Tertiary Structure Motif of *Oxytricha* Telomere DNA†

Ke Yu Wang,‡ S. Swaminathan,§ and Philip H. Bolton*,‡

Chemistry Department, Wesleyan University, Middletown, Connecticut 06459, and Gilead Sciences, 346 Lakeside Drive, Foster City, California 94404

Received December 9, 1993; Revised Manuscript Received February 8, 1994®

ABSTRACT: The tertiary structure of a single-stranded DNA containing the sequence of *Oxytricha* telomere DNA has been determined. This DNA adopts a compact tertiary structure that consists of four base-paired tetrads of guanine residues which are connected by three loops. The tetrads show significant deviations from planarity, and two of the loops exhibit significant loop–loop interactions. The structure of this telomere contains *syn*-thymine residues, which are in the loops, as well as an intraloop pyrimidine–pyrimidine base pair between residues that are separated by a single residue. The tertiary structure of the telomere DNA is consistent with prior results that showed that two thymines distant in sequence could be photo-cross-linked. The overall folding pattern of this telomere DNA is similar to that previously determined for a DNA aptamer, which binds to and inhibits thrombin, though the details of the two structures are quite distinct.

Telomere DNA occurs at the ends of the chromosomes of eukaryotes and has several highly conserved features (Blackburn, 1990, 1991, 1992). One of the prime functions of telomere DNA is to protect the chromosomes from shrinking during each round of replication (Blackburn, 1990, 1991, 1992; Zakian, 1989). The standard model for DNA replication predicts that linear chromosomes are shortened in each round of replication due to the nature of leading and lagging strand replication. This shortening of the DNA is known as the "end-replication problem". Telomere DNA is both noncoding and synthesized independently of other DNA, and these features can solve the problem of chromosome shrinkage. Telomere DNA is synthesized by telomerases that are reverse transcriptases that have a captive RNA (Zakian, 1989; Romero & Blackburn, 1991).

Telomere DNA may also be important in the organization of chromosomes in the nucleus, in preventing fusion between chromosomes, and in aligning the chromosomes during cell division (Blackburn 1990, 1991, 1992; Price, 1990). Evidence for the importance of the ends of chromosomes was known long before the discovery of telomere DNA (McClintock, 1939, 1942). Many of the studies performed on telomere DNA have used *Oxytricha* since this organism has $\approx 2 \times 10^7$ macronuclear chromosomes and each has the same telomere DNA (Gray et al., 1991).

There has been considerable interest in telomere DNA and not only for gaining understanding of its fundamental roles in protecting and organizing chromosomes. The length of the telomere DNA may decrease as a cell ages, and when telomere DNA is significantly degraded, the cell dies (Levy et al., 1992). There is evidence that without telomere DNA cell death occurs (Blackburn, 1990, 1991, 1992; Levy et al., 1992). Varmus and co-workers have studied the length distribution of human telomere DNA and have shown that the length of the telomere DNA decreases with cell age (de Lange et al., 1990). There have also been suggestions that preservation of the telomere DNA of cells could prevent cell aging and death (Blackburn, 1991; Levy et al., 1992). There have also been speculations that immortalized cells have sufficient telomerase activity to maintain their telomere DNA indefinitely and that this activity is crucial to retaining their immortalized state (Zakian, 1989; Blackburn, 1992; de Lange et al., 1990). These possible relationships to cell aging and cancer have considerably added to the interest in telomere DNA.

The length of telomere DNA varies from organism to organism as well as during the lifetime of a cell. The current consensus picture for telomere DNA is

$$5'\text{-}G_nT_mG_nT_mG_nT_m\ldots G_nT_mG_nT_mG_nT_mG_nT_m\text{-}3'$$

$$3'\text{-}C_nA_mC_nA_mC_nA_m\ldots C_nA_mC_nA_m\text{-}5'$$

Telomere DNA consists of up to hundreds or thousands but in some organisms only a few repeats of a simple DNA sequence that ends in a 3′ "overhang" of two repeats on the G-rich strand. In some organisms the sequences contain A or C residues in the G-rich strand (Blackburn, 1990, 1991, 1992; Zakian, 1989). Cech and co-workers have shown that the overhang appears to associate with a protein heterodimer in *Oxytricha* (Hicke et al., 1990; Gray et al., 1991). The sequence and structure of duplex telomere DNA may be important in the specific recognition by proteins (Blackburn, 1990, 1991, 1992; Hicke et al., 1990; Gray et al., 1991). In yeast, the repressor activator protein RAP1 appears to associate with duplex telomere DNA (Conrad et al., 1990; Gilson et al., 1993).

There have been a number of investigations of the NMR (Hardin et al., 1991, 1992; Macaya et al., 1993; Wang et al., 1991; Henderson et al., 1987; Smith & Feigon, 1992; Wang & Patel, 1992; Scaria et al., 1992), ultraviolet absorption (Hardin et al., 1991, 1992; Sen & Gilbert, 1990; Henderson et al., 1987), circular dichroism (Guschlbauer et al., 1990; Hardin et al., 1991; Zahler et al., 1991), gel electrophoresis (Guschlbauer et al., 1990; Hardin et al., 1991; Guo et al., 1992, 1993; Williamson et al., 1989; Sen & Gilbert, 1988, 1990; Sundquist & Klug, 1989), nuclease sensitivity (Voloshin et al., 1992; Blackburn, 1991), photo-cross-linking (Williamson et al., 1989), chemical protection (Williamson et al., 1989; Henderson et al., 1987), and potassium dependence (Guschlbauer et al., 1990; Hardin et al., 1991, 1992; Williamson et al., 1989; Sen & Gilbert, 1990; Wang et al., 1991, 1993a,b;

† The research at Wesleyan was supported, in part, by the Patrick and Catherine Weldon Donaghue Medical Research Foundation.

‡ Wesleyan University.

§ Gilead Sciences.

® Abstract published in *Advance ACS Abstracts,* June 1, 1994.

0006-2960/94/0433-7517$04.50/0

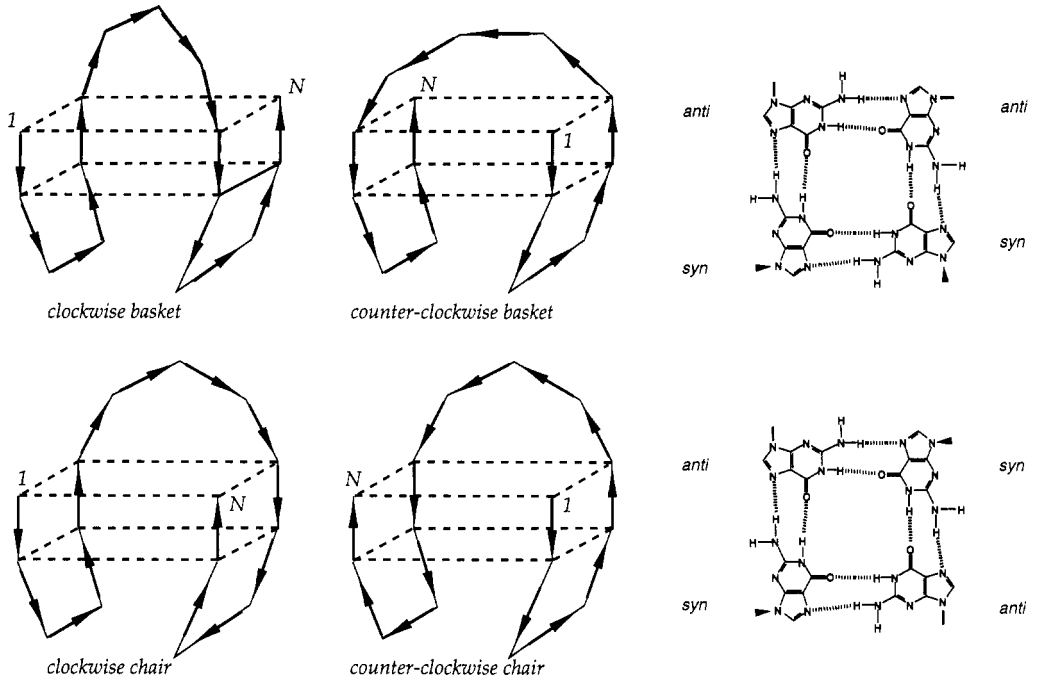

FIGURE 1: Schematic depictions of the two enantiomorphs of the basket form of the telomere DNA as well as the two enantiomorphs of the chair form of the telomere DNA are shown. The arrows show the 5′ to 3′ direction of the DNA. The figure also contains a depiction of the syn–syn–anti–anti tetrad associated with the basket forms as well as the syn–anti–syn–anti tetrad associated with the chair forms. The DNA aptamer d(GGTTGGTGTGGTTGG) has been previously shown to adopt a counterclockwise chair structure in solution.

Henderson et al., 1987; Zahler et al., 1991; Kang et al., 1992; Wang & Patel, 1992; Smith & Feigon, 1992, 1993; Scaria et al., 1992) of the properties of DNAs containing the general sequence G_mT_n.

The structure of d(GGTTGGTGTGGTTGG) was recently determined (Wang et al., 1993a,b). This DNA molecule binds to and inhibits thrombin both in vitro and in animal models (Bock et al., 1992; Griffin et al., 1993). An aptamer is a DNA that binds to a specific molecular target (Bock et al., 1992; Wang et al., 1993a,b) following the nomenclature and methodology originally proposed for RNA molecules that bind to specific molecular targets (Ellington & Szostak, 1990). The tertiary structure of this aptamer consists of two G tetrads that are connected by three loops as shown in Figure 1. The G tetrads alternate syn–anti–syn–anti as depicted in Figure 1. The novel structural motif of the DNA aptamer was proposed to be relevant to that of telomere DNA. These basic structural features of the aptamer were recently confirmed (Macaya et al., 1993).

