The coherence of synthetic telomeres

O.L.Acevedo, L.A.Dickinson, T.J.Macke[1] and C.A.Thomas Jr*
Helicon Foundation, 4622 Santa Fe St, San Diego, CA 92109 and [1]Scripps Clinic and Research Foundation, Department of Molecular Biology, 10666 N. Torrey Pines Road, La Jolla, CA 92037, USA

Received January 14, 1991; Revised and Accepted April 10, 1991

ABSTRACT

The chromosomal telomeres of Oxytricha were synthesized and their ability to cohere examined on non-denaturing acrylamide gels containing the stabilizing cation K^+. At least 5 different mobility species were observed, in addition to that of the monomeric telomere. By cohering synthetic telomeres containing different lengths of subtelomeric DNA, we showed that each of the different mobility species was a dimer of two telomeres. Since the different mobility species did not differ in numbers or sequences of nucleotides, they must correspond to different molecular shapes probably caused by different degrees of bending of the dimer. Paradoxically, telomeres with longer subtelomeric stems cohered more efficiently. In the presence of K^+, solutions had to be heated to over 90° before the telomeres separated. Various synthetic constructs, restriction endonuclease and dimethyl sulfate protection experiments showed that the only nucleotides involved in the cohered structures were the 16 base 'tails' of sequence 3'$G_4T_4G_4T_4$. Extension of this motif was actually inimical to coherence. Oligomers containing 2 G_4T_4 motifs protected their GN7 positions by forming dimers, those with 5 G_4T_4 could do so by internal folding, but the 3' terminal group of G_4 was left unprotected. This suggests that only four groups of G_4 are necessary for the cohered structure. Single-chain specific nuclease, S1, as well as osmium tetroxide, which oxidizes the thymine residues of single chains, reacted less efficiently with the cohered structures. Synthetic telomeres containing inosine replacing guanosine were not observed to cohere, indicating that the C2-NH_2 is strongly stabilizing. The cohered structures appear to be unusually compact and sturdy units in which four G_4 blocks form quadruplexes stabilized by K^+. A new model for the cohered structure is presented.

INTRODUCTION

It is generally agreed that the telomeres of the macronuclear DNA of Oxytricha have the following sequence and structure:

$$\begin{array}{l} XXXG_4T_4G_4T_4G_4T_4G_4T_4G_4\ 3' \\ XXXC_4A_4C_4A_4C_4\ 5' \end{array} \quad (1, 2).$$

Yet, as first observed by Lipps (3, 4), these terminals were able to stick together in spite of the absence of any obvious base-pairing opportunities. Oka and Thomas (5) reinvestigated this phenomenon and concluded that protein-free DNA could cohere provided that the 16-base 'tail' was intact and unpaired with complementary oligomers and the indented 5'- terminated chain was also intact. The necessary nucleotides are marked by '^^^' above. Further, they demonstrated that the thermal stability of the cohered forms was increased by 25° in the presence of K^+. The only prior example of a specific cation-stabilization of a nucleic acid structure was polyinosine (6), which earlier X-ray diffraction experiments indicated was a quadruplex structure (7). The possibility that the high thermal stability was the result of unusual nucleosides was eliminated by the direct analysis of macronuclear DNA. The results indicated that there was less than 0.3 non-standard nucleosides per telomere (8).

At the time the present work was begun, it was not known how many telomeres joined to form the cohered products; nor was it known which portions of the telomere and interior DNA were essential for the formation of the unusually stable cohered structure. During the intervening time a number of papers have appeared which bear directly or indirectly on this problem. These are briefly summarized here.

The studies of Henderson *et al.* (9) made it clear that oligomers having the sequence of the G-rich strand were capable of self-association to form intramolecular hairpins. Working with synthetic sequences corresponding to the IgG switch regions, Sen and Gilbert were able to show that four oligomers associated in a parallel manner in such a way as to protect their N7 positions from the action of dimethyl sulfate (DMS) (10). In a subsequent publication these authors demonstrated the tetramerization of oligomers containing two groups of G_4 sequences and studied the competing roles of Na^+ and K^+ ions (11). Sundquist and Klug synthesized oligomers corresponding to the telomeres of Tetrahymena and observed two dimeric species in neutral gels (12). In this process the N7 positions of every guanine in the single chain tails became resistant to methylation by DMS. They proposed that the dimerization occurred by the association of

* To whom correspondence should be addressed

hairpin loops to form an antiparallel quadruplex of hydrogen bonded guanines. The single chains containing thymine formed unassociated loops and connected the double helical portion of the molecule. Williamson *et al.* (13) also studied self-association of oligomers corresponding to the G-rich stands of Tetrahymena and Oxytricha and found that in moderate levels of salt the oligomers would fold up to migrate rapidly on neutral acrylamide gels and at the same time protect from DMS the N7 positions of certain of the guanines. By exposing these structures to UV light, they found that specific T residues were crosslinked, a fact that indicated that they were adjacent to each other in the structure. On the basis of these results they proposed that telomeric association is the result of the formation of an antiparallel quadruplex structure, similar to that proposed by Sundquist and Klug, but involving only two of the four possible guanine residues from each set of four. Panyutin *et al.* studied the folding of long stretches of $(dG)_n$ by DMS protection experiments (14). On the basis of electrophoretic mobility in neutral gels, Raghuraman and Cech measured the rate of folding and unfolding of $(T_4G_4)_4$ and found that the telomeric protein would not bind to the folded forms (15).

In the present study, we have focused only on the synthetic telomeres of Oxytricha and studied them mainly in the presence of the highly stabilizing cation K^+. We have been able to demonstrate that coherence involves *only* the 16-nucleotide tails of the telomeres and that they join to produce at least five different dimeric structures the most stable of which is remarkably resistant to chemical and enzymatic probes. To account for these properties we propose a structure based upon a parallel G-quadruplex plate with T-residues hydrogen bonding to the G-quadruplex. This structure, called 'the crown' is stereochemically permissible and should be considered as an additional possibility to those previously proposed.

MATERIALS AND METHODS

Oligomers

Oligomers were synthesized by Operon Technologies, Alameda, CA and often used without additional purification. Occasionally the oligomers were purified by electrophoresis on denaturing acrylamide gels. Two complementary oligomers were annealed by heating to 90° in 0.1×Tris-EDTA (1 mM Tris.HCl, pH 7.5; 0.1 mM EDTA) and allowed to cool slowly over a period of several hours. The lowered electrolyte concentration appeared to retard the subsequent coherence.

Standard cohering treatment

The annealed synthetic telomeres were adjusted to a concentration of 5 to 10 μM in Tris-EDTA and 100 mM NaCl or 100 mM KCl, incubated at 70° for 10 min and slowly cooled to room temperature for 12 to 24 hrs.

Acrylamide gel electrophoresis

Electrophoresis was performed in vertical 8 to 10% acrylamide gels cast in a buffer containing 45 mM Tris-borate buffer (pH 8.0) containing 1 mM EDTA and 25 mM KCl. The KCl was essential for the preservation of cohered synthetic telomeres during electrophoresis. These gels were generally 20−28 cm high and run at 200 V, 60 mA for 3−4 hrs at room temperature. The glass plates were clamped to a flat 0.25' thick aluminium plate during electrophoresis and no warming could be detected. The slabs were then stained with ethidium bromide and photographed using conventional procedures (16). Sequencing gels were run in the standard manner using 8 M urea and 45 mM Tris-borate buffers (17).

