsis	(Richards and Ausubel, 1988)
d(TTTTGGGG)	Oxytricha	(Klobutcher et al., 1981)
d(TTTTGGGG)	Stylonychia	(Oka et al., 1980)
$d(T_{1-3}(T/A)G_{3-4})$	subset consensus	

Telomeric DNA is mainly double stranded but, in the cases examined most carefully, terminates at the 3′ end with two additional copies of the G-rich repeat (Klobutcher et al., 1981; Henderson and Blackburn, 1989). Single-stranded oligonucleotides composed of two or four repeats of the G-rich strand of telomeric sequences from a variety of organisms have been shown to serve as substrates for telomere terminal transferase (Greider and Blackburn, 1985, 1987, 1989; Zahler and Prescott, 1988; Shippen-Lentz et al., 1989; Morin, 1989) and to form specific complexes with the telomere binding protein from Oxytricha (Gottschling and Zakian, 1986; Raghuraman et al., 1989; Raghuraman and Cech, 1989). Thus, while molecules containing four repeats of the telomeric sequences from Oxytricha (Oxy-4) or Tetrahymena (Tet-4) are artificial in the sense that four-repeat G-strand overhangs have not been demonstrated in vivo, they serve as good substrates for two fundamental activities of telomeres. This may reflect the fact that extended 3′ terminal tails are common intermediates in telomere replication (Greider and Blackburn, 1985; Zahler and Prescott, 1988, 1989). An important observation was made by Henderson et al. (1987), who found that oligonucleotides derived from the G-rich strand of most telomeric sequences form structures with enhanced electrophoretic mobility in nondenaturing, or native, polyacrylamide gels. In addition, the oligonucleotide Tet-4 was found to contain G-G base pairs as well as several (4–6) guanosines in the *syn* conformation (Henderson et al., 1987).

Our initial studies are focused on the role of the dG-strings in the structure of telomeric oligonucleotides (Cech, 1988). The fact that strings of 3–4 dG's are present in most telomeric sequences suggests that this common sequence element may be the common structural element. Although the telomeric sequences in Tetrahymena and Oxytricha are different from the yeast sequence, they are recognized as functional telomeres when introduced into yeast (Szostak and Blackburn, 1982; Pluta et al., 1984; Pluta and Zakian, 1989). This also suggests that different telomeric sequences may adopt common structures.

Sequences rich in guanosine are known to exhibit unusual properties. Polyguanosine and polyinosine form four-stranded helices (Zimmerman et al., 1975), 5′-GMP at 1 M concentration is known to form tetrameric aggregates (Gellert et al., 1962; Miles and Frazier, 1972; Pinnavaia et al., 1975), and G-rich regions of immunoglobulin switch regions have been found to form parallel four-stranded structures (Sen and Gilbert, 1988). The stability of four-stranded structures adopted by poly(I) exhibits a strong dependence on the identity of the monovalent cation present (Miles and Frazier, 1978; Howard and Miles, 1982b). The fact that the T_m of Tet-4 has been reported as markedly higher (by about 40°C) in K^+ than in Na^+ (I. Tinoco, Jr., and C. C. Hardin, personal communication) suggested that similar structure might be forming from telomeric G-strings.

We have used a variety of chemical and physical techniques to assess the contribution of G-G base pairs to telomere structure. Modification of guanosine by dimethyl sulfate (DMS) is a direct probe of the accessibility of N7. For



Key for Figure 1:

1 -- $d(T_4G_4)_4$ Oxy-4
2 -- $d(T_4G_4)_2$ Oxy-2
3 -- $d(T_4G_4)$
4 -- $d(TTGIGG)_4$
5 -- $d(TTIGGG)_4$
6 -- $d(TTGGGG)_4$ Tet-4
7 -- $d(T_2I_4T_2I_4)$
8 -- $d(T_2G_4T_2I_4)$
9 -- $d(T_2I_4T_2G_4)$
10 -- $d(T_2G_4T_2G_4)$ Tet-2

M -- markers: dT_{20}, dT_{12}, T_2G_4

Figure 1. Electrophoresis of Telomeric Oligonucleotides in Nondenaturing Gels

(a) 0.5× TBE. (b) 0.5× TBE + 50 mM NaCl. (c) 0.5× TBE + 50 mM LiCl. All gels were run at 5°C until the bromophenol blue migrated 16 cm. In each gel, the mobility of the dT_{12} marker was approximately the same, migrating ~22 cm.

example, a Hoogsteen G-G base pair involves a hydrogen bond to N7 and should confer protection from methylation by DMS. Substitution of inosine for guanosine can probe the involvement of the exocyclic amino group of the guanine base in hydrogen bonding. Native gel electrophoresis is a powerful tool for assaying the presence of structure in oligonucleotides. Finally, we have found a UV-induced cross-link in the DNA that has provided structural clues and has permitted the study of the association of different monovalent cations with telomeric oligonucleotides. Based on the results of these experiments and on previous data (Henderson et al., 1987), we propose a model for the structure of telomeric oligonucleotides that involves hydrogen bonding similar to that observed for four-stranded poly(I) and poly(G).

Understanding the possible structures of telomeric DNA is the first step in addressing the relationship of structure to function. In a separate study, we will examine the relevance of the structured forms of telomeric DNA to Oxytricha telomere protein binding.

Results

Monovalent Cation Effect on the Electrophoretic Mobility of Telomeric Oligonucleotides

Oxytricha Oligonucleotides

A puzzling exception to the ability of telomeric oligonucleotides to form high mobility forms on native gels was Oxy-4, an oligonucleotide composed of four repeats of the telomere sequence from Oxytricha. Although it is very similar in sequence to Tetrahymena telomeres, only a small fraction (<10%) of the Oxy-4 molecules showed enhanced mobility in native gels (Henderson et al., 1987). The electrophoretic mobilities of oligonucleotides composed of one, two, and four repeats of the Oxytricha telomere sequence under conditions similar to those reported by Henderson et al. (1987) are shown in Figure 1a, lanes 1–3. Oxy-4 migrated as would be expected for a 32-mer oligonucleotide, and Oxy-2 ran as would a 16-mer. However, addition of 50 mM NaCl to the gel and the running buffer resulted in a dramatic increase in mobility of Oxy-4 and Oxy-2, as shown in Figure 1b, lanes 1 and 2. Two distinct bands of similar mobility were seen for Oxy-4.