The crystal structure of a dimer of d(GGGGTTTTGGGG), which contains one and a half repeats of the telomere DNA sequence of *Oxytricha*, was determined by Rich and co-workers (Kang et al., 1992). The G tetrads in the crystal structure are anti–syn–anti–syn, the 5′-GG-3′ alternate between syn and anti, and the structure contains potassium. These are the same basic structural features observed for the tetrads of the aptamer in solution. This crystal structure is of a DNA dimer and has two loops, whereas the aptamer is a monomer and has three loops. The foldback structure with three loops of the aptamer structure is reminiscent of ones proposed by Cech and co-workers for single-stranded telomere DNA (Hardin et al., 1991; Zahler et al., 1991), and this structural motif may be one which telomere DNA can adopt. An intramolecular structure with three loops is consistent with the data on telomere DNA photo-cross-linking and the observed pauses in telomerase synthesis that occur when four G regions are synthesized (Hardin et al., 1991; Zahler et al., 1991).

The aptamer structure and the *Oxytricha* crystal structure results are distinct from a structure of a dimer of d(GGGGTTTTGGGG) proposed on the basis of NMR results that has G-G-G-G tetrads that are syn–syn–anti–anti (Smith & Feigon, 1992, 1993). There have also been studies on a range of tetrad structures formed by tetramers of oligonucleotides (Wang et al., 1992; Scaria et al., 1992; Cheong & Moore, 1992).

It was unexpected to find that the structure of telomere DNA and an aptamer that binds to and inhibits thrombin may have similar sequence and structural features. This remarkable coincidence suggested that this aptamer may have interesting interactions with the proteins responsible for synthesizing and stabilizing telomere structure and that this structural motif may be useful in inhibiting enzymes other than thrombin (Wang et al., 1993a,b). This similarity also suggested that the tetrad structure of the aptamer may turn out to be a basic structural element of DNA and may be exhibited by aptamers to targets other than thrombin as well as in naturally occurring systems including telomeres (Wang et al., 1993a,b). The structure of the aptamer and the structure of the *Oxytricha* telomere DNA, presented below, have suggested a rationale for all known telomere sequences in terms of their secondary and tertiary structures (K. Y. Wang and P. H. Bolton, in preparation).

In this article we report on the structure of the single strand of DNA containing the *Oxytricha* telomere DNA sequence. The structure determination used NMR data obtained from 400 to 750 MHz and under a variety of experimental conditions. The structure of the telomere DNA is quite similar to the structure of the DNA aptamer. The structure of the *Oxytricha* telomere DNA is consistent with the prior photo-cross-linking and biological data on this DNA.

EXPERIMENTAL PROCEDURES

The NMR experiments were carried out on DNA samples of the sequence 5′-$G_1G_2G_3G_4T_5U_6T_7U_8G_9G_{10}G_{11}G_{12}U_{13}$-$T_{14}U_{15}T_{16}G_{17}G_{18}G_{19}G_{20}U_{21}U_{22}T_{23}T_{24}G_{25}G_{26}G_{27}G_{28}$-3′. The

DNA was obtained from Pharmacia and the purity of the sample checked by both NMR and HPLC. This sequence was chosen to aid the assignment process since each of the three pyrimidine tracts has a unique sequence. The DNA samples were annealed by heating to 90 °C and allowed to cool to room temperature. The proton NMR spectrum of this DNA was obtained with the sample in the presence and absence of potassium. The results indicated that in the presence of potassium multiple conformations of the telomere DNA are present. The NOESY, ROESY, TOCSY, and other NMR experiments were carried out in the usual fashion (Ernst et al., 1987; Clore & Gronenborn, 1989; Van de Ven & Hilbers, 1988; Wüthrich, 1986). All of the data were processed using Varian VNMR except where indicated. The 750-MHz experiment was carried out with a sample of 90 OD_{260} of DNA and the other experiments with a sample of 180 OD_{260}. The extinction coefficient at 260 nm is 2.7×10^5 to an accuracy of 15%. The samples were all in 0.6 mL of solution.

For the 750-MHz experiments, the sample of the telomere DNA was in 90% H_2O–10% ^2H_2O in a buffer of 140 mM NaCl and 20 mM perdeuteriated Tris at a pH of 7.2. The sample for all the other experiments was at a concentration of 180 OD_{260} in a buffer of 140 mM NaCl, 20 mM perdeuteriated Tris, 1 mM EDTA, and 1 mM EGTA at a pH of 6.0. A shaped and shifted pulse was used for water suppression (Smallcombe, 1993).

A 200-ms NOESY in 90% H_2O–10% ^2H_2O was obtained using a Varian Unityplus 750 spectrometer at 10 °C using the States–Haberkorn method. The data were collected into 4096 × 2 complex points in t_2 and 838 × 2 complex points in t_1 with spectral width of 12999.7 Hz in each dimension. A total of 32 transients were obtained for each t_1 increment. The time domain data were multiplied by a Gaussian weighting in both dimensions. The final size of the spectrum was 4096 by 4096 real points.

A 250-ms NOESY in 90% H_2O–10% ^2H_2O was obtained using a Varian Unityplus 500 spectrometer at 20 °C using the States–Haberkorn method. The data were recorded with a spectral width of 9000.9 Hz, 2048 × 2 complex points in t_2, and 300 × 2 FIDs in t_1, and 160 transients were obtained for each increment. Linear prediction was used to calculate the first three points in t_2 and the first five points in t_1. Recursive linear prediction was used to extend the number of points from 281 to 1024 in t_1. A phase-shifted sine weighting was used in t_2 and a Gaussian function in t_1. The spectra were zero-filled to 4096 by 2048 real points.

A 400-ms NOESY and a 80-ms ROESY were obtained using a Varian Unityplus 400 spectrometer at 30 °C in ^2H_2O using the States–Haberkorn method. In both cases, the data were collected into 1024 × 2 complex points in t_2 and 256 × 2 points in t_1 with spectral widths of 3399.9 Hz in each dimension. The NOESY data set was acquired with 128 transients for each t_1 value and the ROESY data set with 32 transients for each t_1 value. A Gaussian weighting was used in both dimensions in both cases, and the spectra were zero-filled to 2048 by 2048 real points.

A 400-ms 600-MHz NOESY spectrum was obtained on the sample in ^2H_2O at 30 °C using a Bruker AM spectrometer using the TPPI method. The spectral width was 6024 Hz in each dimension with 2048 complex points in F_1 and 600 increments of t_1, and 128 transients were acquired for each increment of the evolution time. The data were processed using FELIX 2.1. The time domain data was processed with sine bells in both dimensions and zero-filled to 2048 points in both dimensions before Fourier transformation.

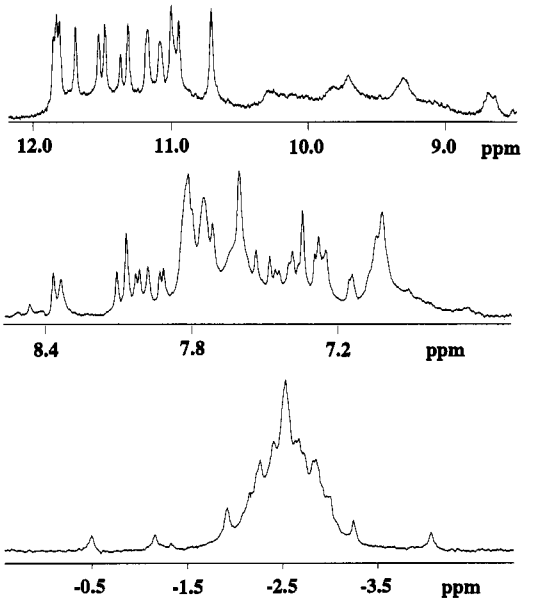

FIGURE 2: The top spectrum is of the imino proton region of the telomere DNA obtained at 500 MHz with the sample in 90% H_2O–10% ^2H_2O. The spectrum shown in the middle is of the aromatic proton region of the telomere DNA obtained at 500 MHz with the sample in 90% H_2O–10% ^2H_2O. The spectrum shown at the bottom is the ^{31}P spectrum of the telomere DNA obtained at 161 MHz. All of these spectra were obtained with the sample at 20 °C.

One-dimensional ^{31}P NMR spectra were obtained at 161.9 MHz with proton decoupling. The spectral width was 2000 Hz with 3264 complex points and 2048 scans.

RESULTS AND DISCUSSION

The first step in the structure determination was to ascertain that only one form of the DNA was present in solution. If only one form is present in solution, then there should be a single resonance for each proton. The resonances of the H8 and H6 occur in the region between 7.0 and 8.5 ppm. Figure 2 contains the one-dimensional proton spectrum of this region. The two-dimensional spectra of this region, discussed below, showed that there are cross-peaks associated with the expected 28 aromatic resonances. The resonances have line widths that are consistent with the single form being a monomer and not a dimer or higher order complex.

The imino proton region offers information about the number of imino protons that are protected from exchange. The imino region of the proton spectrum is shown in Figure 2. The total intensity of the imino region, from 10.5 to 12 ppm, is consistent with there being ≈20 imino protons protected from exchange. The number of imino resonances was determined by comparison of the integral of the imino region with that of the aromatic, H1′,H5, and H2′,H2″ regions. The imino region of the spectrum also indicates that there is only one form of the molecule in solution as the imino resonances all have unit intensity.