RESULTS

Coherence and stability of synthetic telomeres

Oligomers 96 and 97 were annealed, adjusted to 100 mM NaCl or 100 mM KCl and subjected to a standard cohering treatment. Aliquots were made and subjected to a 15 min heat treatment at increasing temperatures before loading on a 10% acrylamide gel containing KCl. Fig. 1 shows that coherence in the presence of NaCl or KCl results in 4 ethidium-staining bands. As will be shown later, the fastest species called '1' (43 bp) is the monomeric telomere. The three slower-mobility species named '2', '3', and '4' are cohered products migrating at positions corresponding to 67, 93 and 100 bp respectively. On other gels, species '3' is resolved into three subspecies '3a', '3b', and '3c' for a total of *five* recognizable cohered species (Fig. 5). Exposure to temperatures up to 75° in buffer containing 100 mM NaCl eliminates '3' and '4', but has little or no effect if 100 mM KCl is present. The apparent stability is not the result of rapid reassociation because exposure to higher temperatures results in complete separation (see Fig. 4). Species '2' is puzzling in that it appears not to be dispersed by the highest temperatures used here. Other experiments indicate that it is not routinely formed, and is more likely to be seen with 'shorter' telomeres, like 96/97 which has only 30 base pairs of double helix. This species, which may be highly interesting, must remain a problem for the future.

Table I. Synthetic oligomers and telomeres

Number	Sequence	Length
90	5′ $C_4A_4C_4A_4C_4G_2C_2G_4T_4G_4T_4G_4T_4G_4T_4G_4$	60-mer
95	5′ $C_4A_4C_4A_4C_4G_2C_2$ I_4T_4 I_4T_4 I_4T_4 I_4T_4 I_3G	60-mer
	tail stem	
96	5′ $C_4A_4C_4A_4C_4$TATGTGCAAA	30-mer
97	3′ $G_4T_4G_4T_4G_4T_4G_4T_4G_4$ATACACGTTT	46-mer
98	5′ CCCCAAAACCCCAAAACCCCTATGTGCAAATTGATTCT	38-mer
99	$G_4T_4G_4T_4$GGGGTTTTGGGGTTTTGGGGATACACGTTTAACTAAGA	54-mer
102	5′ CCCCGGAACCCCAAAGCTCCTGGCCGTACATTGATTCT	38-mer
103	$G_4T_4G_4T_4$GGGGCCTTGGGGTTTCGAGGACCGGCATGTAACTAAGA	54-mer
104	5′ TGGCCGCCAGCTGAAGCTGAGGAAATAGAGTACTGAAA	38-mer
105	$G_4T_4G_4T_4$ACCGGCGGTCGACTTCGACTCC TTTATCTCATGACTTT	54-mer
102	5′ C_4GGAAC$_4A_3$GCTCCTGGCCGTACATTGATTCT	38-mer
106	ACCGGCATGTAACTAAG$_4$CCTTG$_4T_3$CGAGGACCGGCATGTAACTAAGA	54-mer

Cohered species are dimers

How many telomeric units are involved in the formation of the various cohered species? To answer this question we prepared telomeres of two different duplex lengths — 96/97 with a stem length of 30 base pairs and 98/99 with a stem length of 38 base pairs. (See Table I for the definition of 'tail' and 'stem'). These constructs were cohered both singly and together, the idea being that dimers would produce a single hybrid species, while trimers and tetramers would produce 2 and 3 hybrid species respectively. Fig. 2 shows that each cohered species formed a single hybrid band and is therefore a dimer. Species '1' formed no hybrid band and therefore must contain the uncohered monomeric telomere. Species '4', '3' and '2' have different mobilities, but each consists of only two telomeres. These results were confirmed by 8% gels and under various other conditions. Since the telomeric dimers have identical DNA sequences, and as will be shown below, are mainly double-helical, the different mobility species must correspond to different molecular shapes. Similar experiments have been done by Sundquist and Klug (11) and by Sen and Gilbert (10) to demonstrate dimerization and tetramerization respectively.

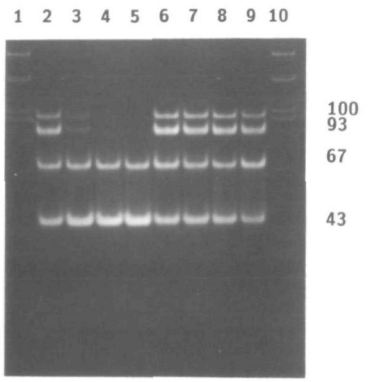


Figure 1. The thermal stability of cohered synthetic telomeres. Oligomers 96 and 97 were annealed and cohered in 100 mM NaCl or 100 mM KCl as described in Materials and Methods. The final telomeric concentration was 4 μM. After coherence, aliquots were heated for 15 min before loading on a 10% acrylamide gel. Lanes 2–5 show the results for coherence in NaCl, lanes 6–9 for KCl. Lanes 2 and 6: room temperature, 3 and 7: 55°, 4 and 8: 65°, 5 and 9: 75°. Lanes 1 and 10 contain pUC19/HpaII fragments which permit the assignment of nominal duplex segment length in bp to the four visible species as shown.

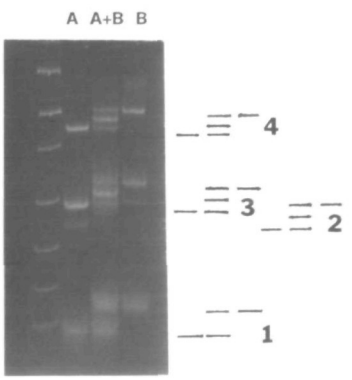


Figure 2. Cohered species are dimers. A pair of synthetic telomeres that differ in stem length were prepared. 'A' was 96/97 having 30 base pairs in the double helix and 'B' was 98/99 which was 8 base pairs longer. These samples were cohered at telomeric concentrations of 10 μM when alone, or a total of 20 μM when cohered together and run on 20% acrylamide gels. The first lane contains pUC9/HpaII fragments. The lanes marked 'A', 'A + B', and 'B' contain the cohered telomeres.

Cation type and stem length effects on coherence

The availability of synthetic telomeres of different double-helical lengths allowed us to compare their efficiency of coherence in the presence of different cations. In the case of native macronuclear DNA incubated at telomeric concentrations of 1 μM, it was found that Na^+ best promoted coherence (5). In the case of synthetic telomeres at a concentration of 10 μM, Fig. 3 shows that cation type and the stem-length of the synthetic telomere *both* have an effect. Na^+, K^+, NH_4^+ and Rb^+ promote coherence of 96/97 (30 bp); if 98/99 (38 bp) is tested, Li^+ and Cs^+ also promote coherence. Ca^{++} and Mg^{++} did not promote coherence of either telomere.

As judged by the proportion of ethidium staining in cohered forms, the fraction of stain in the dimeric forms is greater for the 38-bp telomeres than for the 30-bp telomeres. This 'stem-

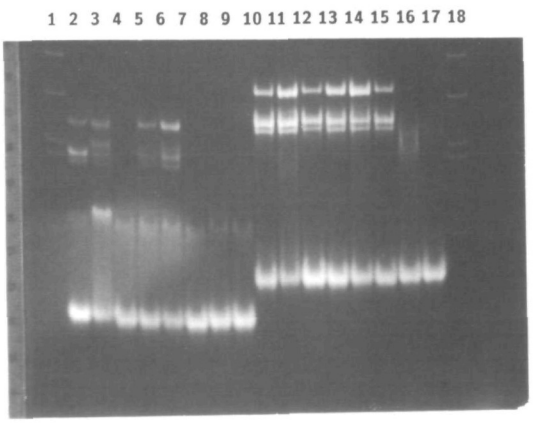


Figure 3. The cation type and stem length affects coherence. Telomere 96/97 (30 bp double helix) and Telomere 98/99 (38 bp double helix), both at concentrations of 10 μM, were cohered in 200 mM of various chloride salts and run on 10% acrylamide gels. Lanes 1 and 18 markers. Lanes 2–9, 96/97 cohered in Na^+, K^+, Li^+, NH_4^+, Rb^+, Cs^+, Ca^{++}, Mg^{++} respectively; lanes 10–17, 98/99 cohered under the same conditions.