Native gels were also run in 50 mM KCl and in 50 mM CsCl. Both Cs^+ and K^+ were competent to convert Oxy-4 and Oxy-2 to faster moving species, although the two bands observed for Oxy-4 in the sodium gel were combined into a single broad band (data not shown). Oxy-4 in 50 mM LiCl, however, was not completely converted to the high mobility form, and appeared as a smear of intermediate mobility (Figure 1c, lane 1). Oxy-2 exhibited the same mobility in Li^+ as in Na^+, K^+, and Cs^+ (Figure 1, lanes 2; data not shown). Thus by investigating the effects of monovalent cations on the gel mobility of Oxy-4, we have

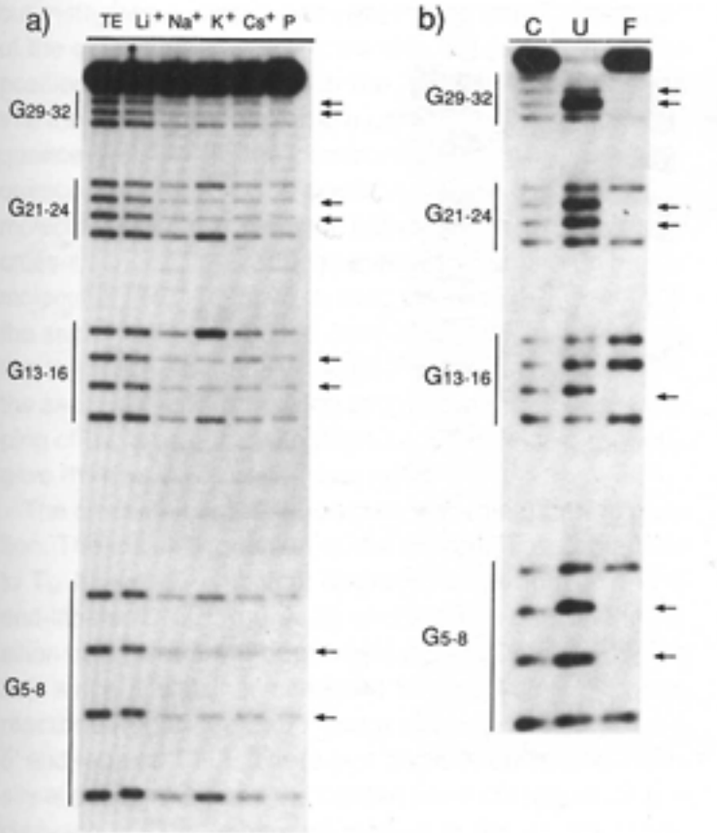


Figure 2. Methylation Protection and Interference of Oxy-4

(a) Autoradiogram of a denaturing 20% polyacrylamide gel showing methylation protection. Oxy-4 end labeled at the 5′ position was methylated by DMS in TE buffer or in TE with 50 mM LiCl, NaCl, KCl, or CsCl added, and then cleaved with pyrrolidine. Lane P shows background pyrrolidine cleavage in the absence of DMS treatment. Arrows indicate the positions of strongest protection observed in K^+ solution.

(b) Autoradiogram of a denaturing 10% polyacrylamide gel of the cleavage products of methylation interference of Oxy-4. Oxy-4 was uniformly methylated at 65°C, and the methylated DNA was subjected to electrophoresis in a native gel at 5°C in TBE with 50 mM NaCl. Folded and unfolded molecules were cut from the gel, eluted, and then cleaved with piperidine. Lane C, uniformly methylated DNA; lane U, unfolded, interfered DNA; lane F, folded, uninterfered DNA. Arrows indicate positions at which methylation strongly interferes with folding.

found mild conditions within the physiological range that convert the unstructured form of Oxy-4 into a structured, high mobility form.

Tetrahymena Oligonucleotides

In native gels containing only TBE as the buffer, Tet-4 has a mobility comparable to that of a 12-mer (Henderson et al., 1987; Figure 1a, lane 6). Addition of 50 mM LiCl enhanced its mobility slightly (Figure 1c, lane 6), but addition of 50 mM NaCl caused a large increase in mobility to an apparent size of a 6-mer (Figure 1b, lane 6). The Cs^+ and K^+ forms of Tet-4 were similar to the Na^+ form; however, the K^+ form exhibited, in addition, several bands of lower mobility (data not shown; Henderson et al., 1989).

Two variants of Tet-4 containing deoxyinosine (dI), $d(TTIGGG)_4$ and $d(TTGIGG)_4$, were tested and are shown in lanes 4 and 5, respectively, of Figure 1. Note that d(TTIGGG) is also related to the human telomere sequence d(TTAGGG). The substitution of four dI's for dG's apparently disrupted the structure of Tet-4, as shown by the electrophoretic mobilities of the oligonucleotide variants. Li^+ converted $d(TTGIGG)_4$ to a form more compact than the lithium form of Tet-4, but the sodium form was slower moving than the sodium form of Tet-4. $d(TTIGGG)_4$ showed a mobility comparable to Tet-4 in the Na^+ form but was slower migrating than Tet-4 in both the Li^+ form and in the absence of added salt.

The effect of inosine substitution on the mobility of Tet-2 was also examined with three sequence variants, $d(T_2G_4T_2I_4)$, $d(T_2I_4T_2G_4)$, and $d(T_2I_4T_2I_4)$, shown in Figure 1, lanes 7–10. Addition of 50 mM Li^+, Na^+, K^+, or Cs^+ enhanced the mobility of all four molecules, but there appeared to be no specific monovalent cation effect. All four sequences exhibited roughly the same mobility in all four salts, greater than or equal to that of a 6-mer, and the mobility was somewhat greater than in the absence of added salt.

Methylation Protection and Interference

The accessibility of the N7 of guanine to solvent can be probed chemically with DMS. When Oxy-4 was methylated at 0°C in TE buffer (10 mM Tris, 0.1 mM EDTA [pH 7.5]) or TE with 50 mM added LiCl, there was no apparent protection of the N7 of any of the guanines, as shown in Figure 2a. However, when 50 mM NaCl, KCl, or CsCl was added, protection was observed. In the presence of NaCl, methylation of all the guanines was reduced almost to the background pyrrolidine cleavage level (Figure 2a, lane P). In the presence of KCl or CsCl, the outer two dG's of each G-string were more accessible to DMS than were the central two. Control experiments on an oligonucleotide with no structure (i.e., d(TTGGGG)) indicated that the rate of methylation by DMS is unaffected by added salt (data not shown).

The ability of an N7 methyl group to disrupt folding was tested by uniform methylation of Oxy-4 at 65°C in TE, followed by electrophoresis of the methylated DNA in a native gel containing 50 mM NaCl. The folded and unfolded forms were recovered from the gel, subjected to strand cleavage with piperidine at the modified positions, and finally subjected to electrophoresis in a sequencing gel (Figure 2b). Since G_5, G_8, G_{13}, G_{15}, G_{16}, G_{21}, G_{24}, and G_{29} appear as strong bands in lane F (folded DNA), we conclude that methylation at these positions does not interfere with structure formation. Bands corresponding to the central two G's in each repeat are overrepresented in lane U (unfolded DNA) and, with the exception of G_{15}, are not represented in lane F. These correspond to guanines where methylation of N7 interferes with structure formation.

UV Cross-Linking

During our attempt to cross-link Oxy-4 to the Oxytricha telomere binding protein via ultraviolet (UV) radiation, a cross-link was found to form in the DNA by itself. The cross-linked DNA was observed as a faster migrating species in a denaturing 10% polyacrylamide gel (Figure 3a). The rate of cross-linking was independent of DNA concentration from 50 nM to 50 μM (data not shown); for this reason, in addition to its increased electrophoretic mobility, the cross-link is thought to be intramolecular.

