# Crystal structure of four-stranded *Oxytricha* telomeric DNA

**ChulHee Kang*, Xiaohua Zhang*, Robert Ratliff†, Robert Moyzis† & Alexander Rich***

* Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
† Center for Human Genome Studies, Los Alamos National Laboratories, Los Alamos, New Mexico 87545, USA

The sequence d(GGGGTTTTGGGG) from the 3′ overhang of the *Oxytricha* telomere has been crystallized and its three-dimensional structure solved to 2.5 Å resolution. The oligonucleotide forms hairpins, two of which join to make a four-stranded helical structure with the loops containing four thymine residues at either end. The guanine residues are held together by cyclic hydrogen bonding and an ion is located in the centre. The four guanine residues in each segment have a glycosyl conformation that alternates between *anti* and *syn*. There are two four-stranded molecules in the asymmetric unit showing that the structure has some intrinsic flexibility.

TELOMERES are located at the ends of chromosomes and contain DNA with specialized sequences. They have an essential role in maintaining the stability and integrity of chromosomes (reviewed in ref.1). The sequences are made up of short nucleotide repeats and characteristically have a G-rich and a C-rich strand. Sequenced telomere ends have two repeats of the G-rich strand which overhang at the 3′ end[2,3] and they have unusual properties[4–7]. For example, in the macronucleus of the ciliated protozoan *Oxytricha*, telomere ends form cohering structures[4,5,8]. Several models have been proposed for the three-dimensional structure of the G-rich strand of telomers[9–11]. These models generally involve four guanine bases that are hydrogen-bonded to each other in a cyclic fashion (G-quartet). The origin of these models stems from experiments over 30 years ago in which the structure of fibres of polyinosinic acid was studied[12]. The fibres had a diffraction pattern consistent with the formation of three-stranded or four-stranded molecules in which the central bases form cyclic hydrogen bonds. Polyinosinic acid was later shown to be four-stranded[13,14]. A similar system of cyclic hydrogen-bonding of four guanine bases was used to interpret the stable gels formed by monomeric guanylic acid[15], as well as polyguanylic acid[13,14]. When guanine is used, eight hydrogen bonds are thought to stabilize the structure at each level of the helix. Structures of this type have also been proposed to be important in meiosis in regions of the chromosomes where there are stretches of G residues with four guanine strands parallel to each other[16].

Experiments dealing with the G-rich telomere sequences have led to the proposals that they can form four-stranded structures in which the two repeats of the 3′-terminal G strand are formed into hairpin-like structures which combine to form quartets of bases[9–11]. These models were based on an analysis of gel electrophoretic patterns, chemical modifications and ultraviolet-induced crosslinks. But until now there have been no single crystal X-ray analyses of any structure containing four guanine bases hydrogen-bonded in a cyclic fashion.

The 3′-terminal overhang of *Oxytricha* has the sequence $d(T_4G_4)_2$ (ref. 4). We have synthesized $d(G_4T_4G_4)$, crystallized it and solved its three-dimensional structure. This telomeric sequence forms a complex containing two $d(G_4T_4G_4)$ molecules held together by cyclic hydrogen bonding of four guanines. The two thymine tracts form loops found at opposite ends of the molecule, and a potassium ion is located at the centre of the quartets of guanine bases. Unlike the proposed models, however, the stretches containing guanine residues alternate in *syn* and *anti* conformation in agreement with nuclear magnetic resonance (NMR) experiments[17–19].

## Crystallization and structure determination

Many G-rich telomeric sequences were synthesized (Applied Biosystems Model 380B), but the most promising crystals were obtained with $d(G_4T_4G_4)$ from *Oxytricha*. The space group was $P2_12_12_1$, $a = 27.73$ Å, $b = 49.57$ Å, $c = 97.27$ Å. There is room for four $d(G_4T_4G_4)$ strands in the asymmetric unit. The structure was solved by molecular replacement and a Patterson correlation refinement[20] using four-guanine quartet models with alternate chains running in various directions both right- and left-handed. The right-handed coordinates were mostly derived from those of Zimmerman[21] with modifications in that alternate strands had either all *anti* or all *syn* conformation of the guanine residues. The right-handed conformation produced two significant solutions. These two positions were related by a pseudo 2-fold axis identical to the solution of the self-rotation search. After rigid body refinement of this partial model, difference maps were constructed which made it possible to trace the loops of the four thymine residues which were found on opposite ends of the G-quadruplex. Molecular replacement and simulated annealing refinement were done with the program X-PLOR[22] and graphics with FRODO[23]. Omit maps were made to determine chain polarity and glycosyl conformation. After restrained individual temperature factors were refined, a number of ions or solvent molecules were located. The final structure contained, in addition to the four strands of $d(G_4T_4G_4)$, 160 ions or solvent molecules, including spermines. Crystallographic data and refinement statistics are presented in Table 1 and full details will be published elsewhere. The coordinates have been deposited in the Brookhaven Protein Data Bank.

