Subramanian, S. (1982) *Arch. Biochem. Biophys. 216*, 116-125.
Subramanian, S., & Kaufman, B. T. (1980) *J. Biol. Chem. 255*, 10587-10590.
Thompson, S. T., & Stellwagen, E. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 361-365.
Thompson, S. T., Cass, K. H., & Stellwagen, E. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 669-672.
Wierenga, R. K., De Maeyer, M. C. H., & Hol, W. G. J. (1985) *Biochemistry 24*, 1346-1357.
Ysern, X., & Prochaska, H. J. (1989) *J. Biol. Chem. 264*, 7765-7767.

Cation-Dependent Transition between the Quadruplex and Watson-Crick Hairpin Forms of d($CGCG_3GCG$)†

Charles C. Hardin,* Thomas Watson, Matthew Corregan, and Charles Bailey
Department of Biochemistry, North Carolina State University, Raleigh, North Carolina 27695
Received June 4, 1991; Revised Manuscript Received October 10, 1991

ABSTRACT: The DNA oligonucleotide d($CGCG_3GCG$) can form either a Watson-Crick (WC) hairpin or a parallel-stranded quadruplex structure containing six G-quartet base pair assemblies. The exchange between these forms and single strands can be monitored using circular dichroism (CD). NMR results verified the assignment of specific CD bands to quadruplex and hairpin species, respectively. Cations stabilize the quadruplex in the order $K^+ > Ca^{2+} > Na^+ > Mg^{2+} > Li^+$ and $K^+ > Rb^+ > Cs^+$, indicating that K^+ has an optimum ionic radius for complex formation and that ionic charge affects the extent of ion-induced stabilization. The quadruplex is stable in the presence of 40 mM K^+ at micromolar DNA concentration and can be kinetically trapped as a metastable form when prepared at millimolar DNA concentration and then diluted into buffer containing 40 mM Na^+. The concentration of K^+ required to reverse the equilibrium from the hairpin to the quadruplex decreases sharply with increased DNA concentration. The quadruplex has an unusual pK_a of ca. 6.8, indicating that $C{\cdot}C^+$ base pairs are probably forming. This system provides insights into some of the detailed structural characteristics of a ["G4-DNA"·ion] complex and an experimental model for the recently proposed "sodium-potassium conformational switch" [Sen, D., & Gilbert, W. (1988) *Nature 334*, 364-366; Sen, D., & Gilbert, W. (1990) *Nature 344*, 410-414]. These results may help to explain the lack of cytidine residues in G-rich telomeric DNAs and suggest that methylation of GC-rich duplex DNAs in "GpC islands" may induce quadruplex formation within heterochromatin domains, resulting in reversible chromosomal condensation.

Telomeres occur at the ends of linear eukaryotic chromosomes and are necessary for faithful genomic heritability (Blackburn, 1986, 1991). Telomeric DNAs consist of repetitive sequences that are unique for a given organism. A G-rich 3′-terminal strand extends an estimated 12-16 nucleotides past the end of the complementary C-rich strand in the telomeric DNAs of organisms that have been analyzed (Blackburn, 1986; Blackburn & Szostak, 1984; Henderson & Blackburn, 1989). Repetitive G-rich telomeric sequences that are composed of G's and T's, G's and A's, or all three nucleotides are known to exist. Specific examples include T_2G_4 (*Tetrahymena*), T_4G_4 (*Oxytricha*), $TGTGTG_3$ (*Saccharomyces*), $AG_7AGAG_6AG_6$ (*Dictyostelium*), T_2AG_3 (*Homo sapiens*), and T_3AG_3 (*Arabidopsis*) (Forney et al., 1987; Richards & Ausubel, 1988; Roberts, 1988). Curiously, C residues occur very infrequently in the G-rich strands of telomeric DNAs (Forney et al., 1987; Richards & Ausubel, 1988; Roberts, 1988).

Electrophoretic chemical-probing experiments and spectroscopic studies have shown that telomeric DNAs can form a number of unusual structures that apparently contain G·G base pairs (Sundquist & Klug, 1989; Oka & Thomas, 1987; Williamson et al., 1989; Sen & Gilbert, 1990; Raghuraman & Cech, 1990; Hardin et al., 1991). The term "G-DNA" was coined for this class of structures (Cech, 1988). G-DNA structures can be classified into two general categories (Sen & Gilbert, 1990; Hardin et al., 1991), duplexes ("G2-DNA") and quadruplexes ("G4-DNA"). Duplex forms include both intramolecular hairpins and bimolecular helices, while quadruplexes can conceivably be constructed from 1, 2, 3, or 4 strands (Oka & Thomas, 1987; Sundquist & Klug, 1989; Sen & Gilbert, 1990; Hardin et al., 1991). Nondenaturing electrophoresis, NMR, and CD[1] results showed that d$(T_2G_4)_4$ can form different equilibrium mixtures of apparent 1-, 2-, or 4-stranded quadruplex species, depending upon the solution conditions (Hardin et al., 1991). Three-stranded G-DNA complexes could conceptually exist as an intramolecular triplex annealed (via G·G base pairing on both "sides") to a second strand. Strand polarities can be classified as parallel or antiparallel. Hairpin-containing complexes are constrained to having antiparallel strands. Strand polarity and nucleoside glycosidic torsion angles appear to be correlated in G-DNA quadruplexes. Syn and anti glycosidic torsion angles have been observed by NMR in the hairpin complex formed by d$(T_2G_4)_4$

† These studies were supported by grants from the North Carolina Research and Development Fund (No. 9223) and the National Institutes of Health (BRSG RR9223). Partial funding was also provided by the North Carolina Agricultural Research Service.
* To whom correspondence should be addressed.

[1] Abbreviations: CD, circular dichroism; DMS, dimethyl sulfate; KP, 20 mM potassium phosphate (pH 7) containing 0.1 mM EDTA; NaP, 20 mM sodium phosphate (pH 7) containing 0.1 mM EDTA; TSP, sodium 3-(trimethylsilyl)-1-propanesulfonate; Tris, tris(hydroxymethyl)aminomethane; WC, Watson-Crick.

0006-2960/92/0431-833$03.00/0

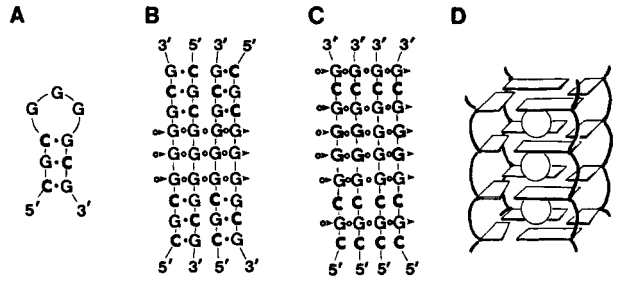

FIGURE 1: Some possible structures formed by d(CGCG$_3$GCG). (A) Watson–Crick hairpin, (B) antiparallel-stranded quadruplex, and (C) parallel-stranded quadruplex. Symbols: WC base pairs (●); G·G base pairs (○). (D) Three-dimensional structure of a "generic" cation/quadruplex complex. The parallel-stranded quadruplex would represent the four adjacent sets of G·G base paired residues shown in panel C.

(Forney et al., 1987; Hardin et al., 1991) and the antiparallel four-stranded complex formed by d($G_2T_5G_2$) (Jin et al., 1990; Wang et al., 1991). In contrast, all of the nucleotides in the parallel-stranded fiber structure formed by poly[r(G)] adopt the anti glycosidic conformation (Saenger, 1984). In addition, the quadruplex has a right-handed helical sense, and the conformational features are similar to those of A-family structures (Guschlbauer et al., 1989; Saenger, 1984).