To determine if the cross-link was specific to the folded

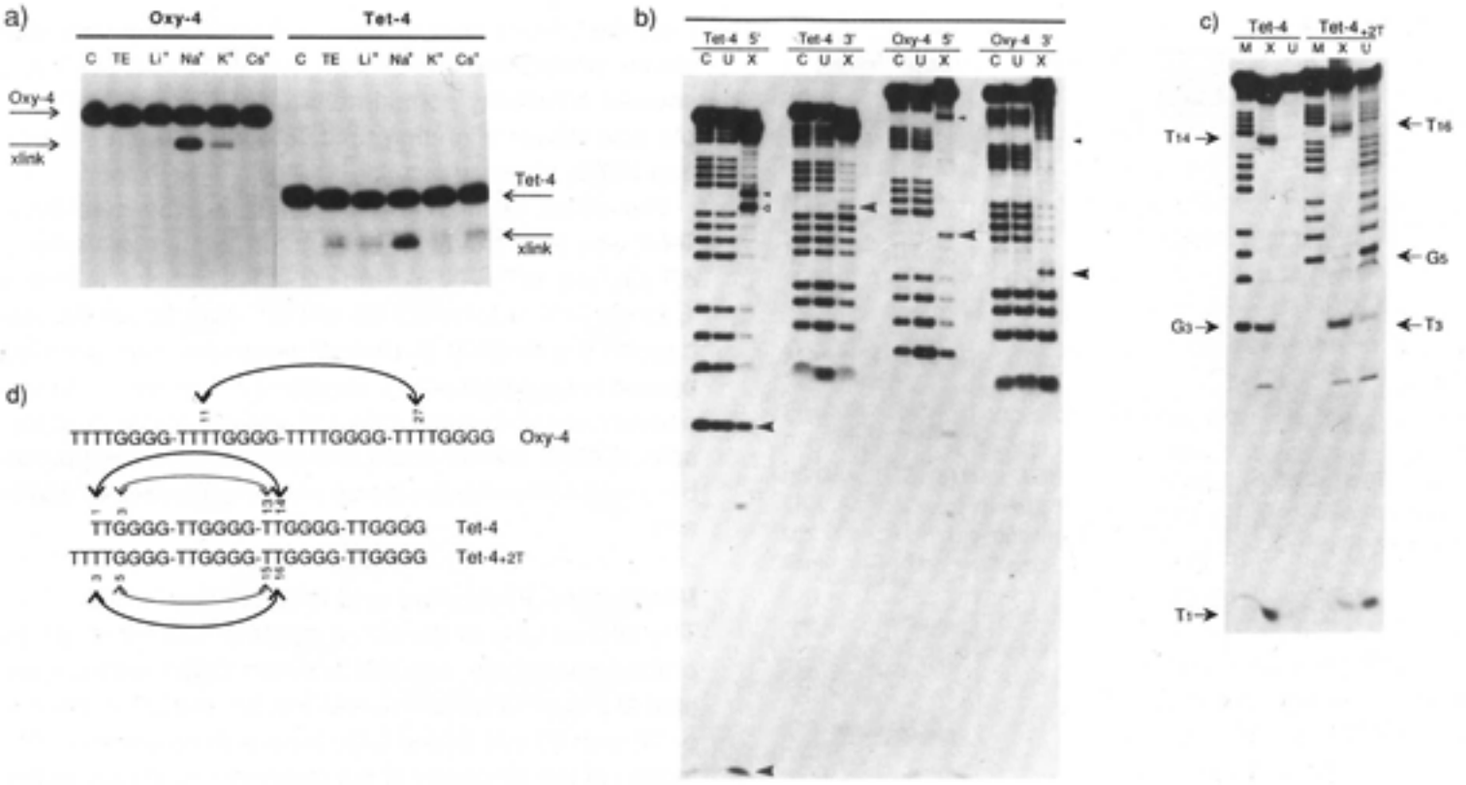


Figure 3. Cross-Linking of Oxy-4 and Tet-4

(a) Autoradiogram of a denaturing 10% polyacrylamide gel of the products of cross-linking upon irradiating Oxy-4 and Tet-4 for 4 min at 254 nm at 25°C. Lane C, unirradiated material. Other lanes contain irradiated material in TE or TE with 50 mM LiCl, NaCl, KCl, or CsCl added.

(b) Autoradiogram of a denaturing 20% polyacrylamide gel of the DMS cleavage products of cross-linked Oxy-4 and Tet-4, end labeled at either the 5′ or 3′ position. C lanes contain unirradiated material, U lanes contain material that has been irradiated but not cross-linked, and X lanes contain the cross-linked material. Large arrowheads mark the residue involved in the cross-link that is closest to the end label. Small arrowheads mark the partner in the cross-link that is farthest from the end label. The band marked with an open triangle is a DMS-dependent cleavage product (see text).

(c) Autoradiogram of a denaturing 20% polyacrylamide gel of the pyrrolidine cleavage products in cross-linked Tet-4 and Tet-4_{+2T}. The M lanes contain a DMS ladder of unirradiated material as a control. The X lanes and U lanes contain cleavage products of cross-linked material and irradiated but un-cross-linked material, respectively. Arrows denote the positions of the cross-links in the X lanes.

(d) Schematic illustration of the location of the cross-links in Oxy-4, Tet-4, and Tet-4_{+2T} as determined by chemical degradation. In both Tet-4 and Tet-4_{+2T}, there is a major cross-linked species indicated by the bold arrows and a minor species indicated by the thin arrows.

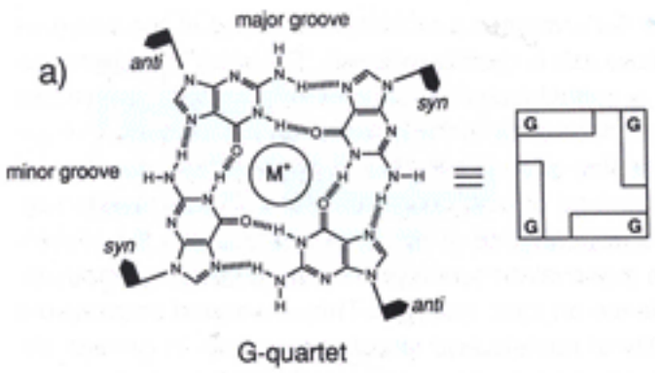
form of the DNA, the effect of different salts on the rate of cross-linking was studied. As shown in Figure 3a, NaCl greatly enhanced the rate of cross-linking of Oxy-4, while KCl had a smaller effect and LiCl and CsCl did not appreciably affect the rate of cross-linking. The initial rates of cross-linking in 50 mM Li^+, Na^+, K^+, and Cs^+ at 25°C were 0.01, 0.15, 0.05, and 0.02 min^{-1} (±20%), respectively (data not shown). The rate of cross-linking in the presence of different salts is not necessarily a reflection of the stability of the folded form. Although folding is required for cross-linking to occur, the rate of cross-linking depends on the details of the π-bond overlap in the structure, which may be different for different ionic forms.