## Views of the telomeres

Each strand folds back on itself forming an intramolecular hairpin stabilized by G · G base pairs[6]. Two of these hairpins associate together in an *anti* parallel manner to form a stack of four guanine quartets, with thymine tract loops at each end of the molecule. The loops join the two G tracts on the same side of the quartet rather than across the diagonal. The structure has two wide and two narrow grooves. The latter are formed by the two 5′ ends of the molecule or the two 3′ ends (Fig. 1). The numbering scheme that was used is shown in Fig. 2. In numbering the residues, we have taken the first guanine residue at the 5′ position to be the one closest to the pseudo 2-fold axis. There are two molecules A and B and the numbering system is shown in Fig. 2. Molecules A and B are joined together by a symmetric pair of hydrogen bonds from the 3′ OH of G12 and G42 to phosphate oxygens of G42 and G12, respectively. The proximal

TABLE 1 Refinement summary

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Resolution (Å) | | 10.0–3.91 | 3.13 | 2.75 | 2.50 | 2.30 | Total |
| Possible | | 1,303 | 1,228 | 1,209 | 1,203 | 1,379 | 6,322 |
| Measured | (all†) | 1,177 | 1,163 | 963 | 905 | 755 | 4,963 |
| | (1$\sigma$‡) | 1,139 | 1,061 | 713 | 621 | 451 | 3,985 |
| | (2$\sigma$) | 997 | 780 | 332 | 244 | 118 | 2,471 |
| *R*-factor* (%) | (1$\sigma$) | 26.3 | 21.2 | 23.2 | 23.9 | 24.7 | 23.8 |
| | (2$\sigma$) | 23.0 | 16.5 | 17.3 | 17.6 | 19.3 | 19.5 |

Owing to the pseudo 2-fold axis parallel to the *y* axis and the resultant pseudo centering, a large proportion of the reflections are weak. We have not used non-crystallographic symmetry constraints between molecules A and B in the refinement. R.m.s. deviation of bond distance: 0.016 Å, bond angle: 3.99°.

* $R$-factor $= \sum |F_o - F_c| / \sum |F_o|$.

† All: above zero data.

‡ $\sigma$: based on I.

thymine loops of molecules A and B are close to each other and several ions are found interposed near the pseudo 2-fold axis. The helical axes of molecules A and B are oriented 125° from each other. A and B have very similar structures and the best molecular fitting between them gives an r.m.s. deviation of 1.48 Å. Comparison of the G-quartets alone has an r.m.s. deviation of 1.29 Å. Thus the largest differences between the two molecules was found in the organization of the thymine loops.

The distribution of electron density and the position of guanine residues in the first and second planes of molecule B is shown in Fig. 3*a*. The cyclically bonded guanine residues do not form a completely square array, but are slightly distorted. This is less so for the second and third planes, which may be associated with the presence of an ion located between them as discussed below. Because of the distortions from a square array in some planes, several hydrogen bonds are three-centred or bifurcated.

The core of the molecule consists of four stacked G quartets with a mean spacing of 3.5 Å. Figure 3*b* shows a projection of molecule A viewed in its electron density showing the four planes of the quadruplex as well as three of the four thymine residues in the loop. The disposition of sugar phosphate chains in electron density are shown in Fig. 4. Rounded segments of electron density found between the chains are solvent molecules or ions. Inspection of the guanine residues attached to the chain shows the manner in which they alternate in *syn* and *anti* conformations. Figure 2 lists the conformation of the glycosyl bonds, either *anti* (A) or *syn* (S). Glycosyl torsion angles are also listed for the guanine residues in both chains. The mean difference in the glycosyl angle of corresponding bases in molecules A and B is 23°.