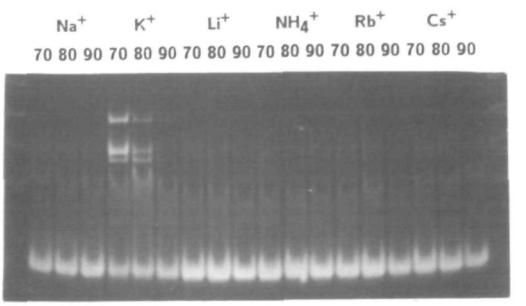


Figure 4. Cohered telomeres are stabilized by K^+. The telomere 98/99 was cohered at a concentration of 8 μM in 100 mM of the indicated cation chlorides (Na^+, K^+, Li^+, NH_4^+, Rb^+, Cs^+). Significant amounts of cohered species were demonstrated in each case (Fig. 3). Aliquots of each sample were then heated for 10 min at either 70°, 80°, or 90°, then rapidly cooled to 0° and loaded on 10% acrylamide gels.

length effect' has been observed under other circumstances to be discussed later.

Cohered telomeres are stabilized by K^+

The cohered form of native Oxytricha macronuclear DNA is uniquely stabilized by K^+, so it was of interest to determine whether the cohered synthetic telomeres behaved in the same way. All the monovalent ions that promote coherence were tested to see whether they increased the thermal stability of the cohered forms. The results shown in Fig. 4 demonstrate that only K^+ affords increased stability to the cohered forms. Some cohered species can be observed after 90° in the presence of K^+, but a 70° treatment demolished the cohered species in all other cases.

What portions of the telomere cohere?

The following experiments demonstrate that at least species '4' and '3' involve only the 16 base 'tails' and no other portion of the molecule. As shown in Fig. 5, synthetic telomeres with stems of arbitrary sequence, unrelated to the G_4T_4 motif, cohered, provided that they have a 'native' tail: $3'G_4T_4G_4T_4$. Telomeres with native stems and arbitrary tails did not cohere. Notice that monomeric telomeres with arbitrary tails run more slowly (80 bp) than monomeric telomeres with native tails (54 bp) which assume a more compact shape. It is curious that the species have different thermal stabilities in the order $4 > 3c > 3b > 3a$.

However, there remained the possibility that *any homologous* double-helical region appended to a native tail will promote coherence; for example, by fusing duplexes. This possibility is ruled out by the experiments shown in Fig. 6. A special telomere was synthesized (102/103) which contained four different restriction sites (HpaII, AluI, HaeIII, and RsaI) located throughout the stem. If the homologous double-helical portions were fused to form an unusual structure, we would expect the restriction sites to be uncleavable. Fig. 6 showed that each of these sites was as readily cleaved in the cohered dimeric forms as in the uncohered monomeric form. We conclude that the double helical portions of the dimer are undisturbed, which implies that stems of any sequence could form cohered structures if they were equipped with the 16 base $3'G_4T_4G_4T_4$ single chain.

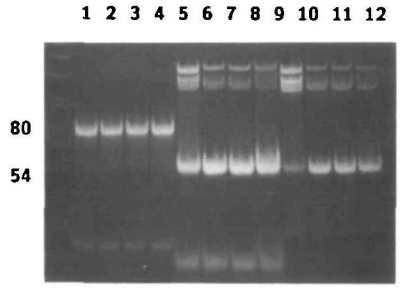


Figure 5. Telomeres with arbitrary stems cohere, but those with arbitrary tails do not. Telomeres 102/106 (arbitrary tail, native stem), 104/105 (native tail, arbitrary stem), and 102/103 (native tail, native stem) were individually cohered under standard conditions in 100 mM KCl, then heated at the indicated temperatures for 10 min, quick chilled on ice and immediately loaded on 10% acrylamide gels. Mobilities are given in nominal bp. Lanes 1–4, 102/106; lanes 5–8, 104/105; lanes 9–12, 102/103. Lanes 1, 5, and 9, room temperature; lanes 2, 6, 10, 70°; lanes 3, 7, 11, 80°; lanes 4, 8, 12, 90°. Band 3a, 3b, and 3c are resolved in order of increasing mobility.

The telomeric dimers have anomalous mobilities

If the dimers are joined by their single-chained tails to form a linear molecule like the double helix, one would expect their mobility in gels to be about equal to restriction segments of comparable length. For example, the dimers of telomere 102/103 should have mobilities corresponding to about 2×38 or 76 bp (not allowing any extra length to be contributed by the cohering structure itself). The cohered structure (species 4) moves much more slowly, corresponding to 165 bp — more than twice the known length of duplex. The various mobilities for the four dimers seen when 102/103 is cohered are shown in Table II.

When these dimers are treated with various restriction enzymes to break off known lengths from the two stems, the shortened cohered structures maintain their same relative mobilities and bands 4, 3a, 3b, etc may be recognized: they form a 'homologous series' (see Fig. 6). This means that once formed, the cohered structures are not altered by the shortening of the dimeric stems.

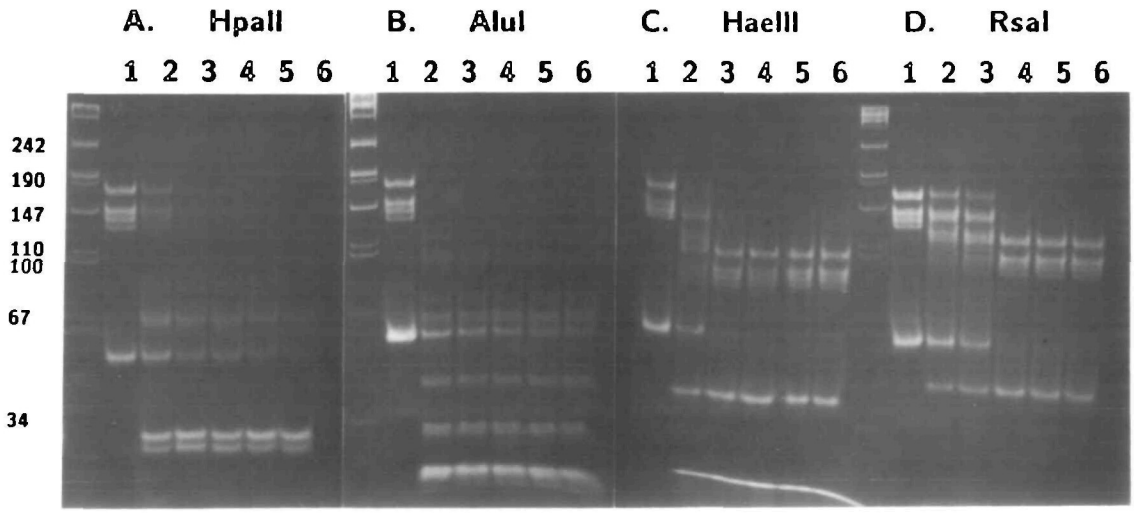


Figure 6. Restriction endonuclease analysis of cohered species. Aliquots of cohered telomeres 102/103 were adjusted to 40 mM KCl and incubated 1 hr at 37° with the indicated amounts of restriction enzyme. The reactions were stopped by adjusting to 20 mM EDTA and analyzed on 10% polyacrylamide gels. A HpaII; B AluI; C HaeIII; D RsaI. The samples in lanes 1–6 in A, C and D were treated with 0, 2.5, 5, 7.5, 10, and 15 units/μg of DNA respectively. The B-series was treated with 0, 1.25, 2.5, 3.75, 5, and 7.5 units/μg DNA respectively. Markers: pUC9/HpaII. Sizes are given in bp.

The stem-length effect

While the tails are the *only* portion of the telomere that are involved in the cohered structures, the double-helical portion may still play a role *in the formation* of the cohered structure. The availability of telomeres containing evenly distributed restriction sites afforded a possibility to prepare a series of telomeres of different stem lengths and to check their ability to cohere. Fig. 7 shows that the longer the double helix, the greater the *proportion* of cohered material. See also Fig. 3 above.