The ^{31}P spectrum was obtained to determine the presence of any upfield- or downfield-shifted resonances. Figure 2 contains the ^{31}P NMR spectrum of the telomere DNA. The spectrum exhibits about a 3.5-ppm dispersion that is similar to that of the aptamer. Both the *Oxytricha* telomere DNA and the aptamer have phosphorus resonances shifted to both higher and lower field than the central region.

The thermal stability of the telomere DNA was determined by examination of the temperature dependence of the line widths of the imino protons. As a DNA melts, the exchange rates of the imino protons with water increase and the line

widths of the imino resonances increase. This method showed that the melting temperature of the DNA is about 75 °C. At room temperature the exchange rates of the imino protons with water are on the time scale of days, or longer, which is consistent with the presence of tetrads (Smith & Feigon, 1992, 1993; Wang et al., 1993a,b). This slow rate of exchange indicates that the structure is kinetically stable. The room temperature exchange rates were determined by monitoring the exchange of the imino protons with ^2H_2O.

The effect of potassium on the NMR properties of the telomere DNA was examined by obtaining the proton NMR spectrum of the telomere DNA in the presence of up to 100 mM potassium. It was found that in the presence of potassium multiple structures of the telomere DNA are present.

These NMR results showed that the telomere DNA adopts a single structure in solution in which about 20 imino protons are most likely involved in base pairing. The backbone apparently contains a range of conformations since there is considerable dispersion of the phosphorus resonances. The structure is both thermally and kinetically quite stable.

Evidence for the Presence of syn-G and syn-T Residues. The determination of the number and types of syn residues has been typically made by the examination of the relative intensity of the NOE cross-peaks of H8/H6–H1′ cross-peaks (Clore & Gronenborn, 1989; Van de Ven & Hilbers, 1988). In syn residues the H6/H8–H1′ internuclear distance is short (Saenger, 1984) and the cross-peak is of high intensity (Wüthrich, 1986), and in anti residues the distance is long and the cross-peak is of low intensity. The NOESY spectrum of the aromatic–H1′ region of the telomere DNA is shown in Figure 3. A careful analysis of the NOESY data of G-tetrad-containing structures can be made to determine which cross-peaks are from syn and which are from anti residues (Smith & Feigon, 1992, 1993; Wang et al., 1993a,b).

The determination of the number of syn residues is more reliable when based on ROESY data such as that shown in Figure 3. In the NOESY experiment there are H8/H6–H1′ cross-peaks from anti residues that arise from spin diffusion. In a ROESY experiment the cross-peaks arising from spin diffusion are very weak, and of opposite algebraic sign, from those arising from direct short-range transfer. The ROESY spectrum of the H6/H8–H1′ region contains 16 strong cross-peaks. Eight of the cross-peaks are from *syn*-G residues, two are from *syn*-T residues, and six are from the H5–H6 cross-peaks of the six U residues. The assignments of these cross-peaks are given in Figure 3. The ROESY and NOESY data indicate that there is only one form of the molecule present in solution since only a single cross-peak is observed for each pair of protons and no exchange cross-peaks are observed.

There are only four possible folding patterns, for the telomere DNA, that are consistent with the NMR data presented in Figures 2 and 3 and the NOE connectivities, discussed below, which demonstrate that *all of the syn-G residues have anti 3′ neighbors* and that there are no imino–imino cross-peaks which can be assigned to G-T wobble base pairs. The imino protons of G-T wobble base pairs have extremely strong NOEs due to the short distance between the G and T imino protons of a G-T wobble base pair. These results indicate that the vast majority of the base pairs can only be G-G. Only G-tetrad-based structures are consistent with these results (Smith & Feigon, 1992, 1993; Wang et al., 1993a,b; Wang & Patel, 1992).

The four possible folding patterns, for the telomere DNA, are illustrated in Figure 1 along with the G tetrads associated with these structures. The counterclockwise chair, so called

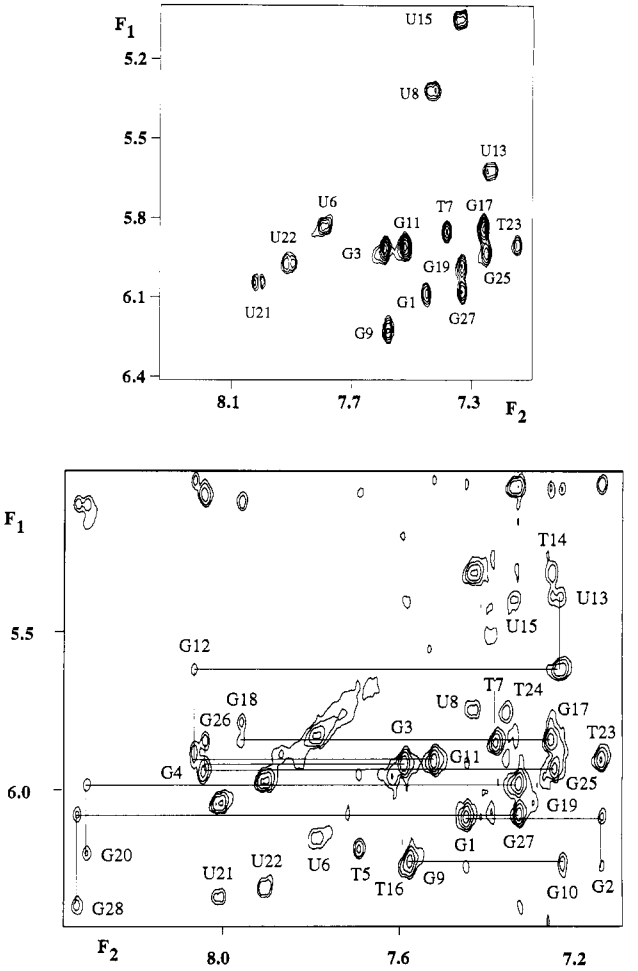

FIGURE 3: The two-dimensional spectrum shown at the top is part of the 400-MHz ROESY spectrum obtained with the sample in ^2H_2O. The G and T cross-peaks are from syn residues and correspond to the H8/H6–H1′ ROEs. The U cross-peaks correspond to the H5–H6 ROEs. The two-dimensional spectrum at the bottom is part of the 600-MHz NOESY spectrum obtained with the sample in ^2H_2O. This region contains intraresidue and interresidue H8/H6–H1′ NOEs as well as the H6–H5 NOEs of U residues. The lines indicate the pathways used to make sequential assignments, and the assignments of many of the cross-peaks are indicated.

because the tertiary structure resembles a chair and the strand orientation is counterclockwise from the viewpoint used, is the structure that is adopted by the aptamer in solution. The clockwise chair is *not* a mirror image of the counterclockwise chair due to the 5′–3′ nature of the DNA. The counterclockwise and clockwise forms also differ in the positions of their "wide" and "narrow" grooves as will be discussed below.

In addition to the chair structures there are also two possible enantiomorphs of the "basket" structure. The basket structure is named for its resemblance to a real basket with a crossover handle. The basket differs from the chair in many respects. First, since the syn–anti alternation is in the 5′–3′ direction, then in the basket form the G tetrads are composed of syn–syn–anti–anti tetrads whereas in the chair form the G tetrads are syn–anti–syn–anti. In a basket structure the top loop must cross over the top G tetrad while in a chair this does not occur. In the basket structures the bottom two loops run parallel to one another whereas in the chair structures the bottom two loops are antiparallel. While the sequential interactions in a basket structure and a chair structure can be quite similar, the interactions between residues distant in sequence can be very distinct. Long-range interactions were used to determine the structure of the aptamer in solution and

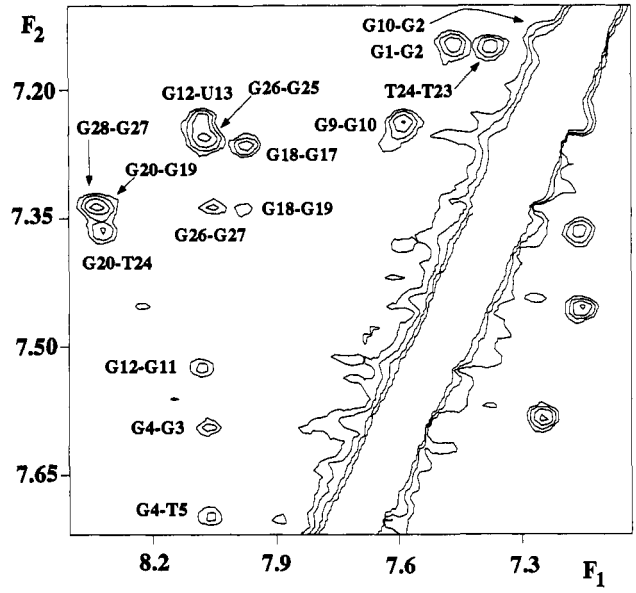

FIGURE 4: The two-dimensional spectrum is part of a 400-MHz NOESY obtained on the telomere DNA in ^2H_2O. The signals in this region are from 3′-GH8–GH8-5′ connectivities and 5′-GH8–TH6-3′ connectivities as well a 3′-TH6–TH6-5′ connectivity. The assignments of the cross-peaks are indicated in the figure.

to distinguish between the possible structures of the *Oxytricha* telomere DNA.

Assignment of the Proton Spectrum. The assignment strategy used for the telomere DNA is similar to that we used for the DNA aptamer (Wang et al., 1993a,b). The 5′-G-G-3′, with G indicating a *syn*-G residue, sequential connectivities were made using 3′ aromatic to 5′ H1′, H2′, H2″ NOEs. These NOEs are indicated in Figures 3 and 5. These NOEs allowed the connection of each of the eight *syn*-G residues to their 3′ neighbors. Three of these 5′-G-G-3′ sequences could be further connected on the basis of the usual anti–anti NOEs connecting the *anti*-G and the adjacent pyrimidine. The G_3-G_4-T_5 sequence could be identified on the basis of the G_4-T_5 NOEs, the G_{19}-G_{20}-U_{21} on the basis of the G_{20}-U_{21} NOEs, listed in Table 3, and the G_{11}-G_{12}-U_{13} on the basis of the G_{12}-U_{13} NOEs.