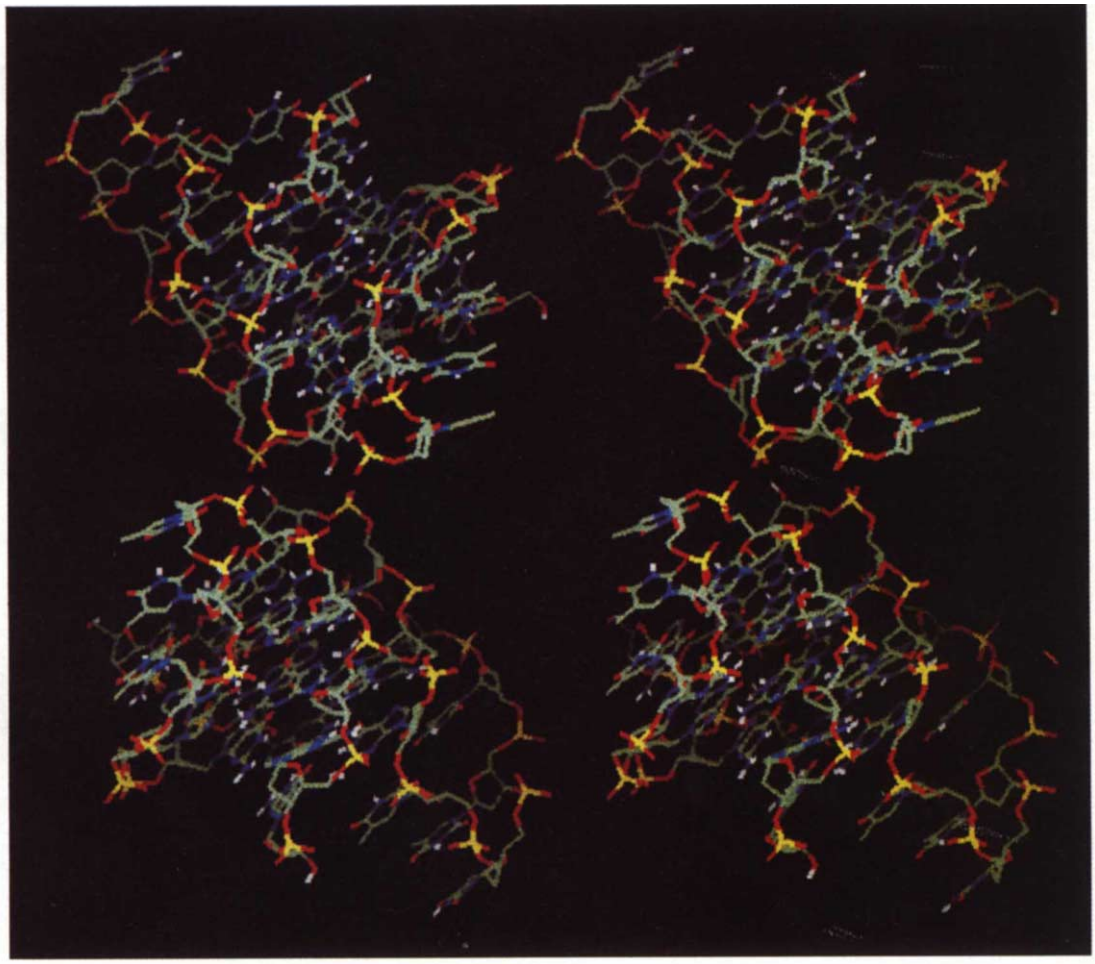


FIG. 1 Stereo view of the *Oxytricha* telomere structure looking down the pseudo 2-fold axis parallel to the *y*-axis. The upper molecule is designated A and the lower B. It can be seen that the molecule has grooves that are wide and grooves that are narrow between the sugar phosphate backbones. The narrow grooves shown here contain the two 3′ ends of the molecules. Crystals were grown by vapour diffusion using the hanging drop technique from a solution that contained 1 mM oligonucleotide, 10 mM $MgCl_2$, 6 mM spermine, 40 mM KCl, 20 mM potassium cadodylate buffer (pH 7.0), 5% 2-methyld-2,4-pentane-diol (MPD) equilibrated with a reservoir of 40% MPD. three-dimensional X-ray diffraction data were collected to 2.3 Å resolution using the multiwire Xuong-Hamlin area detector at the UCSD Research Resource.

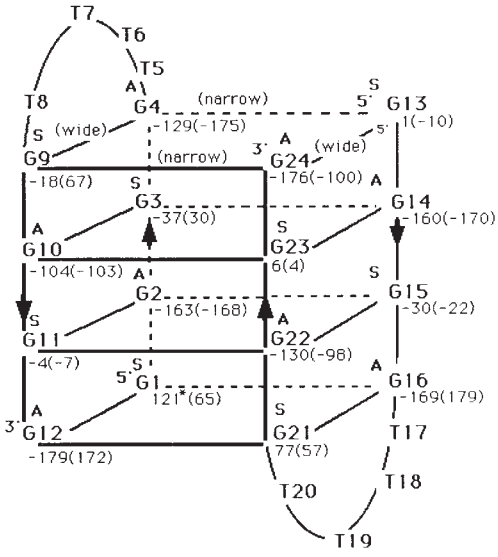
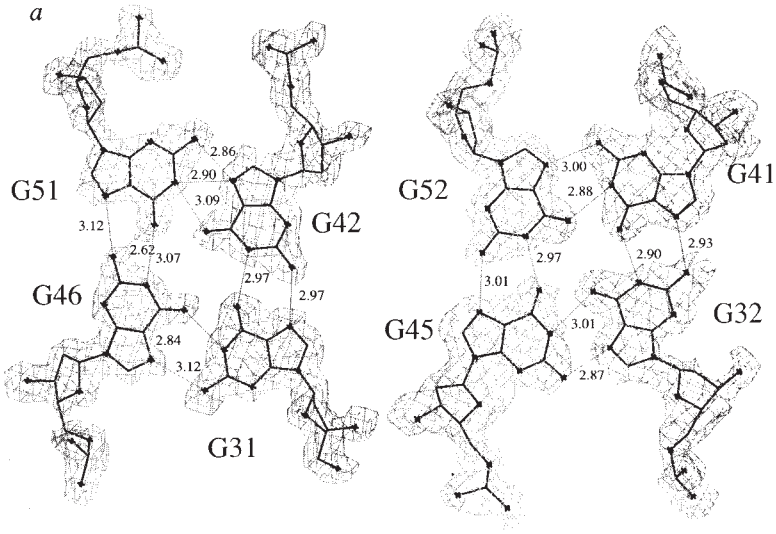
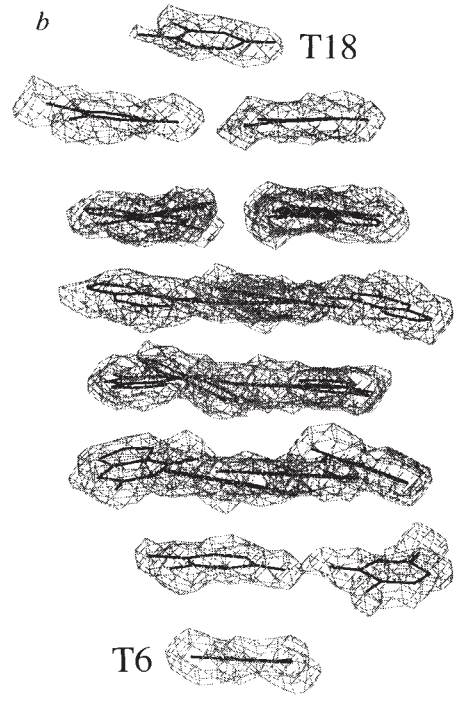