Interconversions between the G-DNA structures formed by d(T_2G_4)$_4$ are very dependent on the concentrations of specific cations (Oka & Thomas, 1987; Williamson et al., 1989; Sen & Gilbert, 1990; Raghuraman & Cech, 1990; Hardin et al., 1991). The differential effects of ions arises from the fact that two "G-quartet" base-pairing assemblies fold around the ion to form a sandwich-like "cryptand" complex (Figure 1D). The stability of the resulting quadruplex depends on both ionic radius and charge (Sundquist & Klug, 1989; Oka & Thomas, 1987; Williamson et al., 1989; Hardin et al., 1991; Zimmerman et al., 1975; Howard & Miles, 1982). The K^+ ion is especially effective in trapping "by-product" species (Sen & Gilbert, 1990) and inducing transitions to exceptionally stable forms (Sen & Gilbert, 1990; Hardin et al., 1991).

A clear understanding of the role of this phenomenon in telomere function has yet to emerge. However, these interconversions could be biologically relevent since they readily occur under physiologically compatible conditions of temperature, pH, and ionic strength and in the presence of the predominant intracellular cations (K^+, Na^+, Ca^{2+}, Mg^{2+}) (Hardin et al., 1991; Zahler et al., 1991). It is especially interesting that under some circumstances a K^+-stabilized form of G-DNA is preferred over the corresponding Watson–Crick DNA structure (Sen & Gilbert, 1988; Hardin et al., 1991; Smith et al., 1989). Nontelomeric G-rich sequences can also form G-DNA structures. Examples include oligonucleotides that contain consensus sequences derived from immunoglobulin switching regions (Sen & Gilbert, 1988), a sequence from the human c-Ha-*ras* gene (Smith et al., 1989) and several sequences of the general form $G_2N_5G_2$ (Jin et al., 1990; Wang et al., 1991). Thus, it has been proposed that G-DNA may participate in functional roles associated with nontelomeric DNAs including alignment of chromosomes in preparation for recombination during meiotic pairing (Sundquist & Klug, 1989), modulation of DNA methylation (Smith et al., 1989), and the control of gene expression (Boeke, 1990; Hardin et al., 1991).

Materials and Methods

Preparation of DNAs. DNA oligonucleotides were synthesized and removed from the support as previously described (Hardin et al., 1991). The oligonucleotides were purified by preparative gel electrophoresis on 20% polyacrylamide gels essentially as previously described. This method was used to purify the oligonucleotides rather than HPLC because they formed mixtures of species, resulting in complicated elution profiles from DEAE HPLC columns using acetonitrile gradients in triethylammonium-HCl buffer at pH 6. The d(CGCG$_n$GCG) (n = 2–4) series molecules migrate predominantly as monomers on 20% polyacrylamide gels containing 8 M urea at pH 8.3. Following electrophoresis, samples were dialyzed extensively against 0.1 mM Li_2EDTA (pH 8) in a BRL 1200MA microdialysis apparatus at ca. 1 mM strand concentration. The sample was stable as a ca. 1:1 mixture of the quadruplex and hairpin species when stored at 4 °C at a strand concentration of ca. 1 mM.

Circular Dichroism Spectroscopy. Circular dichroism spectra were collected on a Jasco J-600 spectropolarimeter interfaced to an IBM PC microcomputer. The sample temperature was maintained by placing the sample in a 1-cm path length cylindrical cell surrounded by an external jacket which was connected to a Neslab recirculating water bath. All CD data were baseline corrected for signals due to the cell and buffer.

NMR Spectroscopy. In order to convert the sample to the hairpin conformation, an 80 A_{260} unit sample was diluted into 20 mL of NaP buffer, heated to 70 °C for 20 min, cooled quickly in an ice bath, lyophilized to dryness, and resuspended into 200 μL of H_2O. The sample was then dialyzed extensively against NaP buffer in a microdialysis apparatus. In order to maintain the quadruplex structure, an 80 A_{260} unit sample was suspended into 200 μL of KP buffer and dialyzed extensively against KP. Both sample volumes were then adjusted to 350 μL. NMR data were collected at 500 MHz on a GE Omega spectrometer using the "1–1 hard pulse" solvent supression method (Otting et al., 1987); spectra contained 16K data points consisting of 256 scans each. Partially deuterated TSP was used as the internal ^1H chemical shift reference.

Results

Differential Effects of Ions on Structures Formed by d(CGCG$_3$GCG). The oligonucleotide d(CGCG$_3$GCG) can potentially form the structures shown in Figure 1. The intramolecular "hairpin" structure (Figure 1A) has a stem containing three C·G base pairs and a loop composed of three guanine residues. The structure shown in Figure 1B represents a four-stranded "quadruplex" containing three guanine "quartets" (Sundquist & Klug, 1989; Figure 1D) and four terminal duplex stems composed of three C·G base pairs. Both of these structures contain antiparallel strands. The second structure consists of two antiparallel duplexes that contain six C·G base pairs and three internal G·G base pairs. The duplex is not considered as a separate structure here since it is likely that it would dimerize in the presence of cations to form the quadruplex. A second type of quadruplex that contains parallel strands and the maximum number of G-quartets (i.e., six) is shown in Figure 1C. The structure has four contiguous G-quartets flanked on each end by cytidine residues and then guanine quartets; the 5′ ends are capped by a third set of cytidine residues. An ion/quadruplex complex containing four G-quartets is shown in Figure 1D to illustrate the nature of the three-dimensional structure represented by the hydrogen-bonding scheme shown in Figure 1C. This "structure" is generic in the sense that the strand polarities and other conformational details are not specified.

CD spectra were obtained in Na^+- and K^+-containing phosphate buffers at pH 7 in the 23–80 °C range to determine if these ions induced differential interconversions between any

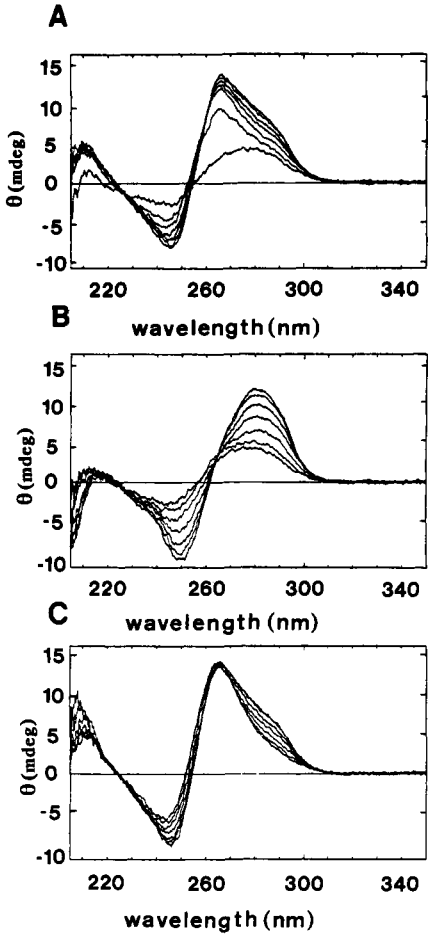

FIGURE 2: Effects of sodium and potassium on the thermal denaturation of d(CGCG$_3$GCG) monitored by CD spectroscopy. DNA (7.5 μM) in NaP (A and B) or KP (panel C) buffer at pH 7 containing 0.1 mM EDTA. (A) Effect of increased temperature on a freshly prepared sample in NaP. Spectra were obtained at 23, 30, 40, 50, 60, 70, and 80 °C; the ellipticity values in the 264–286-nm range decreased as the sample temperatures were raised. (B) The sample used to obtain the results shown in panel A was allowed to cool to 4 °C, incubated overnight at that temperature, and then subjected to a second thermal denaturation. (C) CD spectra of a freshly prepared sample obtained in KP. Identical results were obtained when the sample was cooled, incubated overnight at 4 °C, and then denatured a second time. The CD bands at 264 and 286 nm were assigned to the quadruplex and WC hairpin species, respectively (see text).

of these structures (Figure 2). The intensities of the CD bands at 264 and 286 nm decreased dramatically with increased temperature in the presence of Na^+ (Figure 2A). Upon raising the temperature of the sodium phosphate (NaP) sample to 80 °C and then allowing it to cool, a large CD band appeared at 286 nm, and the band at 264 nm essentially disappeared (Figure 2B). The sample was heat-denatured a second time to determine whether the newly obtained structure was thermodynamically stable. A nearly identical spectrum was obtained upon cooling, verifying that the same structure reformed and the corresponding equilibrium was reversible in NaP buffer.