Cohered structures resist S1 nuclease

Fig. 8 shows a gel of the cohered telomeres (96/97 and 98/99) that were treated with increasing amounts of S1 nuclease. The S1 was demonstrably active on single chains, because the monomeric unit increased in mobility, presumably by losing its tail to the action of 0.5 U/μg. However, band '4' and bands '3' were resistant to 5 U/μg S1. Band '2' (not seen with 98/99) was partly sensitive, becoming altered between 1 and 3 U/μg; the products were stable to 5 U/μg.

The above experiment leaves open the possibility that S1 breaks the chain continuity, but the cohered structure remains intact. Therefore the integrity of the G-rich single chain was examined after treatment with S1. ^{32}P-labelled telomere (98/99*) was cohered, run on a 10% acrylamide gel and the DNA recovered from bands 4, 3 (a, b, c) and 1. After adjusting to S1 buffer conditions, each was treated with increasing levels of S1 as described under Fig. 9. As can be seen, when recovered from band 1 (monomer), oligomer 99 was shortened by treatment with as little as 0.01 units/ml. In contrast the DNA in band 4 required 10 to 50 times higher levels to produce the same products. The DNA in band 3 appeared to have an intermediate sensitivity. Thus, the 16 nucleotides in the cohered structures — particularly in species '4' — were quite resistant to S1 and in species '3' were partly so.

Table II. Observed and expected mobilities assuming linearity of cohered dimer (mobilities in bp)

Band Number	Known Length	Observed Length	Observed / Expected
4	76	165	2.17
3a	76	146	1.92
3b	76	140	1.84
3c	76	134	1.76
1	38+16	54	1.00

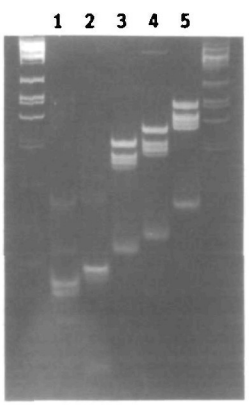


Figure 7. Telomeres with longer stems cohere better. Telomeres with decreasing stem lengths were obtained by cutting telomere 102/103 with RsaI, HaeIII, AluI and oligomers 104/105 with PvuII. The resulting telomeric fragments were gel purified, adjusted to the same molar concentration and cohered in 100 mM KCl. The lengths of the double-helical portion of the telomere in lanes 1–5 were 10, 16, 23, 27 and 38 bp (unbroken) respectively.

Guanine N7 is totally blocked in cohered tails

If the G residues forming the cohered structures were involved in a quadruplex, it would be likely that their N7 positions would

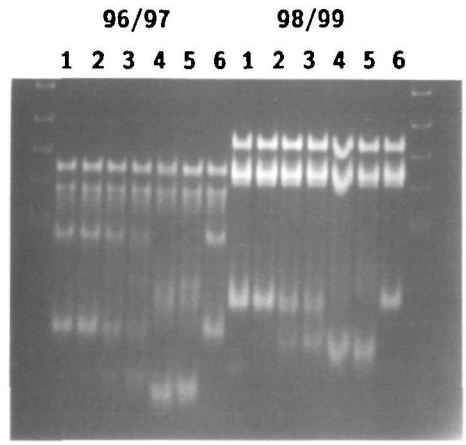


Figure 8. Cohered telomeres are resistant to S1 nuclease. Telomere 96/97 and 98/99 were each cohered in 200 mM KCl, adjusted to S1 buffer conditions, reacted with increasing amounts of S1 nuclease at 37° for 30 min then loaded on 10% acrylamide gels. Lanes 1 and 6 in both sets contained no S1 nuclease. Lanes 2–5 in both sets show the digestion products with 0.1, 0.5, 1.0, 3.0, 5.0 units/μg DNA. Flanking lanes: pUC9/HpaII markers.

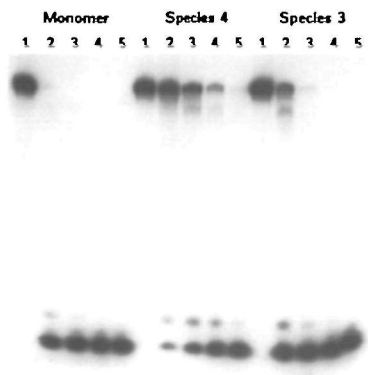


Figure 9. Cohered single chains are protected from the action of S1 nuclease. Oligomer 99 was labelled with ^{32}P at its 5' end and annealed to oligomer 98 (98/99*). Following coherence and separation on a 10% gel, the DNA in bands 4, 3 (including 3a, 3b, 3c) and 1 (monomer) were isolated and adjusted to S1 buffer conditions. Five aliquots (corresponding to lanes 1–5) were incubated for 2 hours at 37° with 0, 0.01, 0.05, 0.10 and 0.50 units/μl in a total volume of 10 μl. Each reaction was terminate by the addition of 2 μl of 0.5 M EDTA, 10 μl 80% formamide buffer and heating for 3 min at 90° then chilled on ice. The total sample was loaded on a prewarmed 20% acrylamide sequencing gel and run for 60 min at 100 V, 50 mA. The gel was exposed to X-Omat film with an intensifying screen for 30 min at 70°.

be blocked to the action of dimethyl sulfate (DMS), the reagent commonly used in DNA sequencing (17). The ^{32}P-labelled telomere (102/103*) was prepared, cohered, run on acrylamide gels and the DNA in bands '4', '3' and '1' recovered in buffer containing 25 mM KCl. These samples were treated with DMS, cleaved by piperidine, run on a 20% sequencing gel containing 8 M urea and autoradiographed. The results are shown in Fig. 10. Based on apparent band densities, (*) indicates the G is readily methylated and cleaved by piperidine; (.) indicates almost total nonreactivity: the (+) indicates partial reactivity.

GGGGTTTTGGGGTTTTGGGGCCTTGGGGTTTCGAGGACCGGCATGTAACTAAGA^{32}P
(Species 4)

GGGGTTTTGGGGTTTTGGGGCCTTGGGGTTTCGAGGACCGGCATGTAACTAAGA^{32}P
(Species 3)

GGGGTTTTGGGGTTTTGGGGCCTTGGGGTTTCGAGGACCGGCATGTAACTAAGA^{32}P
(Monomer)

As expected, the G residues in the double helix were reactive in all samples as expected. The two blocks of four G's in the tail (G43−46 and G51−54) were completely blocked in the two cohered species (4 and 3), but they were more available, but not completely so, in the monomer. Notice that the two cohered species were equivalent by this test, in contrast to the partial sensitivity of Species 3 DNA to S1 nuclease (Fig. 8) and OsO_4 (Fig. 11).

OsO_4 reacted less with thymines in cohered tails

OsO_4 reacts with the 5−6 double bond of thymine when it is in a single-chained (unpaired) condition; but is much less reactive with this base in the double helix (18, 19, 20). We applied this test for the pairing of thymine in the cohered tails. The same materials were prepared as for the DMS experiment above and reacted with OsO_4. As seen in Fig. 11, this reagent was relatively unreactive with T31 and T32 in the double-helical regions although it was highly reactive with the blocks of T residues in the single-chain tails of the monomer (T39−42 and T47−50). The cohered tails were intermediate in reactivity as diagrammed below, where (.) was least reactive; (:) somewhat more reactive; (+) more reactive; and (*) was most reactive of all.