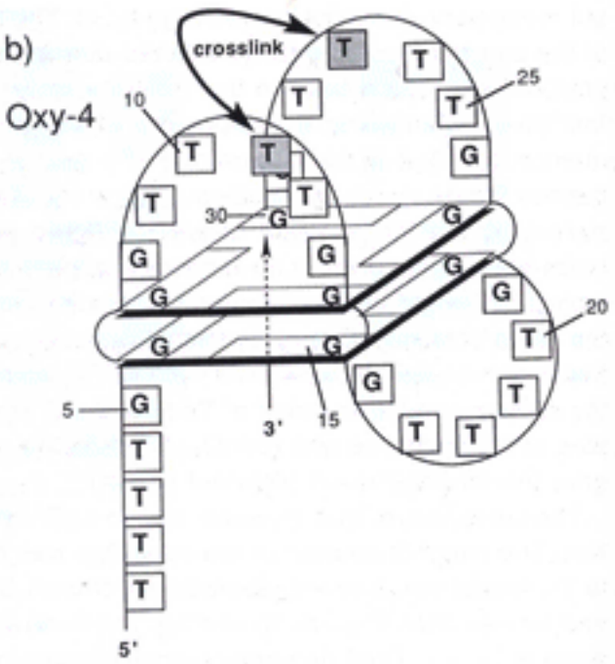
Other evidence supports the conclusion that the cross-link reflects the folded form of the DNA. When a native gel containing Na^+ was run at elevated temperature, ~38°C, folded and unfolded forms of Oxy-4 were observed as separate bands. Upon UV irradiation of this gel, only the folded form was cross-linked (data not shown). In addition, the cross-linked species comigrated with the folded form of un-cross-linked DNA in native gels containing Na^+ (data not shown).

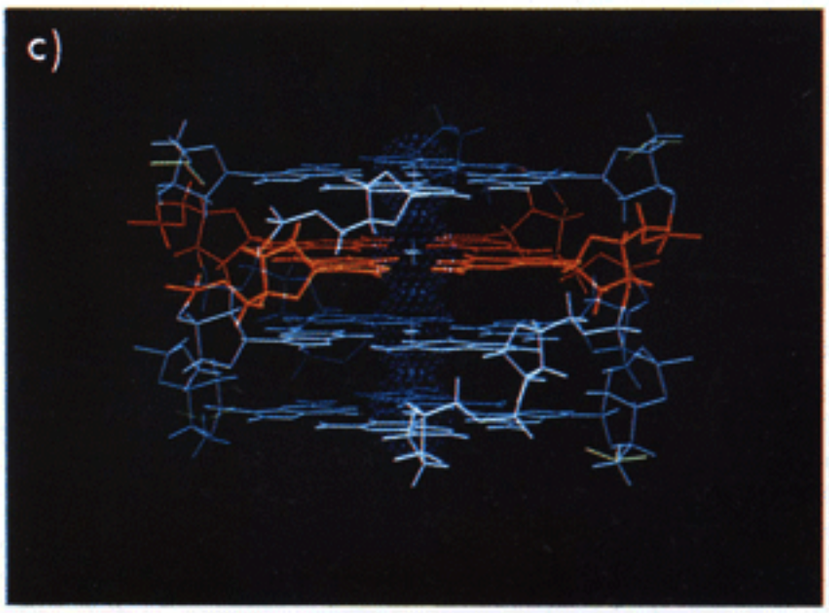
If the structure adopted by Oxy-4 were common to telomeric sequences in general, then cross-linking might occur in other telomere sequences. In fact, a cross-link was observed in Tet-4 as well, as shown in Figure 3a. The cross-linking efficiency in the presence of different salts followed a slightly different pattern from Oxy-4, but Na^+ was the most efficient ion for cross-linking in both cases.

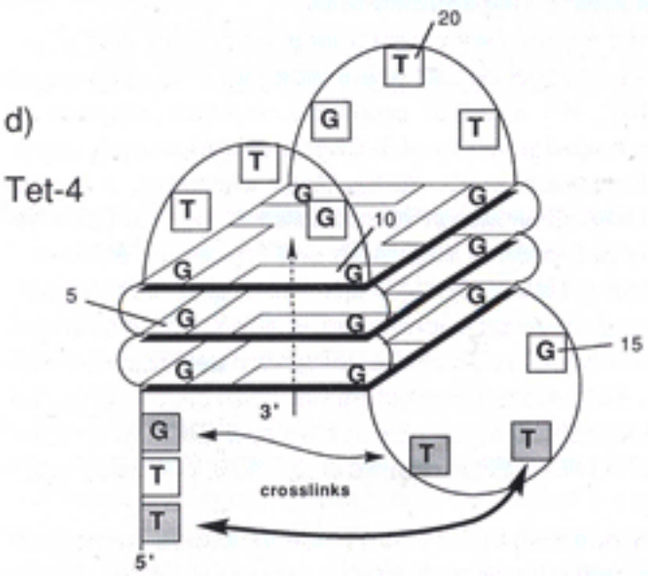
The positions of the cross-links were determined by chemical degradation of the cross-linked species. Unlabeled Oxy-4 and Tet-4 were irradiated at 254 nm in the presence of Na^+ in a preparative reaction. The cross-linked and un-cross-linked species were gel purified and then 5′ end labeled with T4 polynucleotide kinase or 3′ end labeled with terminal deoxynucleotidyl transferase. The labeled DNAs were treated with DMS or potassium permanganate ($KMnO_4$), followed by pyrrolidine-induced strand scission. The sequence ladders from the DMS reaction are shown in Figure 3b (X lanes). Control lanes containing unirradiated material (C lanes) and irradiated but uncross-linked material (U lanes) are also shown.

The presence of an extra band in the T-strings from a G-specific chemical cleavage reaction (Figure 3b, X lanes, filled arrowheads) suggested that pyrrolidine induced strand scission directly at the cross-linked nucleotides. This was confirmed by experiments in which cross-linked Oxy-4 or Tet-4 was subjected to pyrrolidine cleavage with-