These assignments were confirmed and extended by the use of 5′-GH8–GH8-3′ connectivities. These connectivities were made on the basis of the NOE cross-peaks shown in Figure 4 and are observed for all 5′-GH8–GH8-3′ linkages in the molecule. The sequential 5′-G-G-3′ NOEs have been previously used to make assignments of the resonances of dimers of $G_4T_4G_4$ (Smith & Feigon, 1993).

NOE connectivities between the H8 protons of G_2 and G_{10}, G_{18} and G_{19}, and G_{26} and G_{27} were observed in addition to their 5′-GH8–GH8-3′ connectivities. The only NOE connectivity observed between *anti*-G residues is between G_2 and G_{10}. The G_{18}-G_{19} and G_{26}-G_{27} are the only 5′-G-G-3′ NOE connectivities observed. These three NOEs arise from the nonplanarity of the G tetrads and are discussed below.

The intraresidue thymine H6–methyl cross-peaks are easily identified in the NOE data that allowed the assignment of these signals as to residue type. Interresidue NOEs are observed from a thymine H6 at 7.59 ppm to a uridine H2′ whose H6 is at 7.34 ppm and from a thymine H6 at 7.25 ppm to a uridine H2′ whose H6 is at 7.24 ppm. These NOEs are indicative of 3′-T-U-5′ sequences. Both of these sequences are isolated in that they do not exhibit NOEs to any of the other T or U residues. The methyl and aromatic protons of these residues do have NOEs to G residues. These T and U

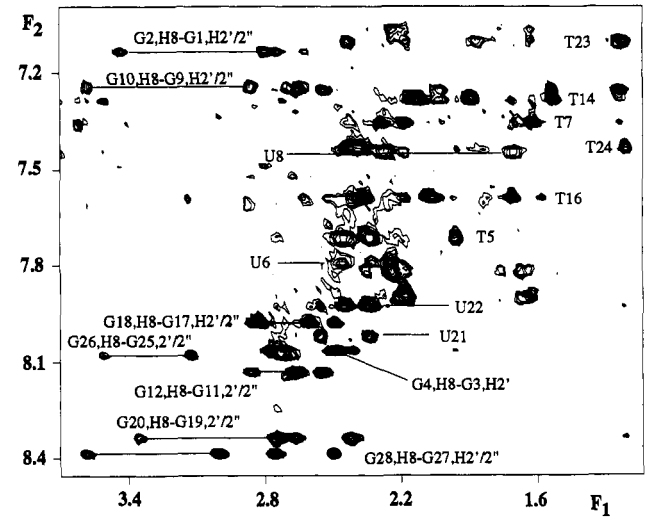

FIGURE 5: The two-dimensional spectrum is part of a 750-MHz NOESY obtained on the telomere DNA in 90% H_2O–10% ^2H_2O. The signals in this region are from intraresidue and interresidue H8/H6–H2′/H2″ NOEs. The lines indicate some of the interresidue cross-peaks, and the assignments of many of the cross-peaks are indicated.

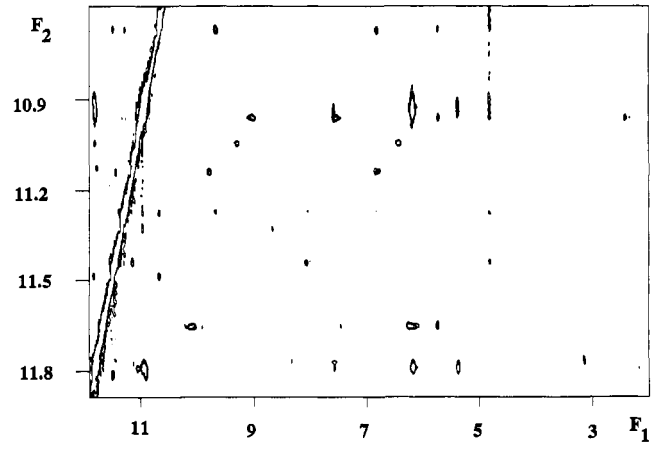

FIGURE 6: The two-dimensional spectrum shown is from a 500-MHz NOESY experiment on the telomere DNA in 90% H_2O–10% ^2H_2O. The spectrum contains the imino proton NOE connectivities to all of the other protons in the sample. The cross-peaks in this region are listed in Tables 2 and 3.

residues have a variety of NOEs, including ones between U_{15} and G_{17}, which has allowed the assignments of the resonances of 3′-T_{16}-U_{15}-T_{14}-U_{13}-5′.

There is a set of 3′-T-T-5′ connectivities observed in the NOE data involving both the H6–H2′/H2″ and H6–H1′ cross-peaks. Since there is only one T-T in the sequence, this connectivity allows assignment of T_{24}-T_{23}. Similarly, a set of typical anti–anti NOEs between two U residues allowed assignment of U_{22}-U_{21}. The U_8-T_7 sequence was assigned on the basis of the NOEs typical of a 3′-anti-5′-syn sequence. The only thymine left was thymine 5, and G4 was assigned on the basis of its connectivities to T_5. The assignment of the 3′-T_{16}-U_{15}-T_{14}-U_{13}-5′ sequence was confirmed by the G_{12}-U_{13}, U_{13}-T_{14}, T_{14}-T_{16}, U_{15}-T_{16} NOE connectivities. These assignments were confirmed on the basis of the long-range NOEs between these residues that are discussed below. In particular, there are NOEs between U_{22} and T_{24}, T_{24} and G_{20}, and T_7 and T_{23}.

Additional long-range NOEs were observed between the exchangeable imino and amino protons of the G tetrads and the pyrimidine residues, and some of these NOEs are shown in the spectrum in Figure 6. The only medium-intensity imino–

Table 1: Proton Chemical Shift Assignments at 30 °C[a]

	H6/H8	H1′	H2′	H2″	H3′	H4′	H5/CH_3
G1	7.45	6.08	3.40	2.84	5.03	4.34	
G2	7.14	6.24	2.65	2.65	5.04	4.35	
G3	7.59	5.92	2.56	2.41	4.94	4.17	
G4	8.04	5.95	2.75	2.50	5.07	4.43	
T5	7.69	6.18	2.33	2.45	4.86	4.25	1.95
U6	7.79	6.15	2.24	2.46	4.75	4.24	5.83
T7	7.38	5.85	2.17	2.26	4.64	4.00	1.65
U8	7.43	5.75	1.75	2.26	4.47	3.99	5.32
G9	7.58	6.22	3.53	2.88	4.95	4.34	
G10	7.23	6.25	2.71	2.63	5.06	4.52	
G11	7.52	5.91	2.86	2.65	5.03	4.17	
G12	8.06	5.88	2.56	2.66	4.94	4.46	
U13	7.24	5.39	1.29	2.06	4.45	3.79	5.62
T14	7.25	5.32	2.10	1.90	4.56	3.61	1.55
U15	7.34	5.40	1.61	2.20	4.44	3.79	5.05
T16	7.59	6.24	2.07	2.37	4.56	3.18	1.70
G17	7.26	5.84	2.51	2.84	4.89	4.26	
G18	7.96	5.78	2.62	2.82	5.10	4.27	
G19	7.33	5.98	3.37	2.68	5.16	4.53	
G20	8.31	6.20	2.75	2.43	5.11	4.53	
U21	8.01	6.34	2.36	2.57	4.92	4.44	6.04
U22	7.90	6.31	2.34	2.46	4.84	4.36	5.97
T23	7.15	5.90	1.90	2.43	4.80	3.91	1.32
T24	7.36	5.75	2.35	2.44	4.90	4.07	1.22
G25	7.25	5.93	3.45	2.73	4.79	4.26	
G26	8.04	5.84	2.71	3.09	5.08	4.19	
G27	7.33	6.08	3.60	2.99	5.10	4.33	
G28	8.33	6.36	2.76	2.50	4.83	4.34	

[a] The ^2H_2O signal was referenced at 4.72 ppm at 30 °C.

imino proton NOEs observed have been assigned to pyrimidine–pyrimidine base pairs since all of the G residues are involved in the G tetrads. The NMR data indicate that there are two pyrimidine–pyrimidine base pairs. At least one of the imino protons of each of these base pairs also has an NOE to an amino proton of a G tetrad.

The imino protons of the G imino protons could not be assigned to specific residues. The intraresidue and interresidue NOE connectivities of the G imino protons did not allow unambiguous assignments of these protons to be made. The G imino protons do exhibit weak imino–imino NOE connectivities, but these NOEs are apparently both intratetrad and intertetrad.

These assignments of the pyrimidines were combined with the assignment of the guanines to give the assignments of all of the nonexchangeable resonances of the telomere with the exception of the H5′,H5″ resonances. The assignments are given in Table 1, the intraresidue NOEs are given in Table 3, and the interresidue NOEs are given in Table 2. The sequential connectivity information, by itself, does not discriminate among the four possible structures. The discrimination among the four structures is based on the information present in the long-range NOEs that can only be interpreted after the sequential assignments have been determined.