GGGGTTTTGGGGTTTTGGGGCCTTGGGGTTTCGAGGACCGGCATGTAACTAAGA^{32}P
(Monomer)

GGGGTTTTGGGGTTTTGGGGCCTTGGGGTTTCGAGGACCGGCATGTAACTAAGA^{32}P
(Species 4)

GGGGTTTTGGGGTTTTGGGGCCTTGGGGTTTCGAGGACCGGCATGTAACTAAGA^{32}P
(Species 3)

The T residues in cohered structure Species 4 were significantly less reactive to OsO_4, suggesting that they are paired in some way. The T residues in cohered structure Species 3 appeared to

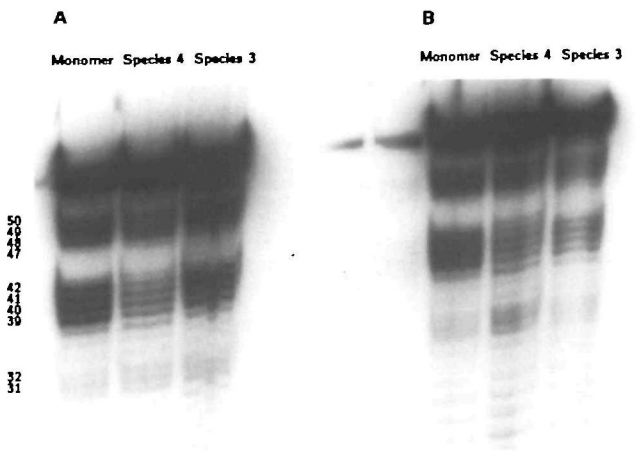


Figure 11. Thymine bases in the tails were less reactive with OsO_4 when cohered than when free in the monomer. ^{32}P-labelled telomere 102/103* was cohered and the species separated on a non-denaturing acrylamide gel as before. The DNA was recovered from bands '4', '3' and the 'monomer', and exactly 450,000 cpm of each was treated with 0.4% OsO_4 at 16° for 1 hour (A) or 2 hours (B). The reaction was stopped, treated with piperidine and run on a 20% acrylamide sequencing gel. The positions of the thymine bases are listed.

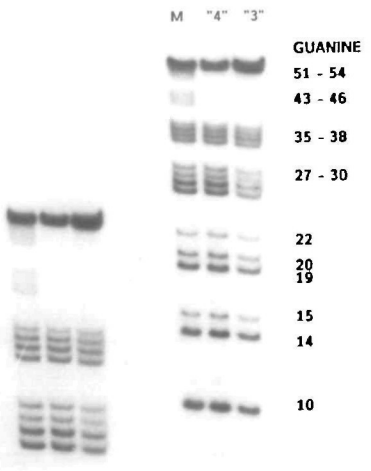


Figure 10. Coherence blocks methylation only in the two groups of four guanine bases in the tails. ^{32}P-labelled telomere 102/103* was cohered, run on an acrylamide gel, and the DNA recovered from the monomer band, 'M', and the two dimer bands, ''4' and '3'. An equal number of cpm of ^{32}P (approximately 10^6) from each band was individually treated with dimethyl sulfate, cleaved with piperidine and run on a 20% sequencing gel containing 8 M urea in 89 mM Tris-borate buffer. A duplicate pattern from an earlier loading appears to the left.

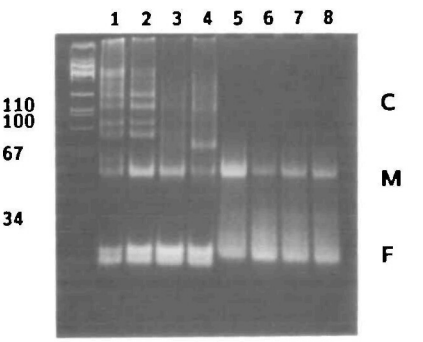


Figure 12. The C2 amino group of guanine stabilizes coherence. Oligomers 90 and 95, the latter having all but one of its Guanosines replaced by Hypoxanthines, were self-annealed at high concentration to make bistelomeres, then cohered in 100 mM KCl. Aliquots were heated for 10 min at increasing temperatures, quick chilled in ice and loaded on 10% polyacrylamide gels. Lanes 1−4, oligomer 90; lanes 5−8, oligomer 95. Samples in lanes 1 and 5 were held at room temperature; lanes 2 and 6, 70°; lanes 3 and 7, 80°; lanes 4 and 8, 90° respectively. C, M, and F denote the 'cohered structures', the 'monomer' and the 'fold-back' species.

be somewhat more reactive, but less so than those in the free tails. These results are comparable to the S1 nuclease sensitivity experiments, but contrast with the DMS protection experiments, where guanines in the tails in Species 3 and 4 were found to be equally resistant to methylation.

Guanine C2-NH_2 stabilizes the cohered forms

While it is possible to build K^+-stabilized quadruplexes with polyinosine (6), we were curious to know whether the cohered species could be formed with telomeres containing inosine instead of guanosine. In this case, we prepared palindromic telomeres, containing guanine (90) and hypoxanthine (95) (see Table I). When hybridized biomolecularly, bistelomeres are formed which can cohere to form multimeric structures seen as a 'ladder' on acrylamide gels. Fig. 12 shows that bistelomeres containing hypoxanthine did not cohere or were unstable to gel electrophoresis, indicating that the C2-NH_2 is essential for the formation or stabilization of the cohered structures. It is interesting that the guanine-containing oligomer 90 formed two mobility species of hairpin structures, while the hypoxanthine-containing oligomer 95 formed only one.

Mobility of monomeric telomeres and single chains

We have estimated the mobility of four different monomeric synthetic telomeres of various types each having a double helical portion of 38 bp and a 16 base single-chained tail. The mobility of their component G-rich single chains was also measured. The results are shown in Table III.

It is clear that those telomeres with native tails moved rapidly (at 54 bp) in comparison with a telomere with an arbitrary tail which ran at 80 bp. Since the double-helical portion of the

Table III. Mobility of synthetic telomeres and constituent G-rich oligomers

Telomere	98/99	102/103	104/105	102/106
	54	54	54	80
Oligomer	99	103	105	106
	34	34	42*	38

*A double band is seen at about 100 bp. (The mobilities given in bp duplex DNA should be considered estimates in view of the paucity of markers in this region of the gel).

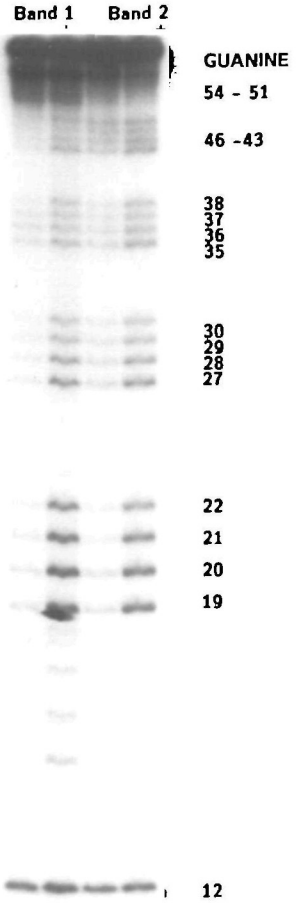


Figure 13. Reaction of DMS with 'cohered' 99. Oligomer 99 was labelled at the 5′ end with ^{32}P and subjected to a cohering treatment and run on non-denaturing gels. Autoradiographs revealed a band migrating at about 34 bp (band 1) and another at 100 bp (band 2). The DNA in these bands was isolated then subjected to methylation, cleavage and electrophoresis on a 16% sequencing gel. One aliquot (marked '+') was denatured (90°, 5′) prior to methylation; the other (marked '−') was not. The identity of each guanine is labelled at the right.

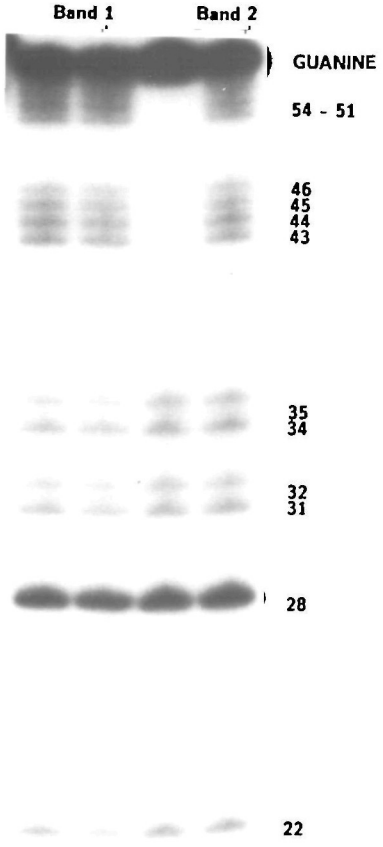


Figure 14. The reaction of DMS with cohered oligomer 105. Labelled oligomer 105 was cohered, run on acrylamide gels and the DNA purified from band 1 (42 bp) and band 2a (100 bp) and 2b (110 bp) taken together. Both samples were subjected to DMS treatment, one aliquot of each being denatured (+) just prior to methylation as described under Figure 13.

molecule cannot change its shape, the faster mobility must result from the more compacted $G_4T_4G_4T_4$ tail. Figs. 6 and 12 show that these 16-bases can assume at least *two* different compact forms differing in mobility.