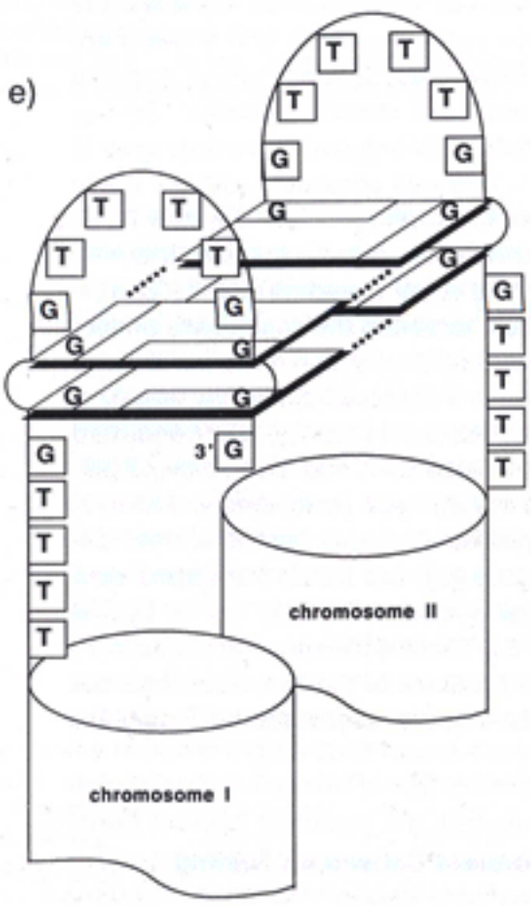


Figure 4. Models for the Structure of Telomeric Oligonucleotides

(a) The G-quartet. Four guanines are hydrogen bonded in a square-planar symmetric array. Each base is both the donor and acceptor in a Hoogsteen base pair. The cavity at the center is a possible binding site for a monovalent cation, M^+. The glycosidic torsions alternate between the *syn* and *anti* conformations, which creates major and minor grooves when several quartets are stacked.

led us to consider the possibility that G-quartets might be involved in the structure of telomeric sequences. We have constructed a model for the structure of telomeric four-repeat molecules that is consistent with most of the available data. The model for Oxy-4, shown in Figure 4b, consists of two stacked G-quartets capped by loops of thymidine and deoxyguanosine residues. Each G-quartet is comprised of one dG residue from each of the four G-strings. The two G-quartets are stacked directly over each other in Figure 4b for artistic convenience, but it is more likely that they are rotated with respect to each other in a right-handed helical manner (Zimmerman et al., 1975; Zimmerman, 1976). The structure in Figure 4b represents the minimum core of hydrogen bonding interactions in Oxy-4 consistent with the data presented. Additional interactions that might be present include additional G-quartets above and below, and intra- and interloop stacking or base pairing.

This model is consistent with the observation of dG residues in the *syn* conformation (Henderson et al., 1987). To date, the proposed four-stranded helices have been modeled as parallel strands, in which all of the guanosines can adopt the usual *anti* conformation (Zimmerman et al., 1975; Howard and Miles, 1982a, 1982b; Miles and Frazier, 1978; Sen and Gilbert, 1988). In our model, any two dG residues involved in a Hoogsteen base pair are equivalent to being on antiparallel strands, and so to adopt the proper geometry, one must be *syn* and one must be *anti*. In any G-quartet in our model, two residues must be *syn* and two must be *anti* (Figure 4a). This arrangement would disrupt the 4-fold symmetry suggested in the simplified drawing in Figure 4b.

A molecular model for the core structure containing four stacked G-quartets is shown in Figure 4c. This model is derived from the fiber diffraction structure of poly(G) (Zimmerman et al., 1975) but is modified such that alternating G-strings are in the *syn* and *anti* conformations. Two features that are not apparent from Figure 4b can be seen in Figure 4c. First, the G-quartets are rotated with respect to each other, resulting in a helical structure. Second, the alternating *syn* and *anti* G-strings give rise to two major and two minor grooves opposite each other (Figure 4a). This reduces the symmetry of the structure to 2-fold. A consequence of this fact is that the loops at the top of the structure are not equivalent to the loop at the bottom.

The observation that the two central dG residues in each G-string of Oxy-4 are protected from methylation by DMS suggests that these residues are involved in adjacent G-quartets. This very symmetrical methylation protection pattern is easily explained by the model in Figure 4b. In the presence of Na^+, all of the dG's are protected from methylation. There are two possible explanations for this observation. There might be a steric effect excluding DMS from methylating the N7's of the outer dG's, even though there is no hydrogen bond. Alternatively, there could easily be four G-quartets stacked in the Na^+ form (Figure 4c). Na^+ is small enough to fit into the center of the G-quartet, and the larger ions, K^+ and Cs^+, might cause slight alterations of the structure that expose the two outer G-quartets to methylation.

The pattern of methylation interference for folding Oxy-4 in the first and third G-strings is as would be predicted from the model of Figure 4b. The methylation interference pattern in the fourth G-string is fairly clear; all four G-s interfere, but the central two interfere more strongly. While methylation of G_{15} (the third dG in the second G-string) does not appear to interfere with structure formation in lane F of Figure 2b, examination of the cleavage pattern of the unfolded DNA (lane U) reveals that position G_{15} does in fact interfere. There are several possible explanations for the special behavior of G_{15}. First, it is possible that DNA methylated at G_{15} can adopt a distinct conformation that has the same mobility as the folded form. Second, the interference pattern may be a superposition of the interference from two distinct but similar conformers present in the Na^+ gel, and G_{15} might be structurally nonequivalent in the two forms.

On the basis of similar considerations, a model for the structure of Tet-4 is proposed in Figure 4d. It consists of three stacked G-quartets instead of two. This difference between these model structures reflects the differences observed in the methylation protection pattern between Oxy-4 and Tet-4 (unpublished data; Henderson et al., 1989). The mobilities of Tet-4 and the inosine variants in native gels support this model. The variant $d(TTIGGG)_4$ exhibited mobility comparable to Tet-4 in a Na^+ native gel, consistent with the dG to dI substitutions occurring in the loop regions of the structure shown in Figure 4d. In contrast, the variant $d(TTGIGG)_4$ did not show enhanced mobility in a Na^+ gel, which might be predicted since two of the three G-quartets are disrupted by the loss of two hydrogen bonds. However, both the K^+ and Li^+ forms of this variant are folded. If the affinity of K^+ for the variant is greater than that of Na^+, K^+ may be able to overcome

(b) Model for Oxy-4. Two G-quartets are stacked over each other and are connected at adjacent corners by loops of d(GTTTTG). The position of a UV-induced cross-link is indicated at the top of the model by the arrows. Variations of the model involve three or four stacked G-quartets with correspondingly shorter loops. Note that the structure does not have the strict 4-fold symmetry that is suggested by this schematic diagram, but rather has 2-fold symmetry about an axis in the plane of the page (see text).

(c) Detailed model for the core structure of telomeric DNA. Four G-quartets are stacked with a 3.4 Å helical rise and a 30° rotation (Zimmerman et al., 1975). One complete G-quartet is highlighted in red. Alternate G-strings have the *syn* and *anti* conformations, which results in a 2-fold symmetric structure with two major and two minor grooves. Ions that could bind in the plane of the G-quartet (i.e., Na^+) are indicated by dotted spheres.

(d) Model for Tet-4. Three G-quartets are stacked over each other and are connected at adjacent corners by loops of d(GTT). The positions of the two UV-induced cross-links are indicated at the bottom of the model by the arrows. Bold arrows indicate the major cross-link, and thin arrows indicate the minor cross-link (see text).

(e) Model for the association of natural Oxytricha telomeric DNA. The model in (b) can be divided vertically down the middle into two two-repeat molecules. Monovalent cations can then induce the dimerization of two telomeric ends to form structures containing G-quartets.

the deficiency in hydrogen bonds of this Tet-4 variant and promote the folding of the molecule. One possible explanation for the unexpected stability of the Li^+ form of $d(TTGIGG)_4$ is that the pocket at the center of the G-I-quartets is able to adjust in size to accommodate the smaller Li^+, since the geometric requirements of the quartet are relaxed by the two missing hydrogen bonds.