There was little change in the intensity of the CD band at 264 nm as the temperature was increased in potassium phosphate (KP) buffer (Figure 2C). However, the intensity of the band at 286 nm decreased to less than half the initial intensity. Comparison with the results shown in Figure 2A demonstrates that the structure corresponding to the 264-nm band is much more stable in the presence of K^+ than in Na^+. In contrast, the structure corresponding to the 286-nm band has essentially the same stability in the presence of both cations. Essentially identical results were obtained when samples were thermally denatured in Tris-HCl buffer at pH 7 containing NaCl or KCl (Figure 3). This demonstrates that the effects produced by Na^+ and K^+ are essentially independent of the buffer (phosphate versus Tris-HCl). Previous studies have shown that quadruplex structures are much more stable in the presence of K^+ than in Na^+ (Sen & Gilbert, 1990; Raghuraman & Cech, 1990; Hardin et al., 1991). On the basis of the differential ionic effects shown here, it was concluded that the peak at 264 nm corresponds to the quadruplex complex and the peak at 286 nm represents the WC hairpin conformation. Given these assignments, these results show that (i) the quadruplex can be kinetically trapped as a metastable form when prepared at millimolar DNA concentration and then diluted ca. 100-fold into buffer containing 40 mM Na^+ and (ii) the WC hairpin is not converted back to the quadruplex structure at a DNA concentration of ca. 7.5 μM in the presence of 40 mM Na^+ (i.e., the equilibrium is poised far to the right).

Ion-Dependent Stabilization of the Quadruplex Structure. In order to assess the effects of cations on the relative stabilities of the hairpin and quadruplex structures more thoroughly, CD thermal denaturation analyses were done at a DNA concentration of ca. 7.5 μM in Tris-HCl buffers containing either LiCl, NaCl, KCl, RbCl, CsCl, $MgCl_2$, or $CaCl_2$. Plots of the ellipticity values at 286 nm and 264 nm are shown in Figure 3. The relatively broad transition profiles obtained at 286 nm (Figure 3A) have similar slopes and ranges, although the Li^+ curve diverges from the other curves at lower temperature and the ellipticities obtained with the divalent ions are lower than seen with the monovalent ions (see below). All of the curves span a large temperature range, indicating that the hairpin dissociation reaction is only moderately cooperative. While the high- and low-temperature baselines are only crudely defined, the curves are clearly sigmoidal in most cases and show an apparent transition temperature in the 50° range. In contrast, the transition temperatures obtained with the different cations for the species corresponding to the 264-nm CD band ranged from 32 to $\gg$80 °C (Figure 3B). Since the entire transition profile could not be obtained in the case of most of the ions, the relative stabilities of the putative quadruplex complexes were estimated by assuming that the ellipticities of the denatured complexes were the same. Given this assumption, the temperature corresponding to 10% denaturation of the complex could be estimated for all of the cases except K^+. The temperatures that correspond to a 10% denatured quadruplex population are plotted as a function of the ionic radius for each of the hydrated ions (Dobler, 1981) in Figure 3C. This temperature (abbreviated as $T_{0.1m}$) is used because it allows a semiquantitative comparison of the relative stabilities of the complexes even though the true T_m ($T_{0.5m}$ in our nomenclature) values could not be measured for the more stable cases. The plot shows how the stability of each ion/quadruplex complex is affected by ionic radius. Results obtained with mono- and divalent cations are plotted separately to show how ionic radius and charge affect the extent of stabilization induced by the ion. The complex formed in the presence of K^+ was the most stable (highest $T_{0.1m}$), demonstrating that of the ions that were tested it has the optimum van der Waals radius for complex formation. A temperature of >90 °C is required to denature this complex. In contrast, the complex formed with the much smaller Li^+ ion denatures by 60 °C! These results show that the monovalent cations stabilize the putative quadruplex in the following order: $K^+ > Na^+ > Li^+$ and $K^+ > Rb^+ > Cs^+$. The order obtained for