The Telomere DNA Does Not Adopt a Basket Structure. The basket structures are inconsistent with a number of observed NOEs, and the basket structures predict a number of NOEs that are not observed. There are a number of NOEs between T_{24} and G_{20} that are observed. These NOEs are not consistent with any basket structure. For example, the T_{24} methyl to $G_{20}H1'$ distance is more than 0.8 nm in the basket structures while a strong NOE is observed between these protons. The basket structures are inconsistent with the number and intensity of the T_{24}-G_{20} NOEs that are listed in Table 2. Furthermore, all basket structures which we constructed predicted a number of NOEs between the aromatic protons of U_{13}-T_{14}-U_{15}-T_{16} and the G_1-G_{12}-G_{17}-G_{28} tetrad which are not observed. The predicted NOEs depend on the details of how the loop crosses over the tetrad. The basket structures also have the 5-6-7-8 loop, and the 21-22-23-24 loops parallel to one another. The interresidue NOEs between the bottom two loops and the bottom G tetrad are schematically shown in Figure 7. We could not construct any basket structure consistent with these NOEs. Therefore, the basket structures were eliminated from consideration. It had been suggested that this DNA might adopt a structure in the basket family by analogy to the solution structure of the dimer of $G_4T_4G_4$ (Smith & Feigon, 1992, 1993).

The Telomere DNA Adopts a Counterclockwise Chair Structure. There are two sets of NOEs that were the primary basis for the elimination of the clockwise chair structure. The observed NOEs between residues T_{24} and G_{20} are not consistent with a clockwise chair structure. The T_{24} methyl to $G_{20}H1'$ distance is more than 0.8 nm in a clockwise chair, as in the basket structures, while a strong NOE is observed between these protons. The other NOEs between T_{24} and G_{20} also

Table 2: Interresidue NOEs of *Oxytricha* Telomere DNA[a]

strong						***T14,H1′–T16,CH₃***	
	G2,H8–G1,H2′	G10,H8–G9,H2′	G18,H8–G17,H2′	G26,H8–G25,H2′	U22,H6–U21,H3′	***T24,CH₃–G20,H1′***	
	G2,H8–G1,H2″	G10,H8–G9,H2″	G18,H8–G17,H2″	G26,H8–G25,H2″			
	G4,H8–G3,H2′	G12,H8–G11,H2′	G20,H8–G19,H2′	G28,H8–G27,H2′			
	G4,H8–G3,H2″	G12,H8–G11,H2″	G20,H8–G19,H2″	G28,H8–G27,H2″			
medium	G2,H8–G1,H8	T5,CH_3–G4,H1′	U13,H6–G12,2′	G4,H8–G3,H1′	G20,H8–G19,H3′	***T24,H6–G20,H8***	
	G10,H8–G9,H8	T5,CH_3–G4,H2′	U13,H6–T14,CH_3	G12,H8–G11,H1′	G28,H8–G27,H3′	***G10,H8–G2,H8***	***T24,CH₃–G20,H8***
	G18,H8–G17,H8	U8,H6–T7,H2′	U13,H5–G12,H8	G10,H8–G9,H1′		***U15,H5–G17,H8***	***T24,CH₃–G20,H2″***
	G20,H8–G19,H8		U13,H5–T14,CH_3	G18,H8–G17,H1′			***T24,CH₃–U22,H1′***
	G26,H8–G25,H8		T14,H1′–U15,2′	G20,H8–G19,H1′			
	G28,H8–G27,H8		T23,H6–T24,CH_3	G26,H8–G25,H1′			
	U13,H6–G12,H8		T24,H6–T23,H2′	G28,H8–G27,H1′			
	T24,H6–T23,H6		T24,H6–T23,H3′				
weak	G4,H8–G3,H8	G2,H8–G1,H1′	T5,H6–G4,H1′	T23,H6–T24,H6		***T23,H6–T7,CH₃***	
	G12,H8–G11,H8	G2,H1′–G1,H8	T5,H6–G4,H2″	T24,H6–T23,H1′		***T23,CH₃–T7,H6***	***T24,CH₃–U22,H6***
	G19,H8–G18,H8	G4,H1′–G3,H8	T5,CH_3–G4,H8				
	G27,H8–G26,H8	G12,H1′–G11,H8	U13,H6–G12,H1′				
	T5,H6–G4,H8	G10,H1′–G9,H8	U13,H6–G12,H8				
		G18,H1′–G17,H8	T16,H6–U15,H1′				
		G20,H1′–G19,H8	T16,H6–U15,H2′				
		G26,H1′–G25,H8	T16,H6–U15,H2″				
		G28,H1′–G27,H8					

[a] The long-range NOEs are indicated by the bold italic letters. Also, since the imino protons are not assigned, NOEs involving some imino protons are not in the table.

Table 3: Intraresidue NOEs of *Oxytricha* Telomere DNA

	G1	G2	G3	G4	T5	U6	T7	U8	G9	G10	G11	G12	U13	T14
strong	H8–H1′ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H8–H2′ H8–H2″ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H8–H1′ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H8–H2′ H8–H2″ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H6–CH_3 H6–H2′ H6–H2″ H1′–H2″ H1′–H2′ H2′–H2″ H2′–H3′ H2″–H3′	H6–H5 H6–H2′ H6–H2″ H1′–H2′ H1′–H2″ H2′–H3′ H2″–H3′	H6–CH_3 H6–H1′ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H6–H5 H6–H2′ H6–H2″ H1′–H2′ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H8–H1′ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H8–H2′ H8–H2″ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H8–H1′ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H8–H2′ H8–H2″ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H6–H5 H6–H2′ H1′–H2′ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H6–CH_3 H6–H2′ H6–H2″ H1′–H2″ H1′–H2′ H2′–H2″ H2′–H3′ H2″–H3′
medium	H1′–H2′ H1′–H4′ H2″–H3′ H3′–H4′	H1′–H2′ H1′–H3′ H1′–H4′ H3′–H4′	H1′–H2′ H1′–H4′ H2″–H3′ H3′–H4′	H1′–H2′ H1′–H4′ H3′–H4′	H6–H1′ H1′–H4′ H3′–H4′	H6–H1′ H1′–H4′	H6–H2′ H6–H2″ H6–H3′ H1′–H2′ H1′–H4′	H6–H1′ H1′–H4′	H1′–H2′ H1′–H4′ H2″–H3′ H3′–H4′	H1′–H2′ H1′–H4′ H3′–H4′	H1′–H2′ H1′–H4′ H2″–H3′ H3′–H4′	H1′–H2′ H1′–H4′ H3′–H4′	H6–H3′ H1′–H4′	H6–H1′ H6–H3′ H1′–H2′ H3′–H4′
weak	H8–H2′ H8–H2″ H8–H3′ H1′–H3′	H8–H1′ H8–H3′ H1′–H3′	H8–H2′ H8–H2″ H8–H3′ H1′–H3′	H8–H1′ H8–H3′ H1′–H3′	H6–H3′ H1′–H3′	H1′–H3′ H6–H3′	H6–H4′ H1′–H3′	H6–H3′ H6–H4′ H1′–H3′ H1′–H4′	H8–H2′ H8–H2″ H8–H3′ H1′–H3′	H8–H1′ H8–H3′ H1′–H3′	H8–H2′ H8–H2″ H8–H3′ H1′–H3′	H8–H1′ H8–H3′ H1′–H3′	H6–H1′ H1′–H3′	H1′–H4′ H1′–H3′

	U15	T16	G17	G18	G19	G20	U21	U22	T23	T24	G25	G26	G27	G28
strong	H6–H5 H1′–H2′ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H6–CH_3 H6–H2′ H6–H2″ H1′–H2″ H2′–H2″ H2′–H3′	H8–H1′ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H8–H2′ H8–H2″ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H8–H1′ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H8–H2′ H8–H2″ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H6–H5 H6–H2′ H1′–H2′ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H6–H5 H6–H2′ H1′–H2′ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H6–CH_3 H6–H1′ H1′–H2″ H1′–H2′ H2′–H2″ H2′–H3′ H2″–H3′	H6–CH_3 H1′–H2″ H1′–H2′ H2′–H2″ H2′–H3′ H2″–H3′ H1′–H4′	H8–H1′ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H8–H2′ H8–H2″ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H8–H1′ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′	H8–H2′ H8–H2″ H1′–H2″ H2′–H2″ H2′–H3′ H2″–H3′
medium	H1′–H4′	H1′–H2′ H2″–H3′ H1′–H3′ H3′–H4′ H6–H1′	H1′–H2′ H1′–H4′ H2″–H3′ H3′–H4′	H1′–H2′ H1′–H3′ H1′–H4′ H3′–H4′	H1′–H2′ H1′–H4′ H2″–H3′ H3′–H4′	H1′–H2′ H1′–H3′ H1′–H4′ H3′–H4′	H6–H1′ H6–H2″ H6–H3′ H3′–H4′	H6–H1′ H6–H2″ H6–H3′ H3′–H4′	H6–H2′ H3′–H4′	H6–H1′ H6–H3′ H3′–H4′	H1′–H2′ H1′–H3′ H1′–H4′ H3′–H4′	H1′–H2′ H1′–H4′ H3′–H4′	H1′–H2′ H1′–H4′ H2″–H3′ H3′–H4′	H1′–H2′ H3′–H4′
weak	H6–H1′ H6–H2′ H6–H2″ H6–H3′ H1′–H3′	H6–H3′ H1′–H4′	H8–H2′ H8–H2″ H8–H3′ H1′–H3′	H8–H1′ H8–H3′	H8–H2′ H8–H2″ H8–H3′ H1′–H3′	H8–H1′ H8–H3′	H1′–H3′ H1′–H4′	H1′–H3′ H1′–H4′	H6–H2″ H6–H3′ H6–H4′ H1′–H4′ H1′–H3′	H1′–H3′	H8–H2′ H8–H2″ H8–H3′	H8–H1′ H8–H3′ H1′–H3′	H8–H2′ H8–H2″ H8–H3′ H1′–H3′	H1′–H4′ H8–H1′ H8–H3′ H1′–H3′