The location of the cross-link observed in Oxy-4 is consistent with the model shown in Figure 4b. The cross-link is between T residues in the second and fourth strings of T's (Figure 3d). These are located in the loops that connect the G-quartets at the top of the model structure. The model of Tet-4 in Figure 4d also has the second and fourth strings of T's on the same side of the molecule. The lack of an efficient cross-link could be due to the shorter loops in Tet-4 compared with Oxy-4; in any case, the failure to observe a cross-link is not a basis for eliminating the model since the ability to cross-link is extremely sensitive to the details of the base overlap.

The cross-link in Tet-4 has been unambiguously mapped to be between the first and the third repeats, although there appear to be two alternative cross-linked species. The DMS-dependent band observed in the cleavage of 5′ end-labeled, cross-linked Tet-4 is not observed in the lane containing 3′ end-labeled material. This suggests that a 5′ terminal lesion reduces the efficiency of end labeling by T4 polynucleotide kinase, as has been observed for thymidine dimers (Weinfeld et al., 1989). Pyrrolidine cleavage of cross-linked material that was 5′ end labeled before cross-linking revealed that the band at G_3 was reduced in intensity relative to the band at T_{14} in Tet-4 (data not shown). The band at G_5 in $Tet\text{-}4_{+2T}$ is also reduced in intensity compared with T_{16} (Figure 3c). These data are consistent with two cross-linked species in Tet-4: a major species joining T_1 and T_{14}, and a minor species joining G_3 and T_{13}. The main point to be derived from these data is that a cross-link can form between the first and third repeats in Tet-4 and that this is consistent with the model structure in Figure 4d.

Previously proposed models for the structure of Tet-4 are variations of a hairpin containing G-G as well as G-T base pairs (Henderson et al., 1987). The cross-link connecting the loops at the top of the structure in Oxy-4, between the second and fourth T-strings, does not provide evidence to discriminate between the current and previous models. Taken together with the cross-link observed in Tet-4 between the first and third T-strings, however, the evidence for the current proposal is compelling. In the hairpin model, the first and third T-strings are at the extreme ends of the molecule and could not give such a cross-link in the absence of additional tertiary interactions.

Structure of Two-Repeat Molecules

Although both Oxy-2 and Tet-2 exhibited an increased electrophoretic mobility in the presence of added salt, the folding did not appear to depend on the specific monovalent cation. Also, neither Oxy-2 nor Tet-2 exhibited any methylation protection or methylation interference (data not shown). The substitution of dI into Tet-4 interfered with structure formation, but substitution of all of the dG's in Tet-2 with dI's did not apparently disrupt the structure adopted by Tet-2. These facts indicate that the most stable structures adopted by two- and four-repeat molecules under these conditions in vitro may be quite different.

However, the G-quartet model does provide a framework to explain a previously observed phenomenon involving two-repeat structures: deproteinized Oxytricha macronuclear DNA has been found to form specific aggregates when incubated in the presence of Na^+ (Oka and Thomas, 1987). Chromosomal DNA from Oxytricha has also been observed by electron microscopy to form linear aggregates that involve the telomeres (Lipps et al., 1982). The junction of the chromosomes appears as a cusp, and this was interpreted as the two chromosomes being aligned in the same direction proceeding away from the junction. Although the natural chromosomal ends contain only two single-stranded telomeric repeats, dimerization, nucleated by sodium or potassium ions, could form a structure containing G-quartets. A model for this dimerization is shown in Figure 4e. In support of this model, Tet-2 was found to form an efficient intermolecular cross-link when incubated in the presence of potassium at 5°C (J. R. W., unpublished data). In addition, Tet-2 forms lower mobility aggregated forms in a concentration-dependent manner in K^+-containing native gels. The cohesion model shown in Figure 4e thus provides a structural basis for telomeric DNA association.

Biological Relevance

There is reason to believe that telomeric DNA has a specific folded structure in vivo. The ability of Tetrahymena and Oxytricha telomeric DNA sequences to support the maintenance of minichromosomes in yeast (Szostak and Blackburn, 1982; Pluta et al., 1984) suggests that telomeric DNA might have a common structural theme despite differences in nucleotide sequence. The models shown in Figure 4 represent a general class of structures that can be adopted by a variety of telomeric sequences. The common sequence elements, G-strings, are incorporated into common structural features, the G-quartets. The variable elements, the T/A-strings, form the loop regions where variations can be easily accommodated.

In the nonreplicating state, Oxytricha macronuclear DNA has two more repeats of $d(T_4G_4)$ than of $d(C_4A_4)$. It has been widely assumed that the $d(C_4A_4)$ is completely base paired and that only two $d(T_4G_4)$ repeats form the 3′ terminal extension, but there is no direct evidence for such a structure in vivo. Considering the stability of the structure formed by Oxy-4 under physiological conditions, a chromosome might terminate in the G-quartet structure, with two repeats of $d(C_4A_4)$ displaced.

Independent of the structure of the nonreplicating telomere, there are times in the cell cycle when additional single-stranded G-strings are expected to be transiently present, enabling the G-quartet structure to form. Telomerase adds large numbers of $d(T_4G_4)$ repeats in vitro, and might do so during telomere elongation in vivo. If the same enzyme is responsible for the initial addition of telomeres during macronuclear development, one might ex-

pect extended 3′ tails at this time as well. In fact, double-stranded telomere regions that are longer than usual have been observed in the developing macronuclei of Euplotes (Roth and Prescott, 1985); according to the current model for telomere replication (Grieder and Blackburn, 1985; Zahler and Prescott, 1988, 1989), single-stranded G-tails would exist as intermediates prior to C-strand synthesis. Formation of the G-quartet structure at these times could serve to transiently stabilize the ends of chromosomes prior to further processing or protein binding.

The methylation protection pattern observed for Oxy-4 in solution is very different from that observed for Oxy-4 bound to the 98 kd telomere protein (Raghuraman et al., 1989), indicating that the final structure of the DNA in the nucleoprotein complex is different from the G-quartet structure presented here. Perhaps the telomere binding protein initially recognizes the G-quartet structure and perturbs it upon binding, or perhaps the structure must be unfolded for the protein to bind. Although the protein can bind two-repeat molecules (Oxy-2), it binds much more effectively to oligonucleotides containing two additional repeats that are either single stranded (as in Oxy-4) or double stranded.

The G-quartet structure is readily formed under physiological conditions. A structured form of telomeric DNAs has been proposed to be an obligatory substrate for the Tetrahymena telomerase enzyme (Henderson et al., 1987; Greider and Blackburn, 1987), and the G-quartet is a candidate for this structure. In contrast, the G-quartet structure is probably not present in the nonreplicating Oxytricha telomeric nucleoprotein complex. In either case, it is likely that the proposed G-quartet is a structure that must be dealt with by the components involved in the replication of telomeres.