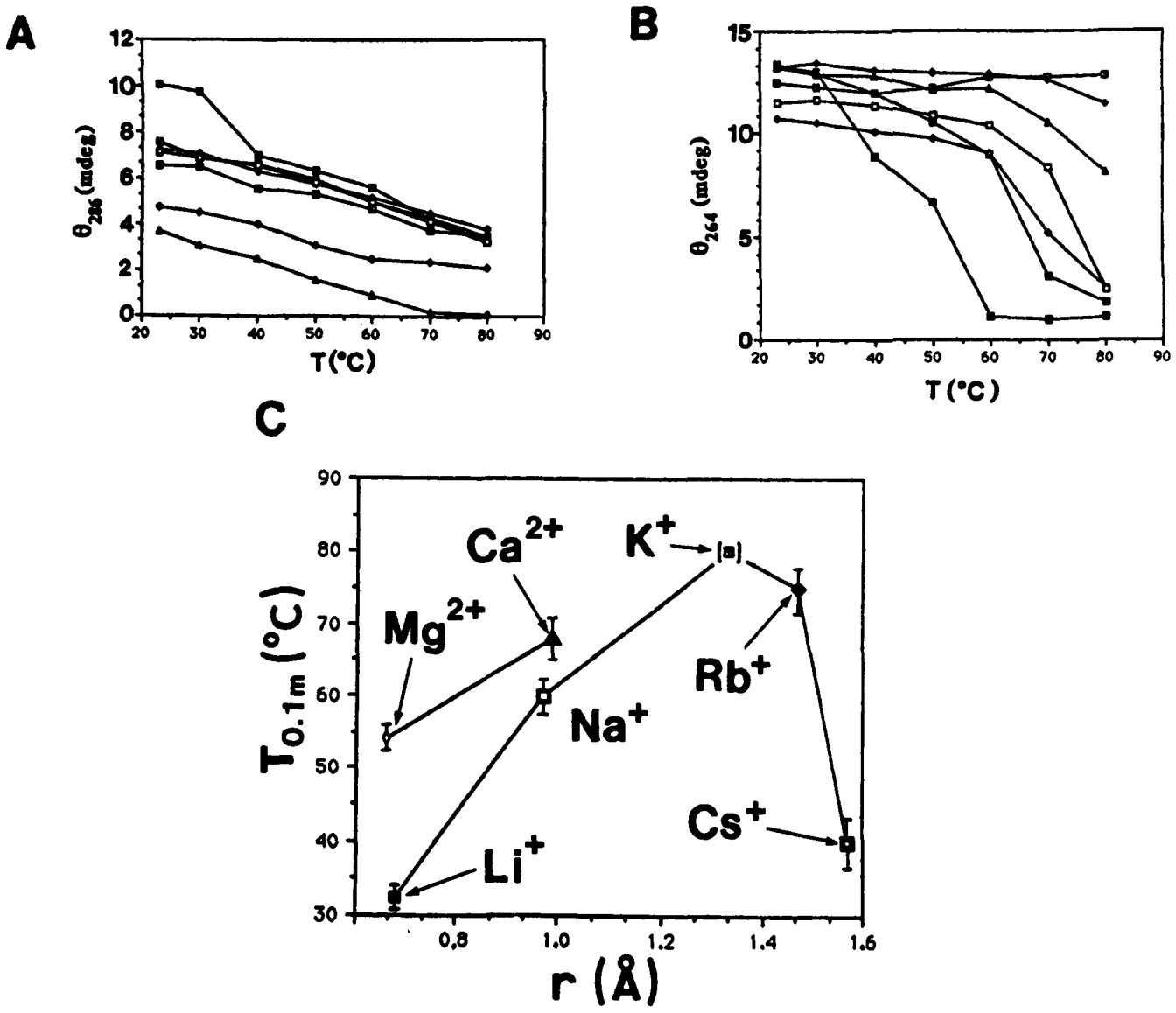

FIGURE 3: Stability of the WC hairpin and quadruplex structures formed by d(CGCG$_3$GCG) in solutions containing different physiologically relevant monoatomic cations monitored by CD. Ellipticities corresponding to either (A) the WC hairpin (286 nm) or (B) the quadruplex (264 nm) are plotted as a function of temperature for the following salts in 10 mM Tris-HCl (pH 7) containing 0.1 mM EDTA: LiCl (■), NaCl (□), KCl (⊡), RbCl (◆), CsCl (◘), $CaCl_2$ (▲), or $MgCl_2$ (◇). (C) Plot of the temperature at which 10% of the quadruplex population has denatured ($T_{0.1m}$) versus the van der Waals radius of the corresponding hydrated ion. Error bars correspond to standard deviations in duplicate determinations.

the divalent cations is $Ca^{2+} > Mg^{2+}$. The difference between the relative stabilities of the Na^+ and K^+ complexes is not as dramatic as seen with d(T$_2$G$_4$)$_4$ (Hardin et al., 1991). This is probably due to the larger number of ion-binding sites in the intermolecular complex(es) formed by d(T$_2$G$_4$)$_4$ (Hardin et al., 1991). Theoretical molecular dynamics/thermodynamic perturbation calculations (W. S. Ross and C. C. Hardin, manuscript in preparation) also yielded a value in the range of the experimentally determined optimum radius for complex formation (ca. 1.3 Å), lending further credence to the conclusion that the complex is a G-DNA quadruplex.

The stabilities obtained with divalent cations are higher than obtained with monovalent ions of comparable radius (Figure 3C), demonstrating that both electrostatic and van der Waals forces affect the ion-dependent stabilization phenomenon. It is interesting that, as the size of the divalent cation increases in going from Mg^{2+} to Ca^{2+}, the differential stabilization relative to the monovalent cation of corresponding size decreases. It is also interesting that the $T_{0.1m}$ values for mono- and divalent cations extrapolate to approximately the same value, i.e., the value obtained with K^+.

Note that the ellipticities at 286 nm, corresponding to the concentrations of the putative hairpin species, decrease in the reverse order of the relative stabilities of the putative quadruplexes (Figure 3C): $Li^+ > Na^+, Rb^+, Cs^+ > K^+$ and $Mg^{2+} > Ca^{2+}$. Since the quadruplex and hairpin conformations denature to form a common structure, the single-stranded species, these results can be attributed to the linked conformational equilibria. Since K^+ stabilizes the quadruplex most effectively, less hairpin should be present. The inverse should be true for Li^+. It is not clear why the θ_{286} values for the divalent species are lower. This result could indicate that the WC hairpins that form in the presence of these ions are less stable than those that form with the monovalent ions. However, this is not consistent with relative "ion exchange equilibrium quotient" values for mono- and divalent cations with duplex DNAs (Paulson et al., 1988). It is more likely that it is due to a decrease in the molar ellipticity of the DNA chromophore (and consequent decrease in CD) in the presence of divalent ions.

A quantitative thermodynamic analysis of these results is not possible when T_m values cannot be accurately determined (e.g., with K^+ and Rb^+). However, van't Hoff enthalpy values were determined for the cases where T_m values could be measured. The analysis was done using the general form of the van't Hoff equation (Chan et al., 1990), which requires information regarding the molecularity of the reactions (i.e., the number of strands in the complex). Values for the all of structures represented by the 286-nm band were calculated assuming either the intramolecular hairpin ($n = 1$) or duplex ($n = 2$) conformation. These values ranged from −6 to −14 (±3–6) kcal/mol per cooperative unit for the hairpin and −9 to −21 kcal/mol for the duplex (standard deviation in duplicate analyses). In contrast, the values ranged from −36 to −110 (±1–11) kcal/mol, assuming that the structure corresponding to the 264-nm band is a quadruplex in the presence of Li^+, Cs^+, Mg^{2+}, Na^+, and Ca^{2+}. These results reflect the large range of denaturation temperatures shown in Figure 3C and demonstrate the large differences in quadruplex stabilization due to the different cations. In contrast, the values calculated for the hairpin or duplex species span relatively narrow ranges.

The Hairpin to Quadruplex Equilibrium Is Reversible. Since the quadruplex is composed of four strands, the equilibrium between the single-stranded species and the putative quadruplex should be strongly dependent on the DNA concentration (Raghuraman & Cech, 1990; Jin et al., 1990; Wang et al., 1991; Sen & Gilbert, 1988). The effect of DNA concentration on the K^+-induced transition was assessed to see if the transition between the hairpin and quadruplex species was reversible. CD spectra were obtained at a ca. 7.5 μM DNA strand concentration in 1-cm cells and at 10- and 100-fold higher DNA concentrations (75 μM and 0.75 mM) using narrow path length CD cells. Concentrated KCl was added to an oligonucleotide sample that had been denatured and then cooled (i.e., converted to the hairpin form). The sample was then allowed to equilibrate at 70 °C for 15 min to overcome any kinetic barriers to the reverse (HP to Q) transition. After the sample was allowed to reequilibrate at 23 °C for 30 min, the CD spectrum was obtained. The ratio of ellipticity values at 264 and 286 nm was used to assay the extent of quadruplex formation. A K^+ concentration of about 630 mM was required to convert the hairpin back to the quadruplex structure at a DNA concentration of ca. 7.5 μM (Figure 4A); a K^+ concentration of ca. 240 mM can induce the conversion at a DNA concentration of 75 μM (Figure 4B); the K^+ concentration required to induce the conversion from a hairpin to the quadruplex at a DNA concentration of ca. 0.75 mM was about 30 mM (Figure 4C). Thus, the concentration of K^+ required to convert the hairpin to the quadruplex decreased dramatically with increased DNA concentration. This result is consistent with the assignment of the 264-nm CD band to a quadruplex structure. Given the limited nature of the results, further quantitative analysis was not pursued.

"Counterion condensation" theory can be used to characterize ion to strand stoichiometries in DNA/counterion systems [see Record et al. (1978)]. However, the analysis is complicated in the present case by at least two factors. Since the hairpin to single-strand equilibrium is coupled to the quadruplex to single-strand equilibrium, it is not a trivial task to determine the equilibrium constants that are necessary to determine the ion to strand stoichiometry. Secondly, it is not clear how the internally bound ions will affect the linear charge density of the molecule. Assuming that no end-to-end associations are present, one can predict the number of complexed