FIG. 2 The numbering system of the A molecule in the *Oxytricha* telomere structure. G1 is the 5′ end of one strand closest to the pseudo 2-fold axis. G1–G12 is one strand, G13–G24 is the other. Next to each guanine residue, S indicates *syn* and A indicates *anti* conformations around the glycosyl bond. Molecule B (not shown) is numbered G30–G42 in one strand, G43–G54 in the other. The two numbers next to the guanine symbols show the glycosyl torsion angles for the base. The number in parenthesis refers to the B molecule. The narrow grooves have either the 3′ ends or the 5′ ends of both strands. The asterisk next to G1 at the 5′ end indicates that its sugar is in a high *syn* conformation because of hydrogen-bonding.



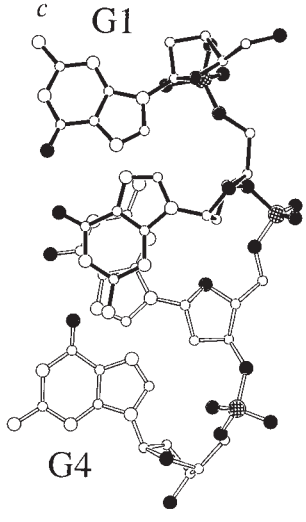




FIG. 3 The organization of guanine bases in the quadruplex. *a*, Four bases from the first two planes of the B molecule are shown in their electron density together with the system of cyclic hydrogen-bonding with bond lengths. Some hydrogen bonds are three centred as shown. The four guanines in the first plane (left) are arranged in an array that is not as square as the second plane (right). The electron density net is a 2 $F_o$–$F_c$ map plotted at 1.0 $\sigma$. *b*, A cross-section of the *Oxytricha* telomere stucture of molecule A at right angles to the G quartets. The bases are shown in their electron density nets calculated at 1.0 $\sigma$. Although the quartets are largely planar, there is some buckling of the bases. Stacking found in the central four G quartets is continued into some thymine residues of the loop sequence. Bases T6 and T18 are oriented parallel to the base stacking planes but are not close enough to give effective base stacking. *c*, Diagram of the four residues G1–G4 viewed perpendicular to the base planes. Guanine bases 2 and 3 are stacked on each other.

Unlike the proposed models, the guanine conformation along the chain alternate in *syn* and *anti*. Some of the explanation for this may be associated with a particularly stable interaction that is found between the bases in the two planes that have an *anti* followed by *syn* linkage. All of the four chains are similar in this regard and Fig. 3*c* shows a view normal to the four guanine residues in one strand. The second and third guanine residues overlap; the six-membered rings lie on top of each other and the 2-amino groups lie on top of the imidazole rings. There is considerable base stacking between planes 1 and 2 and planes 3 and 4, even though it is not as great as that found between 2 and 3. The average rotation angle between the guanine quartets is 28°.

## An axial ion in the quadruplex

In the electron density map there is a peak located in the centre of the four guanines between levels 2 and 3 in both molecules. An omit map was calculated in which the central peak was not included, and the results are shown for molecules A and B in Fig. 5. The peak is slightly off-centred, somewhat closer to level 2 rather than level 3. There is a strong likelihood that this is a potassium ion in an axial position. An axial ion was initially postulated from early studies with guanylic acid gels[24] and polyinosinic acid[25]. Note that the potassium ion is not well defined. When lower resolution data are included, a column of electron density is found filling the region between these two planes. Thus the potassium ion appears to be disordered in this central location.