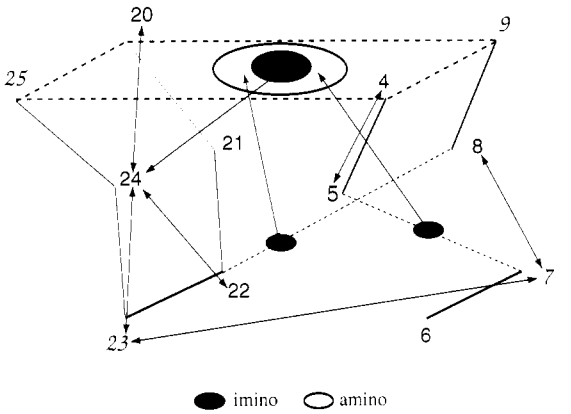

FIGURE 7: A schematic depiction of the bottom two loops of the telomere DNA along with the interresidue NOEs is shown. The interresidue NOEs are indicated by the double-headed arrows, and the dashed lines connecting residues 8 and 22 as well as 5 and 7 indicate pyrimidine–pyrimidine base pairs. The imino protons of the G quartet and the pyrimidine–pyrimidine base pairs are indicated by the shaded ellipsoids and the amino protons of the G tetrad by the unfilled ellipsoid. These interresidue NOEs were used to obtain the structure of the telomere DNA. The residue numbers of the syn residues—7, 9, 23, 25—are given in italics.

rule out a clockwise chair structure. The G_2H8–G_{10}H8 NOE is also not consistent with the clockwise chair structure. In a clockwise chair structure the G_2 and G_{10} are on adjacent tetrads and separated by a wide groove while in a counterclockwise chair the separation is across a narrow groove. The clockwise chair is in significant disagreement with the observed NOEs, and this structure cannot be correct. There were only four structures possible, and three have been eliminated by the comparison of the predicted and the experimental results. Therefore, the structure of the telomere DNA in solution must be a counterclockwise chair.

The Counterclockwise Chair Structure. The structure of the telomere DNA in the counterclockwise chair folding pattern was determined by the structure determination procedures described previously (Wang et al., 1993a,b). The NOE information and the number and types of base-pairing interactions were used as constraints. The nonsequential NOEs provided the key information. The backbone of the molecule is shown in Figure 8 in three different orientations. The "front" view corresponds to that shown in Figure 1 with residue 1 to the front of the structure in the upper right. This view illustrates the compact nature of the structure of the telomere DNA. This view also shows the relative widths of the wide groove of the DNA at the front and the narrow grooves on the sides. The second view corresponds to a rotation of the DNA by 45° about the z axis. This view emphasizes the symmetry of the structure of the DNA. The top loop, however, can be seen to be quite asymmetric relative to the G tetrads. The third view corresponds to a further 45° rotation about the z axis. In this view the irregularities in the backbone of the DNA structure can be seen.

The four G tetrads of the structure are shown in the same orientations in Figure 9. It is seen that the individual G tetrads are not planar, nor do the G tetrads stack in a simple pattern. The structure shows that the top two G tetrads are curved toward the center of the molecule as do the bottom two G tetrads. Thus, the top two G tetrads and the bottom two G tetrads apparently have opposite curvature, which suggests that two G tetrads are a structural repeat unit. This curvature of the G tetrads brings the H8 protons of G_{18} and $\underline{G}_{19}$ as well as those of G_{26} and $\underline{G}_{27}$ relatively close together. 5′-*anti*-GH8–3′-*syn*-GH8 connectivities are not observed from any other sites in the molecule.

A consequence of the nonplanarity of the G tetrads is that the H8 protons of residues G_2 and G_{10} are placed in close spatial proximity. This close approach of the H8 protons of nonadjacent G residues only occurs for these two G residues in the structure, and the only nonsequential G-G NOE observed is between G_2 and G_{10}. This nonsequential G-G NOE is across the narrow groove of the DNA. Structures that separate these two G residues by a wide groove are not consistent with the NMR data.

Some of the most interesting structural features of the telomere DNA are found in the bottom two loops. A number of nonsequential NOEs were observed for these residues including NOEs between the two loops. The key NOEs are illustrated in Figure 7. These NOEs were combined with the NOE evidence for two pyrimidine–pyrimidine base pairs to construct the structure of the loops shown in three orientations in Figure 10. This structure has T_{24} stacked on the bottom G tetrad as is T_5. T_{24} is more stacked on the G tetrad than is T_5. One pyrimidine–pyrimidine base pair is formed by the 5′-anti–anti–syn-3′ sequence T_5-U_6-$\underline{T}_7$ with the first and third pyrimidines base paired and the middle base pushed out. This base pair is more or less parallel to the bottom G tetrad. The

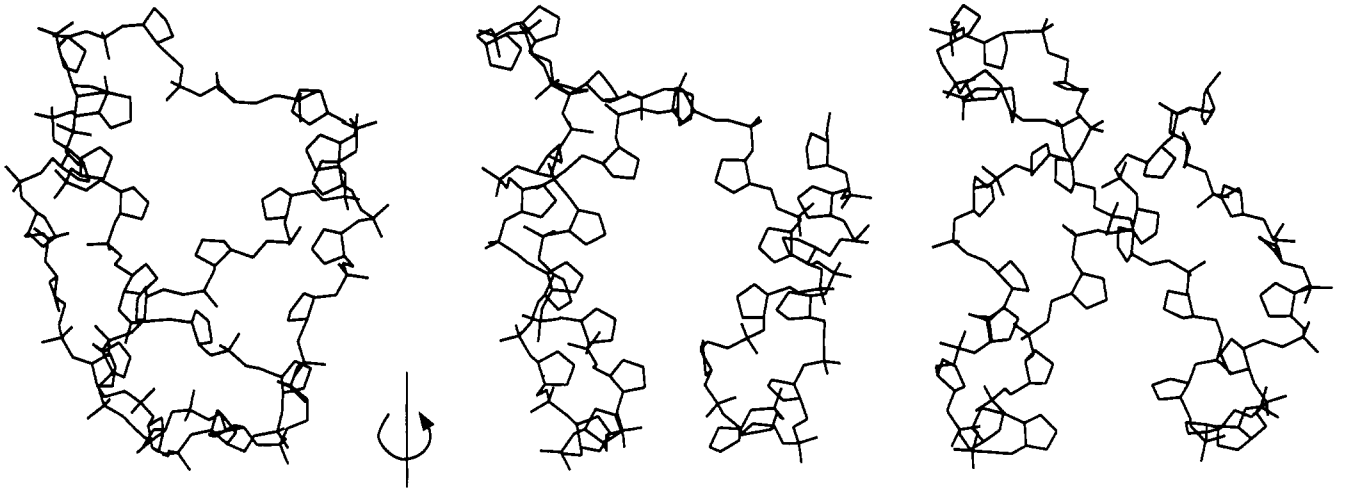

FIGURE 8: The structure of the backbone of the telomere DNA is shown from three viewpoints. These structures do not contain the bases of the telomere DNA. The structure on the left is shown from the same viewpoint as used in Figure 1. The structure in the middle is that obtained by rotation of the structure about the indicated z axis by 45°, and the structure on the right is that obtained by an additional rotation by 45°.

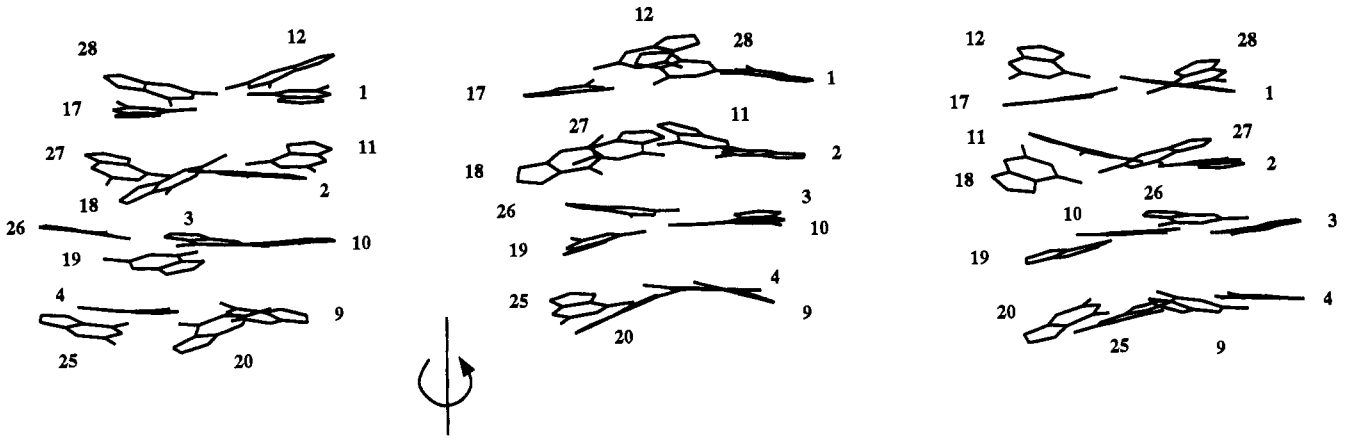

FIGURE 9: The structures of the four G tetrads are shown from three viewpoints. Only the guanine bases are shown in these structures. The structure on the left is shown from the same viewpoint as used in Figure 1. The structure in the middle is that obtained by rotation of the structure about the indicated z axis by 45°, and the structure on the right is that obtained by an additional rotation by 45°.