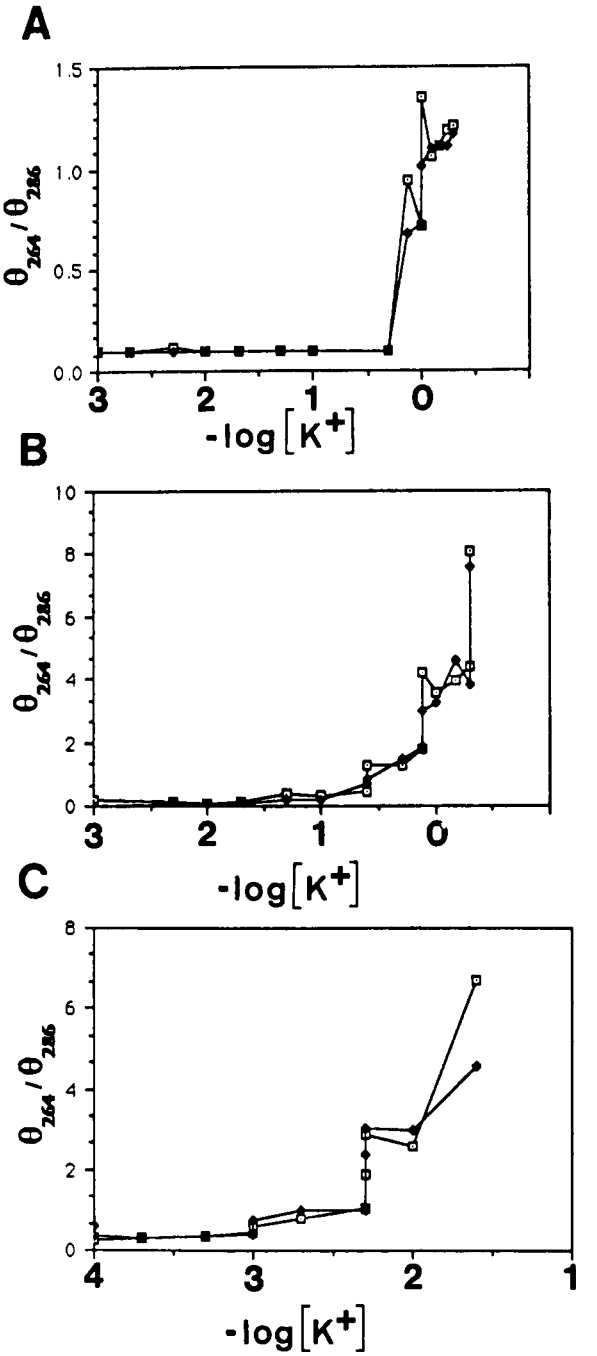

FIGURE 4: Effect of DNA concentration on the potassium-induced conversion from the WC hairpin conformation to the quadruplex form of d(CGCG$_3$GCG). (A) A 7.5 μM DNA sample was incubated in NaP buffer at 23 °C at different added KCl concentrations, and then the CD spectrum was acquired (closed symbols). The temperature was then raised to 70 °C for 10 min, cooled, and reequilibrated at 23 °C, and a second spectrum was obtained (open symbols). Data points correspond to the ratio of measured ellipticity values at 264 nm (quadruplex) and 286 nm (WC hairpin) plotted versus the logarithm of the added KCl concentration. Since the extinction coefficients are of the same order of magnitude, a ratio >1 indicates that the equilibrium has shifted away from the WC hairpin and toward the quadruplex. (B) Same as panel A except that the DNA concentration was increased 10-fold (75 μM) and the spectra were obtained in a 1-mm path length cell. (C) Same as panel A except that the DNA concentration was increased 100-fold (0.75 mM) and the spectra were obtained in a 0.1-mm cell. Ratio values of 1.0 were obtained at approximate K^+ concentrations of (A) 630 mM, (B) 250 mM, and (C) 30 mM.

ions. The four internal stacked G-quartets could presumably bind three ions (Figure 1D), and if the intervening cytidines do not preclude them, at least two more ions could presumably

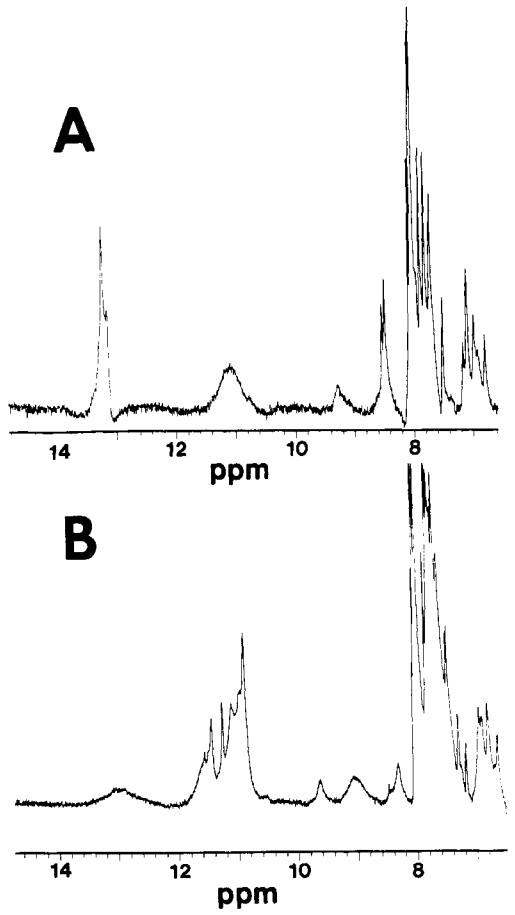

FIGURE 5: Imino ^{1}H NMR spectra of d(CGCG$_3$GCG) in (A) NaP at 5 °C (predominantly WC hairpin) and (B) in KP at 35 °C (predominantly quadruplex). Samples were prepared as described under Materials and Methods.

bind between the second and external sets of G-quartets.

NMR Evidence for WC Hairpin and Parallel-Stranded G-DNA Quadruplex Structures. The results shown in Figure 4C demonstrated that the quadruplex structure was stable in KP at the DNA strand concentrations typically used for NMR analysis (0.5–10 mM) at temperatures up to at least 70 °C. Equivalent DNA samples were dialyzed in parallel into NaP or KP buffer, and the concentrations were adjusted to ca. 1 mM for NMR analysis. CD spectroscopy was used to verify that the NMR samples remained in either the hairpin (in NaP) or quadruplex form (in KP). Aliquots were removed and diluted 2-fold with the appropriate buffer, and then CD spectra were obtained at 23 °C before and after heating to 70 °C. These results showed that a limited amount (<15%) of each type of complex was present in the "other" sample. On the basis of the ellipticity values, the concentrations of the hairpin and quadruplex species did not change appreciably due to the heating/cooling process. This indicated that the complex to single-strand equilibria were relatively stable in each case.

The downfield region of the NMR spectrum in H_2O buffer contains resonances that correspond to the exchangeable imino and amino hydrogen atoms. Spectra for the hairpin and quadruplex forms of d(CGCG$_3$GCG) are shown in Figure 5A,B. Imino proton peaks that integrate to give a ratio of approximately 2:1 are shown in Figure 5A at 13.23 and 13.18 ppm. This is in the chemical shift range for imino protons that are involved in WC C·G base pairs (Arnold et al., 1987; Hare & Reid, 1986; Pardi & Tinoco, 1982). These results add credence to the interpretation that the complex formed in NaP upon melting (corresponding to the 286-nm CD band) is a hairpin complex containing three C·G base pairs.

In contrast, the NMR spectrum obtained in KP (Figure 5B) has a group of six relatively discrete peaks dispersed from 10.96 to ca. 11.64 ppm. Imino proton peaks were assigned to G·G base pairs in the same region of the NMR spectra for the G-DNA hairpin and quadruplex forms of d(T$_2$G$_4$)$_4$ (Henderson et al., 1987; Hardin et al., 1991). These results are consistent with the interpretation that the structure that is formed by d(CGCG$_3$GCG) in KP is a G-DNA quadruplex. The results in Figure 5B provide two pieces of evidence for the somewhat unexpected conclusion that the quadruplex is *parallel-stranded*. First, there are six peaks in in the 10–12 ppm range rather than three as expected for the antiparallel-stranded form. Since the parallel-stranded complex has a 4-fold helical axis of symmetry (Figure 1D), all four imino protons in each G-quartet should be in the same environment. Thus, the chemical shifts would be degenerate, and one would expect six peaks corresponding to 24 protons.