## Flexibility in the molecule

The guanine residues in the quartet planes generally fit in one plane with a small amount of buckling. But a significant exception to that is found in residue G23 as well as its pseudo symmetry-related G53, which are buckled roughly 25° from the plane of the other three guanine residues. The buckled residues are still in hydrogen-bonding distances of the other two guanines in the quartet. Two factors seem to be involved in stabilizing the buckling. The N2 group of G53 (also 23) is hydrogen-bonded to the imidazole N7 of G40 (also 10) in the normal quartet. But its N2 is also in hydrogen-bonding distance of N3 of guanine 39 (also 9) immediately above it. Thus there is a three-centred hydrogen bond from N2 of G53 (and 23) to two guanine residues in both the third and fourth quartet level. In addition, an ion or solvent molecule is coordinated to the phosphate oxygen of P53 (and 23) and P39 (also 9), both of which have rotated somewhat towards the narrow groove. That ion is also coordinated to N2 of G53 (and 23) as well as to a fourth ion or solvent molecule. The N2 amino group of G53 (and 23) is thus involved in both three-centred hydrogen bonds as well as a coordination interaction.

There are two places in which the molecules are in close van der Waals contact with other molecules in the lattice. These close contacts could be related to some of the distortion seen in the molecule and it could also be responsible for the buckling of the bases. Buckling is frequently found in double-stranded nucleic acids, and in the case of the nogalamycin–DNA complex, a G · C base pair has a 26° buckling while still hydrogen-bonded together[26].

Figure 5*b* shows the side of molecule A with the 5′ end narrow groove in front. The terminal phosphates are rotated towards the narrow groove in a conformation different from that found in the other phosphate groups in the groove. Thus phosphate of G2 is hydrogen-bonded to the N2 of G15, whereas phosphates 14 and 44 are hydrogen-bonded to N2 of G3 and 33, respectively. This accounts for the marked narrowing of the groove on the 5′ side of the molecule. This type of phosphate hydrogen-bonding to amino groups is also found in the structure of polyadenylic acid[27].

## Loop conformation differences

The conformation of the four loops is best illustrated in Fig. 5, which shows two different sides of the two molecules. The loop conformations differ somewhat from each other, but there is a general organization. The first and fourth thymine bases in the loops are loosely stacked on the guanine quartets (3.7 Å), but are not in hydrogen-bonding distance. Instead, hydrogen-bonding bridging water molecules are found between the first and fourth base in three loops. The second thymine residue is aligned largely parallel and overlapping the first thymine residue, but at a spacing (4.1 Å) which is somewhat greater than found in normal stacking (Fig. 3*b*). The second thymine is intercalated between the third and fourth thymine and one (T6) makes a hydrogen bond with the phosphate group between the third and fourth thymine. At the apex of the loop, the third thymine residue is found in different orientations. The temperature factors (B) are significantly larger for the thymine residues than for the guanine residues, indicating significant loop mobility. Of the 16 thymine residues in the four loops, one residue, T38, appears to be in the *syn* conformation, whereas all of the remaining 15 residues are in *anti*. One ion is observed in all of the loops, positioned between the phosphates of the second and third thymine residues. These phosphates are close to each other because the chain makes a sharp turn at that point and the ion stabilizes this close interaction.

There are small but significant differences when comparing the conformation of molecule A and molecule B. The greatest difference is found near the 5′ end of one strand where phosphate 3 in molecule A is rotated into the narrow groove hydrogen bonding to the 5′ hydroxyl group of residue 13. But phosphate 33 in molecule B is rotated in the opposite direction into the wide groove where it hydrogen bonds to N2 of G32 in the second guanine plane. Phosphate 2 in molecule A is rotated into the narrow groove where it hydrogen bonds with N2 of G15 as mentioned above. But phosphate 32 in molecule B points out and is not involved in back hydrogen-bonding to the guanine residues. These conformational differences between molecules A and B reinforce the impression we have about the flexibility of the telomeric structure.

## Discussion

This structure demonstrates that two *Oxytricha* G-rich strands can associate to form a G-quadruplex as a structural motif. This organization is generally consistent with models that have been proposed for telomere four-strand structures[9–11]. Unlike the models, however, the sugar phosphate backbone has an alternating *syn–anti* conformation. Three guanine-containing molecules with G tracts capable of forming quartet structures, including one with the sequence studied here, have been studied by NMR[17–19]. These have all identified alternation of *anti* and *syn* conformation and our findings in this structure are consistent with those studies.

This is the second DNA structure found using alternating *anti* and *syn* conformations. In left-handed Z-DNA, the residues alternate in *anti* and *syn* conformations, with purine residues generally in *syn* conformations[28]. In the present structure, the purine (G) residues adopt both *syn* and *anti* conformations. But the reasons for adopting alternating *syn* and *anti* conformations are not the same in the two structures. A four-stranded guanine model has been proposed in meiosis[16]. In that model, all four strands are parallel and the conformation of its backbone remains to be determined.