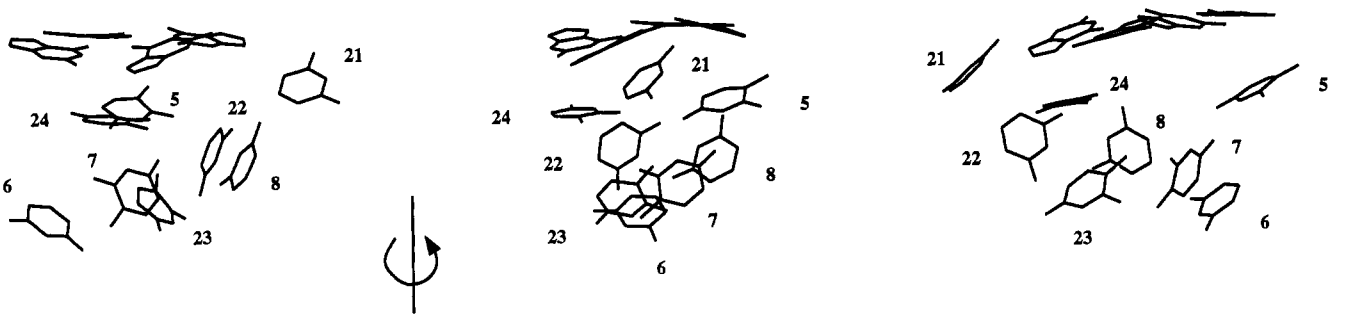

FIGURE 10: The structure of the G tetrad G_4-$\underline{G}_9$-G_{20}-$\underline{G}_{25}$ and the loops T_5-U_6-$\underline{T}_7$-U_8 and U_{21}-U_{22}-$\underline{T}_{23}$-T_{24} are shown from three viewpoints. Only the bases are shown in these structures. The structure on the left is shown from the same viewpoint as used in Figure 1. The structure in the middle is that obtained by rotation of the structure about the indicated z axis by 45°, and the structure on the right is that obtained by an additional rotation by 45°.

T_5-$\underline{T}_7$ base pair is partially stacked on the bottom G tetrad. The other pyrimidine–pyrimidine base pair is between U_{22} and U_8. U_6 is the only residue without significant interactions with other residues in this region. This structure is consistent with the observed NOEs between $\underline{T}_7$ and $\underline{T}_{23}$. Cech and coworkers have shown that $\underline{T}_7$ and $\underline{T}_{23}$ can be photo-cross-linked (Williamson et al., 1989), and this photo-cross-linking event is entirely consistent with the structure in Figure 10. The photo-cross-linking evidence was not incorporated into the model building process but was considered to be an important piece of information.

Insofar as we know, there have been no prior proposals for base pairing within a 5′-anti–anti–syn-3′ sequence. This type of base pairing cannot be accommodated by C or A residues, in their normal protonation states, and may be a reason why telomere DNAs tend to have only T residues between the G repeats. The U22-U8 base pair cannot be formed with neutral A or C residues.

Thus, the structures of the two bottom loops consist of partial stacking on the bottom G tetrad, intra–inter-loop base pairing and $\underline{T}_7$ and $\underline{T}_{23}$ being syn to accommodate the maximal stacking and base pairing. It is worthwhile to point out that while the two loops have symmetric sequences, their structures are quite distinct.

The U_{13}-T_{14}-U_{15}-T_{16} loop is somewhat less well defined in the structure since there were relatively few interresidue NOEs observed from these residues. NOEs from the sugar, but not the base, protons of these residues to the top G tetrad were observed. This indicates that the bases of this loop are not located over the top G tetrad and that the bases are oriented away from the G tetrad. The important NOEs in the model building were those between T_{14} and U_{15} and those between U_{15} and $\underline{G}_{17}$. These NOEs limited the number of possible conformations for the top loop.

The structure of the whole telomere DNA is shown in Figure 11. The four views can be considered to be the front, left, back, and right views of the molecule. The front view illustrates the highly compact nature of the structure as well as the relative widths of the wide and narrow grooves. The back view also shows these features as well as how residue 6, at the very bottom, is exposed as well as the nonplanarity of the G tetrads. The two side views show that the narrow groove on the $\underline{1}$-$\underline{2}$-$\underline{3}$-4:$\underline{9}$-10-$\underline{11}$-12 side is somewhat more regular than that on the $\underline{17}$-18-$\underline{19}$-20:$\underline{25}$-26-$\underline{27}$-28 side. This feature may be a result of the G tetrads not stacking directly on top of one another. Thus, at least one of the narrow grooves will need to be stretched to accommodate the cumulative offset of the stacking of the G tetrads.

These views of the structure of the telomere DNA also illustrate some of the differences between this structure and the crystal structure of Kang et al. (1992) for a dimer of $G_4T_4G_4$ that also contained four G tetrads. The crystal structure is considerably more regular and symmetric than is the structure presented here. The two loops in the crystal structure of the dimer are spatially distant and do not have direct interactions. The structure presented here contains four G tetrads connected by three loops, and the loop–loop interactions may contribute to making the structure less regular and symmetric.

The biological importance of the structural motif determined here is not known at the present time. Fang and Cech (1993a,b) have recently shown that the β-subunit of the *Oxytricha* telomere binding protein catalyzes G-tetrad formation, and they have taken this as evidence for the biological

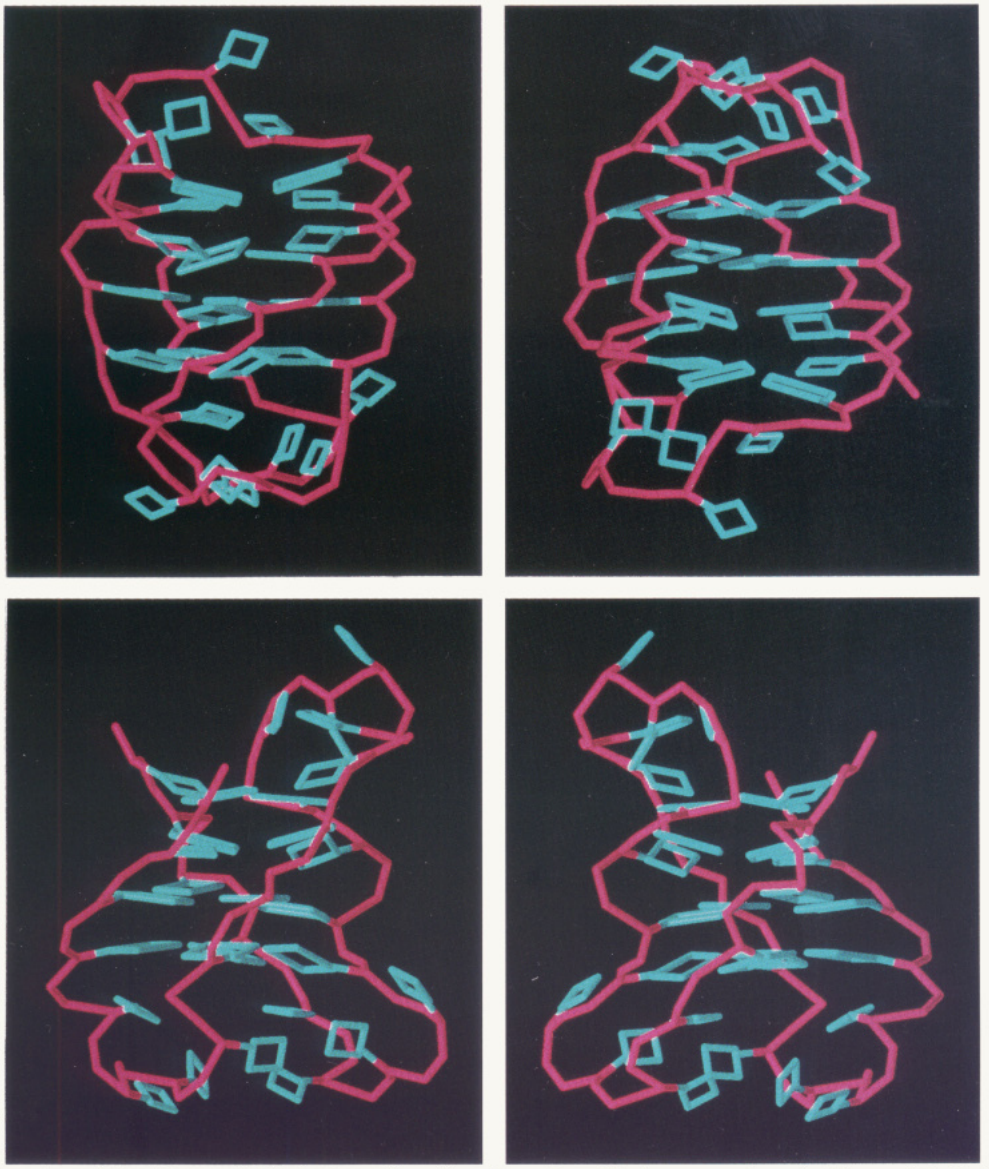

FIGURE 11: The structure of the telomere DNA is shown from four viewpoints. The bases are shown in blue and the backbone in pink. At the top left the structure is shown from the same viewpoint as in Figure 1. At the top right the structure has been rotated 180° about the *z* axis; at the bottom left the structure has been rotated 90° about the *z* axis and in the bottom right by 270° about the *z* axis. The viewpoint in the top right is the same as the unrotated structures in Figures 8–10, and the 90° viewpoint is the same as the 90° viewpoint used in Figures 8–10.

importance of G-tetrad structures. However, it has not escaped our attention that this structure can be adopted by other telomere DNA sequences and that this structure can form a variety of junctions with duplex DNA, and some of these junctions may be important in the biological activity of telomere DNA (K. Y. Wang and P. H. Bolton, in preparation). We are currently investigating the structures and other properties of DNAs that contain tetramer–duplex junctions.