This presents a somewhat confounding situation in trying to solve the structure unambiguously using two-dimensional NMR methods. Due to the (assumed) 4-fold molecular symmetry, one effectively "loses" one dimension of information, and all strand-to-strand "connectivities" appear to be "intramolecular". Thus, the results could also correspond to axially symmetric complexes with other strand stoichiometries. The same problem arose in early attempts to solve the fiber diffraction structures of poly[r(I)$_4$] and poly[r(G)$_4$] complexes (Saenger, 1984; Zimmerman et al., 1975; Howard & Miles, 1982). Two elegant approaches have been used to solve the strand stoichiometry problem in other G-DNA systems. In the electrophoretic demonstration (Sen & Gilbert, 1990), oligonucleotides of two different lengths were mixed and the tetrameric nature of the complexes was confirmed by the expected permutation of products. The thermodynamic proof (Sen & Gilbert, 1988) involved comparing spectrophotometrically determined van't Hoff enthalpy values ($\Delta H°_{vH}$) with results obtained from differential scanning calorimetry measurements ($n\Delta H°_{cal}$) to solve for the strand molecularity (n).

Note that the two downfield resonances are less intense and broadened relative to the others, even when one accounts for overlapping peak intensity (Figure 5B). This would be expected for the imino protons in the 3′- and 5′-terminal G-quartets since "ring-current" effects usually result in upfield shifts for peaks corresponding to stacked imino protons (Arnold et al., 1987; Hare & Reid, 1986; Pardi & Tinoco, 1982). Also, kinetic "fraying" events (exchange between "open" and "closed" forms) typically result in decreased resonance intensities for terminal residues (Giessner-Prettre et al., 1977). Since the farthest downfield "peak" is least intense, it is likely to correspond to the imino protons in the less constrained (see next section) 3′-terminal G-quartet. Therefore, the other downfield "peak" probably correspond to imino protons in the more constrained 5′-terminal G-quartet.

There are small broadened envelopes but no discrete resonances in the 13 ppm range of the NMR spectrum in KP buffer at temperatures ranging from 10 to 80 °C (e.g., at 35 °C, Figure 5B). This indicates that terminal WC base pairs do not occur in near-equimolar amounts in the quadruplex complex under conditions in which the six G·G base-pairing interactions are relatively stable. This result is expected for the parallel-stranded complex and is not consistent with the antiparallel structure assignment (Figure 1B). Broad ill-defined peak envelopes are present in the chemical shift range of the WC C·G resonance (~13 ppm) in the KP sample (Figure 5B) and in the range of the G·G peaks (~11 ppm)

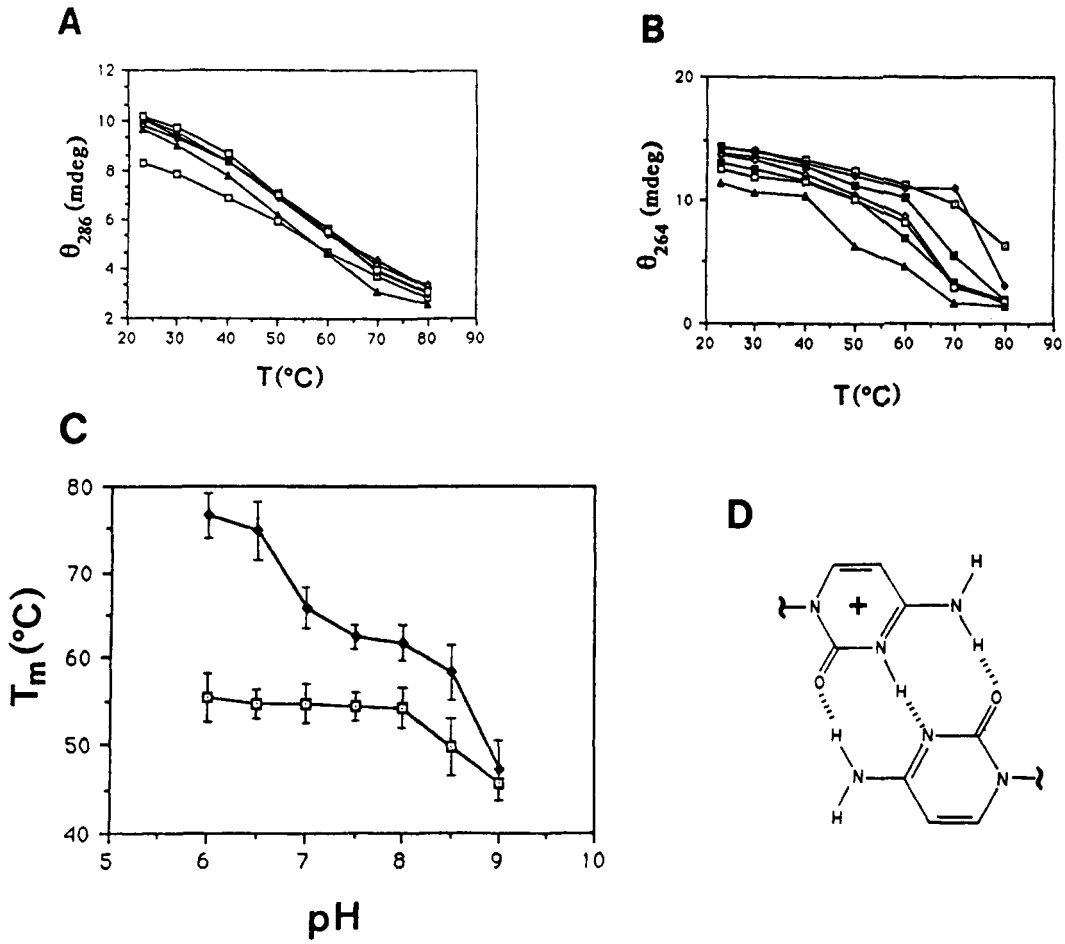

FIGURE 6: Effect of pH on the structures formed by d(CGCG$_3$GCG) in NaP buffer monitored by CD. Ellipticities at (A) 286 nm (WC hairpin) and (B) 264 nm (quadruplex) obtained as a function of temperature corresponding to sample pH values ranging from 6 to 9 are plotted from right to left (i.e., decreasing stability). Values obtained at the given pH correspond to the following symbols: pH 6, ⊡; pH 6.5, ◆; pH 7, ■; pH 7.5, ◇; pH 8, □; pH 8.5, ■; pH 9, ▲. (C) Temperatures corresponding to the average of the maximum and minimum ellipticities (T_m) plotted as a function of the pH value for the 286-nm (WC hairpin, ⊡) and 264-nm (quadruplex, ◆) CD bands, respectively. Error bars correspond to standard deviations in duplicate determinations. (D) Base pair formed between the protonated and neutral forms of cytidine. The population of quadruplexes containing the N–H···N hydrogen bond, which is mediated by the imino and protonated imino nitrogens, exceeds 50% when the pH is greater than the pK_a for the process of protonating N3.

in the NaP sample spectrum (Figure 5A). These results are in agreement with the CD data which indicated that a small population of the "other" type of complex was present in each of the NMR samples due to the linked equilibria (linked via the denatured single-stranded intermediate). After these results were obtained, the samples were combined, lyophilized, resuspended into 200 μL of H_2O, dialyzed into KP, and the volume was adjusted to 350 μL. The imino ^{1}H NMR spectrum of the resulting sample, which had almost twice the DNA concentration of the sample used to obtain the results shown in Figure 5B, had the same six-peak pattern in the 11 ppm region. However, the peaks in the 13 ppm region were essentially gone, demonstrating that the parallel-stranded quadruplex forms almost exclusively in KP (pH 7) at a DNA strand concentration of ca. 2 mM.