The stabilization between levels 2 and 3 in both molecules has several factors contributing to it. (1) The presence of the potassium ion between these levels; (2) the stacking overlap between guanine residues on the same chain (Fig. 3*c*); (3) the hydrogen-bonding between the guanine N2 groups in levels 2 and 3 (for example G3 and G15) to phosphate groups near the 5′ end of the chain (P14 and P2) which are rotated into the groove to form hydrogen bonds; (4) an intricate system of

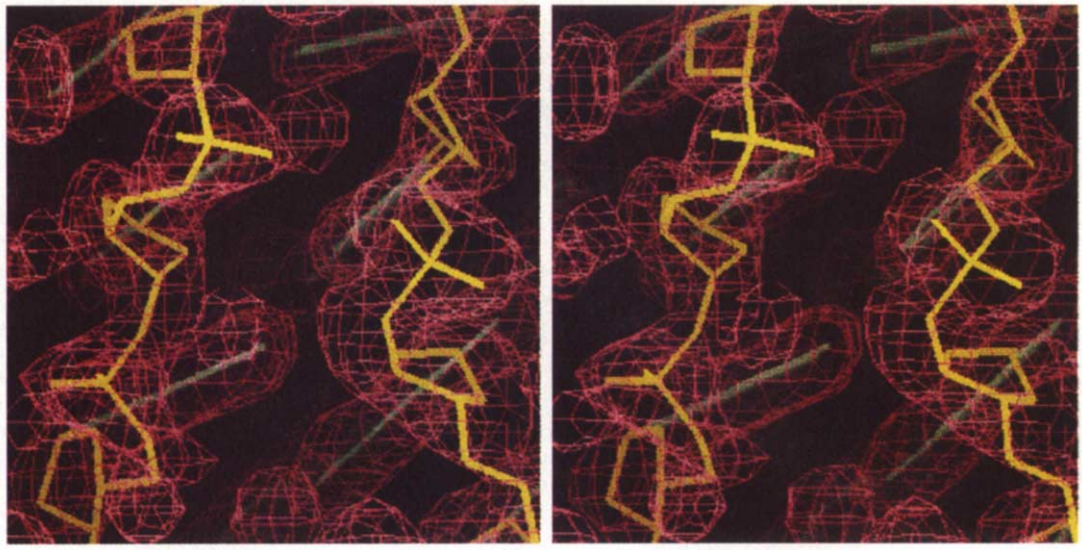


FIG. 4. Stereo view of the sugar phosphate backbone in an electron density net calculated at 0.9$\sigma$. The view shows part of the narrow groove of the A molecule with phosphates 10 (upper) and 11 on the left, phosphate 22 on the right. Rounded segments of electron density not connected to the molecule represent ions or solvent molecules.