Experimental Procedures

Oligonucleotide Synthesis

Oligonucleotides were synthesized on an Applied Biosystems model 380A DNA synthesizer and were purified by electrophoresis on denaturing 10%–20% polyacrylamide gels. Both unlabeled and end-labeled oligonucleotides were recovered from the gel by soaking the gel slice in 0.5 M NH_4OAc and then loading the solution onto a Waters Sep-Pak C_{18} cartridge. The cartridge was washed with 3 ml of water, and the oligonucleotide was eluted with 2 ml of 50% acetonitrile.

Nondenaturing Gel Electrophoresis

Native gel electrophoresis was performed in 12% polyacrylamide gels (29:1 ratio of acrylamide to bisacrylamide) using 0.5× TBE (0.45 M Tris–borate [pH 8.3], 1 mM EDTA) buffer. The gels were run at 5°C in a Hoefer SE-620 electrophoresis apparatus, in which the gel is completely immersed in the bottom buffer tank, which is in turn cooled by a heat exchanger connected to a circulating cold bath. The gels were run at 8–10 V/cm at a maximum power of 10 W. At this power the temperature in the gel was measured to be 1°C–2°C higher than the bath temperature. Gels run with added salt contained salt in the gel as well as the running buffer. DNA samples were in 5 μl of TE plus salt at the same concentration as in the gel, and prior to electrophoresis were heated to 95°C for 2 min, cooled to room temperature, and mixed with 1 μl of 30% glycerol containing marker dyes. Samples were loaded onto the gel and equilibrated for a futher 2 min at the gel temperature. All native gels were run until the bromophenol blue marker migrated 16 cm (10–14 hr).

DNA Modification and Cleavage

G-specific cleavages were performed either by the standard procedure (Maxam and Gilbert, 1980) or by a modification of the standard procedure. In the modified procedure, the oligonucleotide was dissolved in 19 μl of TE buffer, and 1 μl of a fresh 1:100 dilution of DMS in water was added. After 10 min at room temperature, the reaction was stopped by the addition of 20 μl of freshly diluted 2 M pyrrolidine. Pyrrolidine has been found to be a useful substitute for piperidine owing to its increased volatility and reactivity (Shi and Tyler, 1989). Cleavage was effected by heating to 90°C for 15 min, after which the sample was dried under vacuum in a Speed-vac. The sample was resuspended in 25 μl of water and dried twice more. The sample was then dissolved in 10 μl of deionized formamide and loaded onto a denaturing 20% polyacrylamide gel.

T-specific cleavages were performed using a modification of the published procedure (Rubin and Schmid, 1980). The oligonucleotide was dissolved in 10 μl of TE buffer, and 10 μl of 1 mM $KMnO_4$, freshly diluted from a 100 mM stock stored cold in the dark, was added. The reaction was stopped after 5 min by the addition of 1 μl of allyl alcohol. Twenty microliters of 2 M pyrrolidine was added and the cleavage was performed as described above.

Methylation Protection

Methylation protection experiments were performed in a manner similar to that described above with the following exceptions. The samples were heated to 95°C in the desired buffer for 2 min before reaction and equilibrated for 5 min at 0°C, and then 1 μl of a fresh 1:16 dilution of DMS in water was added. After pyrrolidine cleavage as described above, the sample was dissolved in 50 μl of deionized formamide containing marker dyes and 0.2 μg/μl carrier RNA, and then a 5 μl aliquot was loaded onto a denaturing 20% polyacrylamide gel.

Methylation Interference

Methylation interference was performed by methylating the DNA with 0.05% DMS in TE at 65°C, as described above, but instead of stopping the reaction with NaOAc–β-mercaptoethanol, the samples were directly loaded on a native gel containing 50 mM NaCl for electrophoresis at 5°C. Bands corresponding to the folded and unfolded forms were cut from the gel. The methylated DNA was eluted from the gel slice, recovered by ethanol precipitation, and then subjected to piperidine cleavage, and the products were analyzed by electrophoresis in denaturing 20% polyacrylamide gels.

UV-Induced Cross-Linking

UV cross-linking was performed with a Spectroline model XX-15F lamp at 254 nm, which has a power of 1100 μW/cm^2 at a distance of 6 inches. Irradiations were performed in microtiter plates with conical wells at a distance of 6 cm from the lamp, in a volume of 5 μl. All samples were heated to 95°C for 2 min, then equilibrated at the desired temperature for 5 min before irradiation. Five microliters of formamide containing dyes was added, and the sample was loaded directly onto a denaturing 10% polyacrylamide gel. The extent of the reaction was quantitated by scanning the gel on an Ambis radioanalytic scanner. The fraction of the molecules cross-linked was taken as the ratio of the radioactivity in the cross-linked band to the sum of the radioactivities in the cross-linked and un-cross-linked bands.

Acknowledgments

We would like to thank Eric Henderson for helpful discussions and for sharing unpublished data. We thank Cheryl Grosshans for synthesis of the oligonucleotides. J. R. W. thanks the Jane Coffin Childs Memorial Fund for Medical Research for a fellowship. This research was supported by grant GM28039 from the National Institutes of Health. T. R. C. is an American Cancer Society Professor.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received September 1, 1989; revised October 18, 1989.

References

Blackburn, E. H., and Gall, J. G. (1978). A tandemly repeated se-

quence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena. J. Mol. Biol. *120*, 33–53.

Blackburn, E. H., and Szostak, J. W. (1984). The molecular structure of centromeres and telomeres. Annu. Rev. Biochem. *53*, 163–194.

Cech, T. R. (1988). G-strings at chromosome ends. Nature *332*, 777–778.

Frensdorff, H. K. (1971). Stability constants of cyclic polyether complexes with univalent cations. J. Am. Chem. Soc. *93*, 600–606.

Gellert, M., Lipsett, M. N., and Davies, D. R. (1962). Helix formation by guanylic acid. Proc. Natl. Acad. Sci. USA *48*, 2013–2018.

Gottschling, D. E., and Zakian, V. A. (1986). Telomere proteins: specific recognition and protection of the natural termini of Oxytricha macronuclear DNA. Cell *47*, 195–205.

Greider, C. W., and Blackburn, E. H. (1985). Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. Cell *43*, 405–413.

Greider, C. W., and Blackburn, E. H. (1987). The telomere terminal transferase of Tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. Cell *51*, 887–898.

Greider, C. W., and Blackburn, E. H. (1989). A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis. Nature *337*, 331–337.

Gruenwedel, D. W., Hsu, C., and Lu, D. S. (1971). The effects of aqueous neutral-salt solutions on the melting temperatures of deoxyribonucleic acids. Biopolymers *10*, 47–68.

Henderson, E. R., and Blackburn, E. H. (1989). An overhanging 3′ terminus is a conserved feature of telomeres. Mol. Cell. Biol. *9*, 345–348.

Henderson, E. R., Hardin, C. C., Wolk, S. K., Tinoco, I., Jr., and Blackburn, E. H. (1987). Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine-guanine base pairs. Cell *51*, 899–908.