Evidence for C·C$^+$ Base Pairs in the Quadruplex Structure. The effect of pH on the quadruplex to hairpin equilibrium was determined using CD and the ellipticity assignments described above. Oligonucleotide samples were thermally denatured in NaP buffer adjusted to pH values ranging from 6 to 9 (Figure 6). T_m values obtained from data shown in Figure 6A,B are plotted as a function of pH in Figure 6C. From the curves it is clear that both forms are destabilized as the pH is increased above ca. 8.5. These effects are interpreted as showing that both structures are deprotonated with pK_a values in the 9–9.5 range. This is consistent with loss of the deoxyguanosine imino proton since the pK_a of the deoxynucleoside is in the 9.2–9.4 range (T'so, 1974). In contrast, the quadruplex has a pK_a in the physiological pH range (at 6.8). Since the evidence indicates that the structure is a parallel-stranded quadruplex, two of the three sets of cytidine residues could be constrained between the stacked G-quartets in a conformation that is compatible with base pair formation. Cytidines have been shown to form dimeric complexes in single-stranded and duplex DNA when the base N3 atom of one of the residues is protonated (Edwards et al., 1990). In these complexes, the pK_a increases from ca. 4.2 for the nucleoside (Ts'o, 1974; Sowers et al., 1987a,b; Williams et al., 1987) to around 7. This kind of complex has been found in d(A_2C_2), d(A_2TC_2), d-($C_2A_2T)_{3-6}$, d($C_4A_4T_4C_4$), poly[d(C)], and poly[d(CT)] with pK_a values in the 6–7.4 range. Thus, the unusual pK_a shown for the d(CGCG$_3$GCG)$_4$ complex indicates that C·C$^+$ base pairs are apparently forming. This is the first time that G-rich DNAs have been shown to contain C·C$^+$ base pairs in solution [see Cruse et al. (1983) for a crystal structure]. Thus, a slight change in pH in the physiological range (Ling, 1984) can significantly modulate the stability of the complex.

Discussion

The Quadruplex Structural Motif: Design and Reality. It was anticipated (Sen & Gilbert, 1990) yet somewhat surprising that the [d(CGCG$_3$GCG)$_4$·ion$_n$] complex formed with

the *parallel-stranded* orientation. In contrast, the $d(G_2T_5G_2)_4$ complex forms an antiparallel complexes (Jin et al., 1990; Wang et al., 1991). The initial plans to use the d-($CGCG_nGCG$) sequence motif to study the quadruplex structure relied upon the assumption that the CGC- and GCG-tracts would form WC base-paired duplex "clamps" and thereby force the quadruplex into an *antiparallel* orientation. The cytidines residues apparently perturb the system by stabilizing the complex in the pH 6–8 range via $C{\cdot}C^+$ base pair formation. It is not clear at present how much this influences the strand orientation in the complex. Studies with d-($TATG_nATA$) complexes are being pursued in order to determine the effect of $C{\cdot}C^+$ base pairs on the stability and stand polarity of the complex. The $n = 3$ molecule only forms a minor amount of quadruplex in NaP, while the $n = 4$ molecule denatures at a slightly lower temperature than the [d-$(CGCG_4GCG)_4{\cdot}ion_n$] complex (unpublished experiments). Thus, (at least some of) the terminal residues in the d-($CGCG_nGCG$)-series oligonucleotides make reasonably large contributions to the stability of the quadruplex. It is not clear whether C or G residues outside of the central set of stacked G's make large contributions.

The oligonucleotide $d(CGCG_4GCG)$ forms the parallel-stranded quadruplex almost exclusively in NaP and KP (unpublished experiments). This demonstrates that the stability of the quadruplex relative to that of the hairpin depends upon the number of contiguous G residues. The same conclusion was reached when the stabilities of the *Tetrahymena* and human telomeric DNAs $d(T_2G_4)_4$ and $d(T_2AG_3)_4$ were compared (Hardin et al., 1991). The results showed that quadruplex stability depends upon sequence. Methylation-protection experiments showed that the guanine N7 atoms in $d(CGCG_4GCG)$ are all protected to approximately the same extent (E. Henderson and C. C. Hardin, unpublished experiments). This is the expected result for a parallel-stranded complex. The oligonucleotide $d(CGCG_2GCG)$ also forms an apparent quadruplex (unpublished experiments). The ratio of the 286 and 264 nm peak intensities in KP is approximately 1.5, compared to a ratio of ca. 0.6 with $d(CGCG_3GCG)$ (Figure 2A). Thus, as the number of contiguous G residues decreases from five to three, the amount of WC hairpin increases from almost none to about 65% of the total under the same conditions.

On the basis of the limited number of cases that have been investigated in detail (Sen & Gilbert, 1988, 1990; Jin et al., 1990; Wang et al., 1991; Zimmerman et al., 1975; Howard & Miles, 1982), one can conclude that the relative stabilities of parallel and antiparallel quadruplex species depend upon sequence context. The difference in energy between parallel- and antiparallel-stranded WC duplex conformations in DNA is not large (van de Sande et al., 1988). Due to structural constraints, inter- and intramolecular complexes that can be formed by telomeric G-DNAs must be antiparallel complexes containing folded chains (Henderson et al., 1987; Sundquist & Klug, 1989; Oka & Thomas, 1987; Williamson et al., 1989; Sen & Gilbert, 1990; Raghuraman & Cech, 1990; Hardin et al., 1991). From these considerations, one can speculate that DNA sequences may evolve to adopt structures with different stabilities and strand polarities in order to optimize specific genetic functions. Since these effects are quite dramatic, environmental conditions at the relevent in vivo location may play an important role in driving a given process [see Hardin et al. (1991)].

Potential Genetic Consequences of G-DNA. The apparent stabilizing influence of cytidine residues found with d-($CGCG_3GCG)_4$ is interesting in light of the fact that cytidine residues are rarely present in the consensus nucleotide sequences of telomeric G-rich DNAs (Blackburn, 1986; Blackburn & Szostak, 1984; Henderson & Blackburn, 1989; Forney et al., 1987; Richards & Ausubel, 1988; Roberts, 1988). Known exceptions include sequences from the yeast *Saccharomyces pombe* ($T_{1-2}ACA_{0-1}C_{0-1}G_{1-6}$) and sporozoan *Plasmodium* (T_2T/CAG_3) (Forney et al., 1987; Richards & Ausubel, 1988; Roberts, 1988). The detailed conformational states of telomeric DNAs in vivo are not known. If the postulated intramolecular structures form in vivo, the strand(s) would have to unfold prior to elongation and incorporation into a WC duplex (Boeke, 1990; Hardin et al., 1991). The relative stabilities of the folded and unfolded forms may be finely balanced and carefully modulated in vivo. It has been suggested that quadruplex formation may function to limit "runaway" elongation of telomeric DNAs (Zahler et al., 1991; Hardin et al., 1991). In effect, complex formation functions as a DNA-mediated "repression" mechanism. Thus, cytidines may be uncommon in telomeric DNAs because they would detrimentally perturb this energetic balance. In contrast, if G-DNA is an active component in other genetic contexts in vivo, the extra stabilization energy provided by $C{\cdot}C^+$ base pair formation may be functionally advantageous. Does this mechanism contribute to the telomere-linked transcriptional repression effects described by Gottschling et al. (1990)?