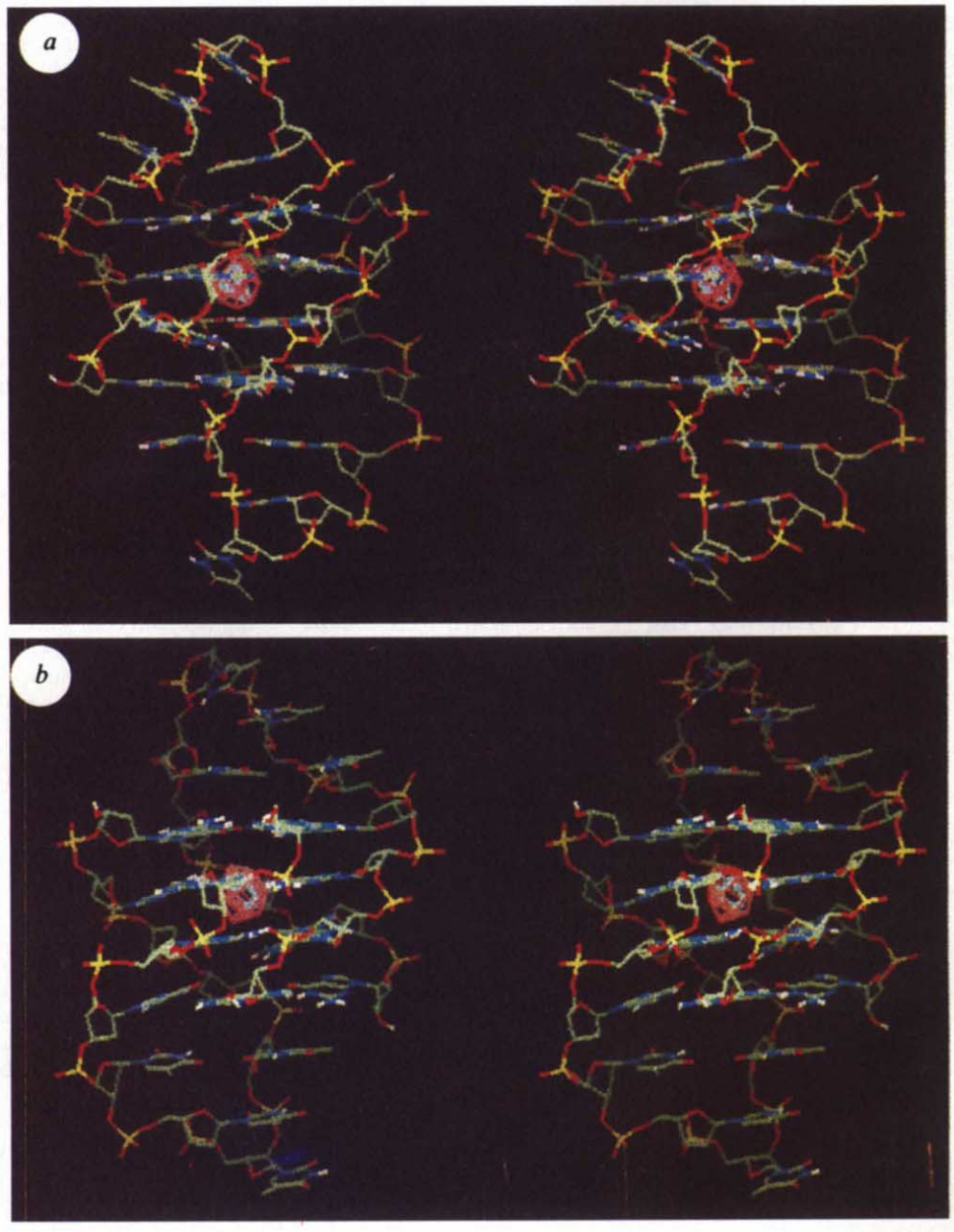



FIG. 5 The structure of the two *Oxytricha* telomere molecules in the asymmetric unit is shown together with an ($F_O$–$F_c$) omit map in the central region between planes 2 and 3. The contour level of 2.5 $\sigma$ is red and 3.5 $\sigma$ is blue. *a*, Molecule A is shown with the narrow groove containing the 3′ ends of the molecule at the front. *b*, Molecule B is shown with the narrow groove containing the 5′ ends at the front. The phosphate at the 5′ ends of both chains is missing and this narrow groove appears shorter than the groove on the other side of the molecule. In both A and B, the proximal end of the molecule closest to the pseudo 2-fold axis is at the top.

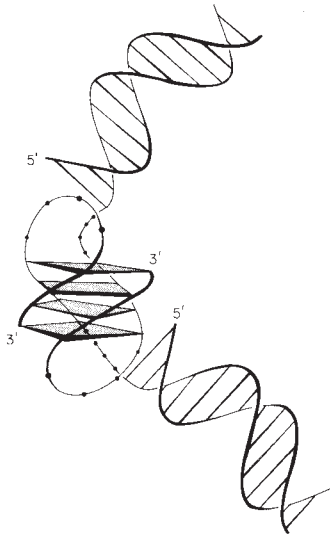



FIG. 6 A schematic diagram illustrating the manner in which the *Oxytricha* telomere structure may link together two DNA molecules or two chromosomes. The shaded squares represent the G quadruplexes and the solid dots represent the thymine residues. Small dots are also added for the four thymine residues immediately adjacent to the DNA duplex which are not found in the structure determined.

hydrogen-bonding and solvent or ion coordination. In the dimethylsulphate methylation protection experiments of Williamson *et al.*[9], an *Oxytricha* telomere model was proposed with only two G-quartets in potassium solutions because only levels 2 and 3 gave complete methylation protection, whereas levels 1 and 4 had much less. It is probable that this result is due to the stabilization between levels 2 and 3 cited above.

The arrangement of the thymine-loop structures is generally similar to a four-base-loop structure deduced from NMR studies[29]. The high-temperature factor and varied conformation of the thymines in the third position may be correlated with the fact that Williamson *et al.*[9] found a photochemical crosslink between two loops in a telomeric construct in which two loops were present at the same end of the molecule. The crosslink formed between two thymine residues in the third position. The varied positions of these thymine residues might facilitate forming this photo crosslink.

Existence of four-stranded guanine complexes has not been established *in vivo*. In *Oxytricha*, the DNA fragments in the macronucleus have telomere ends with two repeats of $d(T_4G_4)$ overhanging on the 3′ end[4]. It has long been recognized that these ends form a cohering structure *in vitro*, under conditions close to physiological[4,5,8]. This structure could be the basis for a linear assembly of the DNA fragments in the *Oxytricha* macronucleus. Using the structure that we have found here, the organization of different DNA fragments held together by the cohering telomeres is shown diagrammatically in Fig. 6. This shows that two DNA strands come together at opposite ends of the quadruplex structure and the two telomeric repeats at the 3′ end of $d(T_4G_4)$ are folded as in the crystal. This model has been drawn to illustrate the manner in which the *Oxytricha* quartet could hold DNA molecules together; however, it could be used for holding two chromosomes together in other systems because there appears to be a common sequence organization found in all telomeres examined.