The G-rich single chains 54 units long with four (or five) groups of G_4 moved quite rapidly — at 34 bp. The single chains having two groups of G_4, moved more slowly at 42 bp if the pair of G_4 groups was located at the 3′ terminal, and at 38 bp if the pair was located more centrally. These results indicate that the single chains themselves can fold up intramolecularly, possibly involving as many as *four* groups of G_4. This led us to examine these structures by DMS protection experiments.

Internal 'coherence' of single chains with G_4T_4 motifs

Since we now know that coherence only involves the 16 terminal nucleotides, one might expect that coherence could take place between single chains provided that they have the necessary $3'G_4T_4G_4T_4$ terminal. We subjected oligomers 99, 103, 105 and 106 to cohering conditions and analyzed the products on neutral K^+-stabilizing 10% acrylamide gels. The results clearly showed that oligomer 105 was the only one to produce what we interpret to be a dimer — a double band — moving at about 100 bp, that was heat stable (not shown). The mobilities of the oligomers themselves are listed in Table III. 105 is the only oligomer to have only a *pair* of G_4T_4 motifs located at the 3′ terminal followed by an arbitrary sequence. 99 and 103 have five and four appropriately spaced G_4 blocks. 106 cannot cohere because it has an arbitrary tail. We think that 99 and 103 cannot cohere because they rapidly form intramolecular structures analogous to the cohered structures themselves and are thus unavailable for bimolecular reaction. If this is the case, these oligomers should show characteristic protection of the guanine N7 positions.

Reaction of DMS with 'cohered' 99

While oligomer 99 produced no detectable level of slower moving material in ethidium stained acrylamide gels, if ^{32}P-labelled, some level of label was observed in the 'dimer' region (ca 100 bp), although most resided in the 'monomer' band. To be cautious, and to be sure that we were dealing with the *monomer*, DNA was isolated from both regions of the gel and subjected to DMS and piperidine treatments and run on 16% sequencing gels. The autoradiograms (Fig. 13) show that both samples had effected a substantial protection of their N7 positions, and that this protection was largely eliminated by exposure to 90° for 5 minutes just prior to DMS treatment. The pattern of protection shown in Fig. 13 and confirmed by other experiments is shown below using (*) for fully accessible, (+) for partly accessible, and (.) for fully protected.

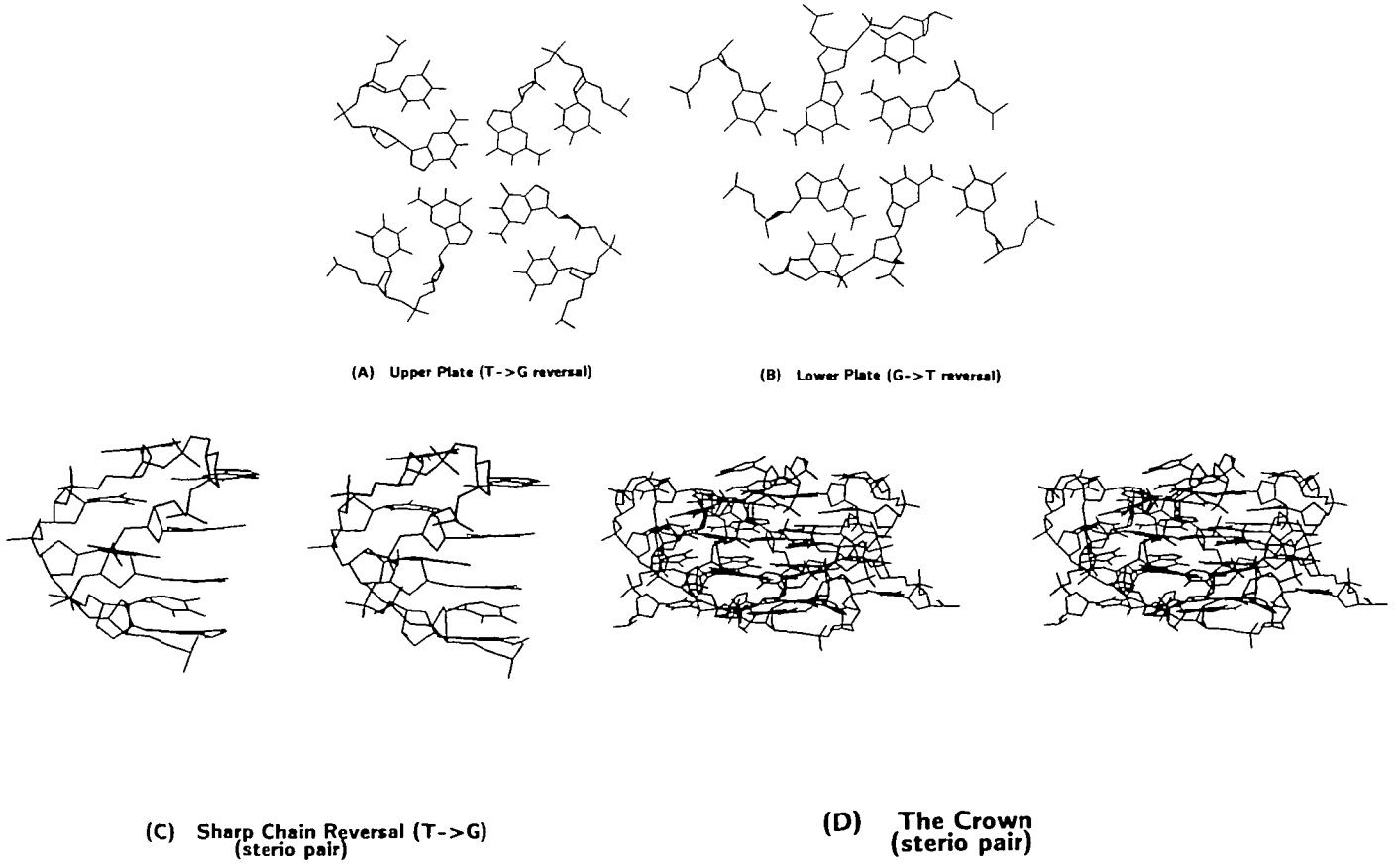


Figure 15. The crown model. It is proposed that the cohered structure is based upon four G4 planar tetramers stacked upon top of each other providing three cavities containing three K^+ ions. The thymines are hydrogen bonded to the remaining donor acceptor positions of guanine, namely N3-H···N3 and O4· H-N2 (reading T to G). Two of the resulting *eight-base* plates are shown in (A) and (B) corresponding to the 'upper and lower' plates. (C) shows a stereo pair of the sharpest of the two chain reversals, which occurs at the 'upper' plate. (D) shows a stereo pair of the completed crown. The double-helical stems are attached to the 'lower' plate to the two 5′ T's.

```
*+++ . . . . . *
GGGGTTTTGGGGTTTTGGGGTTTTGGGGTTTTGGGGATACACGTTTAACTAAGA-32P
(Monomer – Band 1)

*. . . +++ . . . *
GGGGTTTTGGGGTTTTGGGGTTTTGGGGTTTTGGGGATACACGTTTAACTAAGA-32P
(Dimer – Band 2)
```

In the monomer four blocks of four guanines were protected, leaving the fifth block (54–51) relatively exposed. Thus the folding and relatively rapid migration of oligomer 99 seen in Table III is associated with the protection of four groups of G_4. This is consistent with the idea that only four blocks are required to form a quadruplex; the fifth is left out. DNA from the dimer region had a different pattern of protection. G54 and G12 were completely available to the action of DMS in all lanes thereby serving as internal controls.