Henderson, E. R., Moore, M., and Malcolm, B. A. (1989). Telomeric DNA structure analyzed by chemical protection, base analog substitution, and telomerase utilization in vitro. Biochemistry, in press.

Howard, F. B., and Miles, H. T. (1982a). A stereospecific complex of poly(I) with ammonium ion. Biopolymers *21*, 147–157.

Howard, F. B., and Miles, H. T. (1982b). Poly(inosinic acid) helices: essential chelation of alkalai metal ions in the axial channel. Biochemistry *21*, 6736–6745.

Klobutcher, L. A., Swanton, M. T., Donini, P., and Prescott, D. M. (1981). All gene-sized DNA molecules in four species of hyoptrichs have the same terminal sequence and an unusual 3′ terminus. Proc. Natl. Acad. Sci. USA *78*, 3015–3019.

Lipps, H. J., Gruissem, W., and Prescott, D. M. (1982). High order DNA structure in macronuclear chromatin of the hypotrichous ciliate Oxytricha nova. Proc. Natl. Acad. Sci. USA *79*, 2495–2499.

Maxam, A. M., and Gilbert, W. (1980). Sequencing end-labelled DNA with base-specific chemical cleavages. Meth. Enzymol. *65*, 499–560.

Miles, H. T., and Frazier, J. (1972). Formation of a new 5′-guanylic acid helix in neutral solution. Biochem. Biophys. Res. Commun. *49*, 199–211.

Miles, H. T., and Frazier, J. (1978). Poly(I) helix formation. Dependence on size-specific complexing to alkalai metal ions. J. Am. Chem. Soc. *100*, 8037–8038.

Morin, G. B. (1989). The human telomere transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. Cell *59*, 521–529.

Moyzis, R. K., Buckingham, J. M., Cram, L. S., Dani, M., Deaven, L. L., Jones, M. D., Meyne, J., Ratliff, R. L., and Wu, J. R. (1988). A highly conserved repetitive DNA sequence, $(TTAGGG)_n$, present at the telomeres of human chromosomes. Proc. Natl. Acad. Sci. USA *85*, 6622–6626.

Oka, Y., and Thomas, C. A., Jr. (1987). The cohering telomeres of Oxytricha. Nucl. Acids Res. *15*, 8877–8898.

Oka, Y., Shiota, S., Nakai, S., Nishida, Y., and Okubo, S. (1980). Inverted terminal repeat sequence in the macronuclear DNA of Stylonychia pustulata. Gene *10*, 301–306.

Pinnavaia, T. J., Miles, H. T., and Becker, E. D. (1975). Self-assembled 5′-guanosine monophosphate: nuclear magnetic resonance evidence for a regular, ordered structure and slow chemical exchange. J. Am. Chem. Soc. *97*, 7198–7200.

Pinnavaia, T. J., Marshall, C. L., Mettler, C. M., Fisk, C. L., Miles, H. T., and Becker, E. D. (1978). Alkali metal ion specificity in the solution ordering of a nucleotide, 5′-guanosine monophosphate. J. Am. Chem. Soc. *100*, 3625–3627.

Pluta, A. F., Dani, G. M., Spear, B. M., and Zakian, V. A. (1984). Elaboration of telomeres in yeast: recognition and modification of termini from Oxytricha macronuclear DNA. Proc. Natl. Acad. Sci. USA *81*, 1475–1479.

Pluta, A. F., and Zakian, V. A. (1989). Recombination occurs during telomere formation in yeast. Nature *337*, 429–433.

Raghuraman, M. K., and Cech, T. R. (1989). Assembly and self-association of Oxytricha telomeric nucleoprotein complexes. Cell *59*, 719–728.

Raghuraman, M. K., Dunn, C. J., Hicke, B. J., and Cech, T. R. (1989). Oxytricha telomeric nucleoprotein complexes reconstituted with synthetic DNA. Nucl. Acids Res. *17* 4235–4253.

Richards, E. J., and Ausubel, F. M. (1988). Isolation of a higher eukaryotic telomere from Arabidopsis thaliana. Cell *53*, 127–136.

Roth, M., and Prescott, D. M. (1985). DNA intermediates and telomere addition during genome reorganization in Euplotes crassus. Cell *41*, 411–417.

Rubin, C. M., and Schmid, C. W. (1980). Pyrimidine-specific chemical reactions useful for DNA sequencing. Nucl. Acids Res. *8*, 4613–4619.

Sen, D., and Gilbert, W. (1988). Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. Nature *334*, 364–366.

Shi, Y., and Tyler, B. M. (1989). Pyrrolidine, a non-controlled substance, can replace piperidine for the chemical sequencing of DNA. Nucl. Acids Res. *17*, 3317.

Shippen-Lentz, D., and Blackburn, E. H. (1989). Telomere terminal transferase activity from Euplotes adds large numbers of TTTTGGGG repeats onto telomeric primers. Mol. Cell. Biol. *9*, 2761–2764.

Strauss, U. P., Helfgott, C., and Pink, H. (1967). Interactions of polyelectrolytes with simple electrolytes. II. Donnan equilibria obtained with DNA in solutions of 1-1 electrolytes. J. Phys. Chem. *71*, 2550–2556.

Szostak, J. W., and Blackburn, E. H. (1982). Cloning yeast telomeres on linear plasmid vectors. Cell *29*, 245–255.

Walmsley, J. A., and Pinnavaia, T. J. (1982). Stability constant of the 1:1 complex of sodium with guanosine 5′-monophosphate. Biophys. J. *38*, 315–317.

Walmsley, J. A., Barr, R. G., Bouhoutsos-Brown, E., and Pinnavaia, T. J. (1984). Ordered forms of dianionic guanosine 5′-monophosphate with sodium ion as the structure director. Proton and phosphorus-31 NMR studies of hydrogen bonding and comparisons of stacked tetramer and stacked dimer models. J. Phys. Chem. *88*, 2599–2605.

Weinfeld, M., Liuzzi, M., and Paterson, M. C. (1989). Enzymic analysis of isomeric trithymidylates containing ultraviolet light–induced cyclobutane pyrimidine dimers. II. Phosphorylation by phage T4 polynucleotide kinase. J. Biol. Chem. *264*, 6364–6370.

Zahler, A. M., and Prescott, D. M. (1988). Telomere terminal transferase activity in the hypotrichous ciliate Oxytricha nova and a model for replication of the ends of linear DNA molecules. Nucl. Acids Res. *16*, 6953–6972.

Zahler, A. M., and Prescott, D. M. (1989). DNA primase and the replication of the telomeres in Oxytricha nova. Nucl. Acids Res. *17*, 6299–6317.

Zimmerman, S. B. (1976). X-ray study by fiber diffraction methods of a self-aggregate of guanosine-5′-phosphate with the same helical parameters as poly(rG). J. Mol. Biol. *106*, 663–672.

Zimmerman, S. B., Cohen, G. H., and Davies, D. R. (1975). X-ray fiber diffraction and model-building study of polyguanylic acid and polyinosinic acid. J. Mol. Biol. *92*, 181–192.