ACKNOWLEDGMENT

The 750-MHz results were obtained with the assistance of Dr. E. Hoffmann of Varian Associates. The 600-MHz spectra were obtained with the assistance of Dr. Frits Abildgaard using the NMR facility at the University of Wisconsin, Madison, which is supported by NIH Grant RR02301 with equipment purchased with support from the University of Wisconsin, NSF DMB-8415048, NIH RR02781, and the USDA. The 500-MHz spectra were obtained with the assistance of Dr. V. V. Krishnamurthy of Gilead Sciences.

REFERENCES

Blackburn, E. H. (1990) Telomeres: Structure and synthesis, *J. Biol. Chem. 265*, 5919–5921.

Blackburn, E. H. (1991) Structure and function of telomeres, *Nature 350*, 569–573.

Blackburn, E. H. (1992) Telomerases, *Annu. Rev. Biochem. 61*, 113–129.

Bock, L. C., Griffin, L. C., Latham, J. A., Vermaas, E. H., & Toole, J. J. (1992) Selection of single-stranded DNA molecules that bind and inhibit thrombin, *Nature 355*, 564–566.

Cheong, C., & Moore, P. B. (1992) Solution structure of an unusually stable RNA tetraplex containing G- and U-quartet structures, *Biochemistry 31*, 8406–8414.

Clore, G. M., & Gronenborn, A. M. (1989) Determination of 3-dimensional structures of proteins and nucleic acids in solution by nuclear magnetic resonance spectroscopy, *CRC Crit. Rev. Biochem. Mol. Biol. 24*, 479–564.

Conrad, M. N., Wright, J. H., Wolf, A. J., & Zakian, V. A. (1990) RAP1 protein interacts with yeast telomeres in vivo: Overproduction alters telomere structure and decreases chromosome stability, *Cell 63*, 739–750.

de Lange, T., Shiue, L., Myers, R. M., Cox, D. R., Naylor, S. L., Killery, A. M., & Varmus, H. E. (1990) Structure and viability of human chromosome ends, *Mol. Cell. Biol. 10*, 518–527.

Ellington, A. D., & Szostak, J. W. (1990) In vitro selection of RNA molecules that bind specific ligands, *Nature 346*, 818–822.

Ernst, R. R., Bodenhausen, G., & Wokaun, A. (1987) *Principles of NMR in One- and Two-Dimensions*, Oxford Univeristy Press, Oxford.

Fang, G., & Cech, T. R. (1993a) The β subunit of *Oxytricha* telomere-binding protein promotes G-quartet formation by telomeric DNA, *Cell 74*, 875–885.

Fang, G., & Cech, T. R. (1993b) Characterization of a G-quartet formation reaction promoted by the β-subunit of the *Oxytricha* telomere-binding protein, *Biochemistry 32*, 11646–11657.

Gilson, E., Roberge, M., Giraldo, R., Rhodes, D., & Gasser, S. M. (1993) Distortion of DNA double helix by RAP1 at silencers and multiple telomeric binding sites, *J. Mol. Biol. 231*, 293–310.

Gray, J. T., Celander, D. W., Price, C. M., & Cech, T. R. (1991) Cloning and expression of genes for the Oxytricha telomere binding protein: specific subunit interactions in the telomeric complex, *Cell 67*, 807–814.

Griffin, L. C., Tidmarsh, G. F., Bock, L. C., Toole, J. J., & Leung, L. L. K. (1993) In vivo anticoagulant properties of a novel nucleotide based thrombin inhibitor and demonstration of regional anticoagulant extracorporeal circuits, *Blood 81*, 3271–3276.

Guo, Q., Lu, M., & Kallenbach (1992) Adenine affects the structure and stability of telomeric sequences, *J. Biol. Chem. 267*, 15293–15300.

Guo, Q., Lu, M., & Kallenbach (1993) Effect of thymine tract length on the structure and stability of model telomeric sequences, *Biochemistry 32*, 3596–3603.

Guschlbauer, W., Chantot J.-F., & Thiele, D. (1990) Four-stranded nucleic acid structures 25 years later: From guanosine gels to telomere DNA, *J. Biomol. Struct. Dyn. 3*, 491–511.

Hardin, C. C., Henderson, E., Watson, T., & Prosser, J. K. (1991) Monovalent cation induced structural transitions in telomeric DNAs: G-DNA folding intermediates, *Biochemistry 30*, 4460–4472.

Hardin, C. C., Watson, T., Corregan, M., & Bailey, C. (1992) Cation dependent transition between the quadraplex and Watson–Crick hairpin forms of d(CGCG$_3$GCG), *Biochemistry 31*, 833–841.

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

Hicke, B. J., Celander, D. W., MacDonald, G. H., Price, C. M., & Cech, T. R. (1990) Two versions of the gene encoding the 41-kilodalton subunit of the telomere binding protein of *Oxytricha* nova, *Proc. Natl. Acad. Sci. U.S.A. 87*, 1481–1485.

Kang, C., Zhang, X., Ratliff, R., Moyzis R., & Rich, A. (1992) Crystal structure of four-stranded *Oxytricha* telomeric DNA, *Nature 356*, 126–131.

Levy, M. Z., Allsop, R. C., Futcher, A. B., Greider, C. W., & Harrley, C. B. (1992) Telomere end replication problem and cell aging, *J. Mol. Biol. 225*, 951–961.

Macaya, R. F., Scultze, P., Smith, F. W., Roe, J. A., & Feigon, J. (1993) Thrombin binding DNA aptamer forms a unimolecular quadraplex structure in solution, *Proc. Natl. Acad. Sci. U.S.A. 90*, 3745–3749.

McClintock, B. (1939) The behavior in successive nuclear division of a chromosome broken at meiosis, *Proc. Natl. Acad. Sci. U.S.A. 25*, 405–416.

McClintock, B. (1942) The fusion of broken ends of chromosomes following nuclear fusion, *Proc. Natl. Acad. Sci. U.S.A. 28*, 458–463.

Price, C. M. (1990) Telomere structure in *Euplotees crassus*: Characterization of DNA–protein interactions and isolation of a telomere binding protein, *Mol. Cell. Biol. 10*, 3421–3431.

Raghuraman, M. K., & Cech, T. R. (1990) Effect of monovalent cation induced telomeric DNA structure on the binding of *Oxytricha* telomeric protein, *Nucleic Acids Res. 18*, 4543–4552.

Romero, D. P., & Blackburn, E. H. (1991) A conserved secondary structure for telomerase RNA, *Cell 67*, 343–353.

Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.

Scaria, P. V., Shira, S. J., & Shafer, R. H. (1992) Quadraplex structure of d(G$_3$T$_4$G$_3$) stabilized by K^+ or Na^+ is an asymmetric hairpin dimer, *Proc. Natl. Acad. Sci. U.S.A. 89*, 10336–10340.

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

Sen, D., & Gilbert, W. (1990) A sodium–potassium switch in the formation of four-stranded G4-DNA, *Nature 344*, 410–414.

Smallcombe, S. (1993) Solvent suppression with symmetrically-shifted pulses, *J. Am. Chem. Soc. 115*, 4776–4785.

Smith, F. W., & Feigon, J. (1992) Quadraplex structure of *Oxytricha* telomeric DNA oligonucleotides, *Nature 356*, 164–168.

Smith, F. W., & Feigon, J. (1993) Strand orientation in the DNA quadraplex formed from the *Oxytricha* telomere repeat oligonulceotide d(G$_4$T$_4$G$_4$) in solution, *Biochemistry 32*, 8682–8692.

Sundquist, W. I., & Klug, A. (1989) Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops, *Nature 342*, 825–829.

Van de Ven, F. J. M., & Hilbers, C. W. (1988) Nucleic acids and magnetic resonance, *Eur. J. Biochem. 178*, 1–38.

Voloshin, O. N., Vaselkov, A. G., Belotserkovskii, B. P., Danilevskaya, O. N., Pavolova, M. N., Dobrynin, V. N., & Frank-Kamenetskii, M. D. (1992) An eclectic DNA structure adopted by human telomeric sequence under superhelical stress and low pH, *J. Biomol. Struct. Dyn. 9*, 643–652.

Wang, K. Y., McCurdy, S., Shea, R. G., Swaminathan, S., & Bolton, P. H. (1993a) A DNA aptamer which binds to and inhibits thrombin exhibits a new structural motif for DNA, *Biochemistry 32*, 1899–1904.

Wang, K. Y., Krawczyk, S. H., Bischofberger, N., Swaminathan, S., & Bolton, P. H. (1993b) The tertiary structure of a DNA aptamer which binds to and inhibits thrombin determines activity, *Biochemistry 32*, 11285–11295.

Wang, Y., & Patel, D. J. (1992) Guanine residues in d(T$_2$AG$_3$) and d(T$_2$G$_4$) form parallel-stranded potassium cation stabilized G-quadraplexes with antiglycosidic torsion angles in solution, *Biochemistry 31*, 8112–8119.

Wang, Y., Jin, R., Gaffney, B., Jones, R. A., & Breslauer, K. J. (1991) Characterization by ^{1}H NMR of glycosidic conformations of the tetramolecular complex formed by d(GGTTTTGG), *Nucleic Acids Res. 19*, 4619–4622.

Williamson, J. R., Raghuraman, & Cech, T. R. (1989) Monovalent cation induced structure of telomeric DNA: the G quartet model, *Cell 59*, 871–880.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.

Zahler, A. M., Williamson, J. R., Cech, T. R., & Prescott, D. M. (1991) Inhibition of telomereases by G-quartet DNA structures, *Nature 350*, 718–720.

Zakian, V. A. (1989) Structure and function of telomeres, *Annu. Rev. Genet. 23*, 579–604.