Although we have focused on telomere organizations, the structure visualized here could also have other biological roles. For example, in sister chromatid exchange, a segment of the two sister chromatid genomes containing two short G-tracts separated by other bases could form a transient four-stranded structure of this type. Its stability would compensate the fact that the normal DNA duplex must undergo strand separation for this to form. But it could be a feature of chromatid alignment. In a similar way, this structure could also be important in keeping two identical polynucleotide chains together, as for example in the two copies of the RNA genome present in all retroviruses, including human immunodeficiency virus[30]. □

Received 6 November 1991; accepted 7 February 1992.

1. Blackburn, E. *Nature* **350,** 569-572 (1991).
2. Klobutcher, L. A., Swanton, M. T., Donini, P. & Prescott, D. M. *Proc. natn. Acad. Sci. U.S.A.* **78,** 3015-3019 (1981).
3. Henderson, E. R. & Blackburn, E. H. *Molec. cell. Biol.* **9,** 345-348 (1989).
4. Lipps, H. J., Gruissem, W. & Prescott, D. M. *Proc. natn. Acad. Sci. U.S.A.* **79,** 2495-2499 (1982).
5. Oka, Y. & Thomas, C. A. *Nucleic Acids Res.* **15,** 8877-8898 (1987).
6. Henderson, E., Hardin, C. C., Walk, S. K., Tinoco, I. & Blackburn, E. H. *Cell* **51,** 899-908 (1987).
7. Cech, T. R. *Nature* **332,** 777-778 (1988).
8. Lipps, H. J. *Proc. natn. Acad. Sci. U.S.A.* **77,** 4104-4107 (1980).
9. Williamson, J. R., Raghuraman, M. K. & Cech, T. R. *Cell* **59,** 871-880( 1990).
10. Sundquist, W. I. & Klug, A. *Nature* **342,** 825-829 (1990).
11. Panyutin, I. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **87,** 867-870 (1990).
12. Rich, A. *Biochim. biophys. Acta* **29,** 502-509 (1958).
13. Arnott, S., Chandrasekaran, R. & Marttila, C. M. *Biochem. J.* **141,** 537-543 (1974).
14. Zimmerman, S. B., Cohen, G. H. & Davies, D. R. *J. molec. Biol.* **92,** 181-192 (1975).
15. Gellert, M., Lipsett, M. N. & Davies, D. R. *Proc. natn. Acad. Sci. U.S.A.* **48,** 2013-2018 (1962).
16. Sen, D. & Gilbert, W. *Nature* **334,** 364-366 (1988).
17. Wang, Y., Jin, R., Gaffney, B., Jones, R. A. & Breslauer, K. J. *Nucleic Acids Res.* **19,** 4619-4622 (1991).
18. Wang, Y. *et al. J. molec. Biol.* (in the press).
19. Smith, F. W. & Feigon, J. *J. biomolec. Struct. Dynamics* **8,** 201 (1991).
20. Brunger, A. T. *Acta crystallogr.* A **46,** 46-57 (1990).
21. Zimmerman, S. B. *J. molec. Biol.* **106,** 663-672 (1976).
22. Brunger, A. T. *X-PLOR Manual, Version 2.1* Yale University (1990).
23. Jones, T. A. *J. appl. Crystallogr.* **11,** 268-276 (1978).
24. Pinnaraia, T. J. *et al. J. Am. chem. Soc.* **100,** 3625-3627 (1978).
25. Howard, F. B. & Miles, H. T. *Biochemistry* **21,** 6736-6745 (1982).
26. Williams, L. D. *et al. Proc. natn. Acad. Sci. U.S.A.* **87,** 2225-2229 (1990).
27. Rich, A., Davies, D. R., Crick, F. H. C. & Watson, J. D. *J. molec. Biol.* **3,** 71-86 (1961).
28. Rich, A., Nordheim, A. & Wang, A. H.-J. *A. Rev. Biochem.* **53,** 791-846 (1984).
29. Hare, D. R. & Reid, B. R. *Biochemistry* **25,** 5341-5350 (1986).
30. Kung, H. J. *et al. Cell* **1,** 609-620 (1976).

ACKNOWLEDGEMENTS. We thank A. M. deVos (Genentech) and N. Xuong (Multiwire Detector Center at University College, San Diego) for assistance in data collection and advice, L. Williams, J. Williamson, N. Usman, M. Egli, S. Zhang, A. Herbert, C. Lockshin, Q. Gao (MIT), C. Ogata and W. Weis (Columbia University) for discussions, and J. Williamson for help in obtaining preliminary coordinates. This research was supported by grants at MIT from the NIH, the National Science Foundation, the Office of Naval Research, the American Cancer Society, the National Aeronautics and Space Administration and a US Department of Energy grant (to R.M.). The UCSD Research Resource is funded by the NIH. C.K. is a Walter Winchell-Damon Runyon fellow. X.Z. is a postdoctoral fellow of the Human Science Frontiers Program.