The reaction of DMS with 'cohered' 105

Oligomer 105, which has only two G_4T_4 motifs at the 3′ end, was subjected to a cohering treatment, run on K^+-stabilizing acrylamide gels and autoradiographed. The DNAs in band 1 (42 bp) and bands 2 (100, 110 bp) were isolated in buffer containing 25 mM K^+ and subjected to DMS treatment, one aliquot being denatured just prior to treatment. As may be seen, the guanines 54–51 and 46–43 in band 1 DNA (monomer) were partly accessible to cleavage (+), both prior to and after denaturation, as would be expected by the rapid reformation of intramolecular structures. However, in band 2 DNA (putative dimer) these guanines were completely blocked prior to denaturation (.) and somewhat accessible (+) after denaturation. The isolated guanine 28 seems to be completely accessible (*), as was the isolated guanine 12 in oligomer 99 (Fig. 13). In contrast with oligomer 99 described above which can protect four blocks of G_4 in the monomer, the terminal blocks of G_4 can be only completely protected when they unite with another oligomer, thereby recruiting two blocks of G_4 from each oligomer. The pattern of protection seen in Fig. 14 is diagrammed below.

```
++++  ++++  ++ ++ *
GGGGTTTTGGGGTTTTACCGGCGGTCGACTTCGACTCCTTTATCTCATGACTTT-32P
(Band 1 – monomer)

. . . . . +++ *  +
GGGGTTTTGGGGTTTTACCGGCGGTCGACTTCGACTCCTTTATCTCATGACTTT-32P
(Band 2 – dimer)
```

DISCUSSION

The evidence presented here demonstrates that the telomeres of Oxytricha cohere to form dimers of at least five different shapes having different mobilities on acrylamide gels containing the stabilizing cation, K^+. A number of different monovalent cations promoted coherence, but Ca^{++} and Mg^{++} did not. Once formed, only K^+ affords a sharply enhanced thermal stability.

Dimerization experiments using a variety of synthetic constructs and probing with restriction endonucleases, indicate that the cohered structure only involves the 3′$G_4T_4G_4T_4$ 'tail' from each telomere. The interior double helix does not participate in the cohered structure. The earlier observation (5) that the 5′$C_4A_4C_4$... chain must be intact probably results from the fact that its removal exposes additional G_4T_4 motifs that permit internal coherence as shown in Fig. 13 (see Introduction).

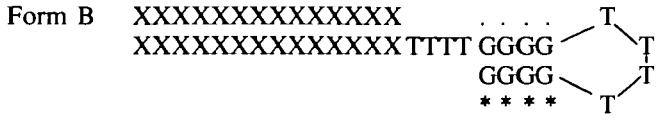
Dimers are probably bent: The cohered telomeres migrate more slowly than does linear DNA of the same molecular weight. The slowest dimer migrates at a position in the gel corresponding to a length of linear DNA that is more than twice its known length (Table II). Since bent DNA is known to migrate anomalously slowly (21), the simplest explanation for this is that the dimers are bent to different degrees — the slowest (species 4) having the sharpest bend. However, it is possible that the dimers are linear and that the 2-fold difference can be attributed to 5 different degrees of flexibility, because little is known about the relative mobility (faster or slower) of duplex segments having flexible joints at their center. Such flexible joints would have to be resistant to S1 nuclease.

The stem-length effect: The telomeres cohere very slowly. Oka and Thomas estimated that the rate of joining of the telomeres of native macronuclear DNA was 10^5-times slower than the rejoining of complementary lambda DNA terminals (5). Native macronuclear DNA will visibly cohere if incubated at room temperature at a telomeric concentration of 1 μM. However, the synthetic telomeres, which have 50-fold shorter stems, will not produce visible cohered product if incubated at this concentration: a 10-fold higher concentration is required, and then only a modest proportion of the monomeric telomeres are converted to cohered forms. Since only the tails are involved in the cohered structures, the lengths of the double-helical stems are unlikely to be affecting the free energy of coherence. Therefore it appears that the fraction of telomeres converted to dimers is limited by the rate of this bimolecular reaction. In view of this, it was of great interest to observe (Figs. 3 and 7) that synthetic telomeres with longer stems appeared to produce a greater proportion of cohered products.

At first this defies common sense, because the longer molecules are certainly not diffusing more rapidly. We propose, however, that telomeres with longer subtelomeric lengths of DNA are encountering one another (anywhere) *more rapidly*. Then ensues a 'scissor sliding' process (in two dimensions), followed by the 'snaring' of the double helix by the tail of the telomere, which now diffuses (in one dimension) until the two telomeres are connected. A similar kind of dimensional argument has been offered to explain the speedy association of the lac repressor with its specific binding sequence (22, 23). If correct, this proposal has a number of general implications which are testable.

Partial N7 protection: Before turning to the protection of the guanine N7 positions in the cohered structure, it is noteworthy that the G N7's are *partially protected* when free in the monomer. One example of this may be seen in Fig. 10 in the monomer lane. Why are these bands lighter than other G's, for example those in the double helical regions? A likely explanation is that G to G pairing can occur in *two* different ways G}G and G{G (where '}' represents a Hoogstein H-bond donor-acceptor pair). In one of these structures (Form A shown below), thie terminal block of (G54–51) has their N7 positions blocked (....), and the subterminal block (G46–43) has their N7 positions completely free (****). In the second structure, the pattern is reversed. The two different kinds of G:G pairing can be seen in Fig. 15. That G:G pairing is responsible for the folding and compaction of telomeric sequences has been proposed earlier by Henderson *et al.* (9, 24).

```
                                    . . . .  T
Form A  XXXXXXXXXXXXXX      GGGG  /     \  T
        XXXXXXXXXXXXXXTTTT  GGGG  \     /  T
                                  * * * *  T
```



This conjecture neatly explains the fact that there are *two* mobility species that can often be seen for the free monomer (Figs. 6, 7, 12).

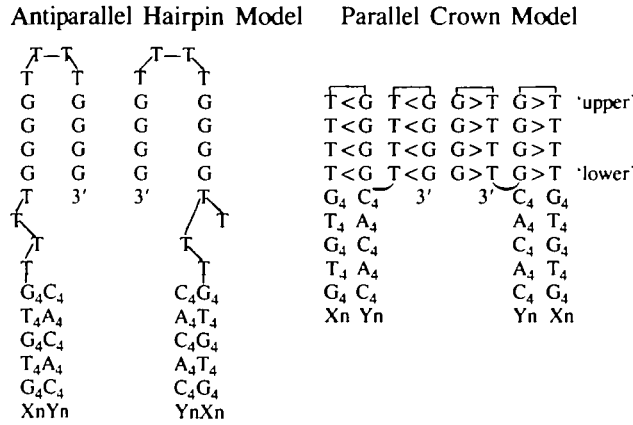
Probing the cohered structures: As seen in Fig. 10, coherence blocks the action of DMS in the two groups of G_4 in the tails of both dimeric species 4 and 3. As seen in Figs. 8 and 9, the cohered structures are quite resistant to the action of the single-chain specific nuclease S1, but species 4 is more resistant than species 3 when the digestion products are examined on denaturing gels. The osmium tetroxide experiments (Fig. 11) support this picture. Taken together, it appears that both cohered structures have their N7 positions blocked, and that the species 4 resists both S1 and OsO_4, while species 3 is somewhat more accessible to these reagents. This fits with the observation that the thermal stabilities assume the order 4 > 3c > 3b > 3a (Fig. 5).

Internally 'cohering' oligomers: The experiments with single chain oligomers that contain *five* G_4 blocks will 'cohere' internally to protect the N7 positions of *four* blocks of G_4 leaving *one* exposed (Fig. 13). In contrast if the oligomer contains only two blocks of G_4, then their N7 positions are not protected unless the DNA is recovered from the dimer region of the gel (Fig. 14). It should be emphasized that the structures formed intramolecularly by these oligomers may *not* be exactly the same structures produced by cohering telomeres. S1 and OsO_4 sensitivity experiments have not yet been performed.

It is noteworthy that the NIH Genebank contains numerous examples of multiple copies of G_mT_n in non-telomeric locations (see ref 25 for an example). One may speculate that these structures are involved in their possible function.

Models for the cohered structures: All this evidence suggests that the blocks of G_4 are associating four at a time to produce the planar quadruplex of guanine bases proposed by Gellert *et al.* (26) and analyzed by X-ray fiber diffraction by Zimmerman *et al.* (7). Stacked quadruplex plates of this kind could provide a cage for cations of specific size as proposed by Howard and Miles (6). These authors were thinking in terms of the parallel association of polynucleotide chains, although it is unlikely that the diffraction experiments of that time could have distinguished parallel or antiparallel association.

For cohering Tetrahymena telomeres, Sundquist and Klug (12) have proposed the antiparallel association of four blocks of G_4 by folding the chain into hairpins with the T_2 forming the loop and connecting the double helical segments. Williamson and Cech (12) have proposed a similar model for Oxytricha telomeres, except that when folded, only two of the four G N7 positions were protected. For simplicity these will be called 'hairpin' models. Shown below is a hairpin model for the cohering Oxytricha telomeres with the G4 quadruplex schematically unfolded and flattened. Notice that the T_4 single chains are substantially exposed and would be expected to be accessible to S1 and OsO_4. For this model to apply to cohering Oxytricha telomeres, some way must be found to render these T_4 single chains less available to S1 and OsO_4.



A second model called 'the crown' is shown above. In contrast with the hairpin models, the crown model achieves considerable protection of the segments of T_4 by sharply reversing their direction from the G_4 blocks (a sharp chain reversal) and hydrogen bonding them to the back of guanine using the G-N3 acceptor and the G-C2NH_2 donor. These are the *only* hydrogen bonding positions that remain available on quadruplexed guanine. Hydrogen bonding to this position, depicted as G>T and called the 'piggy-back position' is not found in the latest collections of all 29 possible base pairings (27). Thus, the crown model protects the T_4 chains, but requires two novel ideas — a sharp chain reversal and an unlisted base pairing.

To see whether the crown model was stereochemically permissible we undertook computer modelling experiments, the results of which are shown in Fig. 15. For this purpose we employed a collection of nucleotides having standard bond-angles and distances and assumed free rotation about all bonds (28). We first constructed the G_4 quadruplex and then paired four T's in the piggy-back position to obtain an eight-base plate similar to that shown in Fig. 15A or B. This plate was replicated three times and each plate displaced by 3.4 Å and rotated 31° about the central helical axis. The G's and T's were then connected together and finally the four connections at the top and two at the bottom of the crown were connected. Each 5' ended chain in the crown was extended by a 16 base-pair B-form double helix. It was possible to closely abut the blunt ends of the double helices to the lower surface of the crown. The resulting structure was energy-minimized by AMBER (29).

The resulting 'crown octaplex' displays T<G piggy-back pairing and sharp chain reversals having permissible bond angles and contacts (Fig. 15C). With 32 hydrogen bonds for 16 guanines, plus 24 additional ligands for the 3 K^+, this could be the most densely hydrogen-bonded nucleic acid (Fig. 15D). The two double- helical stems are attached to the lower plate of the crown to produce a sharply bent structure as predicted by the retarded gel mobility.

Molecular dynamics in the presence of water molecules and appropriate ions are now required. If structures of this kind do exist, it suggests that crowns of various heights can be constructed. Thus, G_mT_m, where *m* is any number greater than some minimum value (3?), would be expected to form stable structures. The structure leads one to inquire whether nucleotides other than T might be able to pair in the piggy-back position. If G_mX_m (where X is any nucleotide) can form crowns, these structures may apply to other telomeric species as well. These ideas are testable.

Just because a model can be built does not mean that the structure actually occurs. NMR and X-ray diffraction analysis is needed.

ACKNOWLEDGEMENTS

We thank Todd Miles for several helpful conversations and correspondence. This work was supported solely by the Helicon Foundation.

REFERENCES

1. Klobutcher,L.A. and Prescott,D.M. (1986) *The Molecular Biology of Ciliated Protozoa*, J.G. Gall Ed. Academic Press, N Y. pp 111–154.
2. Karrer,K.M. (1986) *ibid.* pp 85–110.
3. Lipps,H.J. (1990) *Proc. Natl. Acad. Sci. USA* **77**, 4104–4107.
4. Lipps,H.J., Gruissem,W. and Prescott,D.M (1982) *Proc. Natl. Acad. Sci. USA* **79**, 2495–2499.
5. Oka,Y. and Thomas,C.A.Jr (1987) *Nucl. Acids Res.* **15**, 8877–8898.
6. Howard,F.B. and Miles,H.T. (1982) *Biochemistry* **21**, 6736–6745.
7. Zimmerman,S B., Cohen,G.H. and Davies,D.R. (1975) *J. Mol. Biol.* **92**, 181–192.
8. Thomas,C.A.Jr., Park,J.W. and Ames,B.N. (1988) unpublished experiments.
9. Henderson,E., Hardin,C.C., Walk,S.K., Tinoco,I. and Blackburn,E.H. (1987) *Cell* **51**, 899–908.
10. Sen,D. and Gilbert,W. (1988) *Nature* **334**, 364–366.
11. Sen,D. and Gilbert,W. (1990) *Nature* **344**, 410–414.
12. Sundquist,W.I. and Klug,A. (1989) *Nature* **342**, 825–829.
13. Williamson,J.R., Raghuraman,M.K. and Cech,T R. (1989) *Cell* **59**, 871–880.
14. Panyutin,I.G., Kovalsky,O.I., Budowsky,E.I., Dickerson,R.E., Rikhirev,M.E. and Lipanov,A.A. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 867–870.
15. Raghuraman,M.K. and Cech,T.R. (1990) *Nucl. Acids Res.* **18**, 4543–4552.
16. Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory.
17. Maxam,A. and Gilbert,W. (1980) *Methods in Enzymology* **65**, 497–559.
18. Burton,K. and Reilly,W.T. (1966) *Biochem. J.* **98**, 70–77.
19. Friedman,T. and Brown,D.M. (1978) *Nucl. Acids Res.* **5**, 615–622.
20. Henderson,E.R. and Blackburn,E.H. (1989) *Molec. and Cell Biol.* **9**, 345–348.
21. Koo,H.-S., Wu,H.-M. and Crothers,D.M. (1986) *Nature* **320**, 501–506.
22. Riggs,A.D., Bourgeois,S. and Cohen,M. (1970) *J. Mol. Biol.* **53**, 401–417.
23. von Hippel,P.H. and Berg,O.G. (1989) *J. Biol. Chem.* **264**, 675–678.
24. Henderson,E.R., Moore,M. and Malcolm,B.A. (1990) *Biochemistry* **29**, 732–737.
25. Ullrich,A., Dull,J.T., Gray,A., Philips,J.A.III and Peters,S. (1982) *Nucl. Acids Res.* **10**, 2225–2240.
26. Gellert,M., Lipsett,M.N. and Davies,D.R. (1962) *Proc. Natl. Acad. Sci. USA* **48**, 2013–2018.
27. Hobza,H. and Sandorfy,C. (1987) *J. Am. Chem. Soc.* **109**, 1302–1307.
28. Saenger,W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag: New York.
29. Seibel,G., Singh,U.C., Weiner,P.K., Caldwell,J., Kollman,P.A., Case,D.A. and Bash,P. (1989) AMBER 3.0 Revision A, University of California, San Francisco.