Different structures form when $d(CGCG_3GCG)$ is denatured and then reannealed in the presence of NaP or KP. This is because the quadruplex is much more stable in the presence of K^+. The Ca^{2+} ion is almost as effective as K^+ in stabilizing the quadruplex structure. These cations are present at relatively high concentrations in vivo (millimolar to several hundred millimolar), and the levels may be subject to biological control (Boynton, 1982; Ling, 1984). For example, selective K^+ channels have been shown to exist in mouse nuclear membranes, showing that the nuclear pore is not freely permeable to ions (Mazzanti et al., 1990; Mazzanti, 1991). The quadruplex is very different from the WC hairpin in that it has a pK_a in the physiological range (ca. 6.8). Decreasing the pH below the pK_a results in a large stabilization of the quadruplex. If G-DNA quadruplexes play a role in vivo, these results indicate that the stability could be changed by modest pH changes (<1 pH unit). It has been shown that under some circumstances G-DNA conformations can occur instead of competing WC structures (Hardin et al., 1991; Sen & Gilbert, 1988; Smith et al., 1989). On the basis of the results presented here, it can be concluded that $d(CGCG_3GCG)$ may occur in at least two different forms in vivo, depending upon the ionic conditions, DNA concentration(s), and the presence of other transition-mediating ligands.

Methylation of cytosine at C5 increases the pK_a for protonation at N3 (Sowers et al., 1987a,b; Williams et al., 1987). Since $C{\cdot}C^+$ base-pairing interactions appear to contribute significantly to quadruplex stability, this suggests that m^5C-containing $d(CGCG_3GCG)$ will form a more stable quadruplex than the unmodified oligonucleotide at physiological pH. If G-rich sequences form functionally relevent structures in vivo, this may provide one explanation for the role of DNA methylation in modulating the function of genes in the vicinity of "GpC islands" (Tazi & Bird, 1990; Bernardi, 1989). This mechanism could provide one explanation for chromosome location-dependent "position effects" on gene expression (Gottschling et al., 1990; Weintraub, 1985; John, 1988) and could conceivably be involved in long-term switching events during development and differentiation.

Addendum. Results of SantaLucia et al. (1991) with the ribooligonucleotide r(UGCG$_3$CA) showed that it can form a RNA quadruplex structure. This raises the possibility that "G-RNA" may also form within a biological context and be stabilized by some or all of the factors described in this paper [also see Kim et al. (1991) and Jobling and Gehrke (1987)].

ACKNOWLEDGMENTS

We thank Professor Eric Henderson for allowing us to report unpublished results and for comraderie. We also thank Todd Whitley and Ericka Charleston for technical assistance, Bill Ross and Professors Paul Agris, Ignacio Tinoco, Jr., and Elizabeth Blackburn for encouragement, and Keith Everett for DNA synthesis and photography.

Registry No. C, 71-30-7; K^+, 7440-09-7; d(CGCG$_3$CGC), 137333-01-8.

REFERENCES

Arnold, F. H., Wolk, S., Cruz, P., & Tinoco, I., Jr. (1987) *Biochemistry 26*, 4068–4075.

Bernardi, G. (1989) *Annu. Rev. Genetics 23*, 637–661.

Blackburn, E. H. (1986) in *Molecular Developmental Biology*, pp 69–82, Alan R. Liss, Inc., New York.

Blackburn, E. H. (1991) *Nature 350*, 569–573.

Blackburn, E. H., & Szostak, J. (1984) *Annu. Rev. Biochem. 53*, 163–194.

Boeke, J. D. (1990) *Cell 61*, 193–195.

Boynton, A. L., McKeehan, W. L., & Whitfield, J. F. (1982) in *Ions, Cell Proliferation and Cancer*, Academic Press, New York.

Cech, T. R. (1988) *Nature 332*, 777–778.

Chan, S. S., Breslauer, K. J., Hogan, M. E., Kessler, D. J., Austin, R. H., Ojemann, J., Passner, J. M., & Wilec, N. C. (1990) *Biochemistry 29*, 6161–6171.

Cruse, W. B. T., Egert, E., Kennard, O., Sala, G. B.; Salisbury, S. A., & Viswamitra, M. A. (1983) *Biochemistry 22*, 1833–1839.

Dobler, M. (1981) in *Ionophores and Their Structure*, Wiley-Interscience, New York.

Edwards, E. L., Patrick, M. H., Ratliff, R. L., & Gray, D. M. (1990) *Biochemistry 29*, 828–836.

Forney, J., Henderson, E., & Blackburn, E. H. (1987) *Nucleic Acids Res. 15*, 9143–9152.

Giessner-Prettre, C., Pullman, B., & Caillet, J. (1977) *Nucleic Acids Res. 4*, 99–116.

Gottschling, D. E., Aparicio, O. M., Billington, B. L., & Zakian, V. A. (1990) *Cell 63*, 751–762.

Guschlbauer, W., Chantot, J.-F., & Thiele, D. (1990) *J. Biomol. Struct. Dyn. 8*, 491–511.

Hardin, C. C., Henderson, E., Watson, T., & Prosser, J. K. (1991) *Biochemistry 30*, 4460–4472.

Hare, D. R., & Reid, B. R. (1986) *Biochemistry 25*, 5341–5350.

Henderson, E., & Blackburn, E. H. (1989) *Mol. Cell. Biol. 9*, 345–348.

Henderson, L., Hardin, C. C., Wolk, S. K., Tinoco, I., Jr., & Blackburn, E. H. (1987) *Cell 51*, 899–908.

Henderson, E. R., Moore, E., & Malcolm, B. A. (1990) *Biochemistry 29*, 732–737.

Howard, F. B., & Miles, H. T. (1982) *Biochemistry 21*, 6736–6745.

Jin, R., Breslauer, K. J., Jones, R. A., & Gaffney, B. L. (1990) *Science 250*, 543–546.

Jobling, S. A., & Gehrke, L. (1987) *Nature 325*, 622–625.

John, B. (1988) in *Heterochromatin* (Verma, R. S., Ed.) Cambridge University Press, New York.

Kim, J., Cheong, C., & Moore, P. B. (1991) *Nature 351*, 331–332.

Ling, G. N. (1984) in *Search of the Physical Basis of Life*, Plenum Press, New York.

Mazzanti, M. (1991) *Biophys. J. 59*, 178a.

Mazzanti, M., DeFelice, L. J., Cohen, J., & Malter, H. (1990) *Nature 343*, 764–767.

Oka, Y., & Thomas, C. A. (1987) *Nucleic Acids Res. 15*, 8877–8898.

Otting, J., Grutter, R., Leupin, W., Minganti, C., Ganesh, K. N., Sproat, B. S., Gait, M. J., & Wüthrich, K. (1987) *Eur. J. Biochem. 166*, 215–220.

Pardi, A., & Tinoco, I., Jr. (1982) *Biochemistry 21*, 4686–4693.

Paulson, M. D., Anderson, C. F., & Record, M. T., Jr. (1988) *Biopolymers 27*, 1249–1265.

Raghuraman, M. K., & Cech, T. R. (1990) *Nucleic Acids Res. 18*, 4543–4552.

Record, M. T., Jr., Anderson, C. F., & Lohman, T. M. (1978) *Q. Rev. Biophys. 2*, 103–178.

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

SantaLucia, J., Kierzek, R., & Turner, D. H. (1991) *Biochemistry 30*, 8242–8251.

Sen, D., & Gilbert, W. (1988) *Nature 334*, 364–366.

Sen, D., & Gilbert, W. (1990) *Nature 344*, 410–414.

Smith, S. S., Baker, D. J., & Jardines, L. A. (1989) *Biochem. Biophys. Res. Commun. 160*, 1397–1402.

Sowers, L. C., Shaw, B. R., Veigl, M. L., & Sedwick, M. D. (1987a) *Mutat. Res. 177*, 201–218.

Sowers, L. C., Shaw, B. R., & Sedwick, W. D. (1987b) *Biochem. Biophys. Res. Commun. 148*, 790–794.

Sundquist, W. I., & Klug, A. (1989) *Nature 342*, 825–829.

Tazi, J., & Bird, A. (1990) *Cell 60*, 909–920.

T'so, P. O. P. (1974) in *Basic Principles of Nucleic Acid Chemistry*, Vol. I, Academic Press, New York.

Van de Sande, J. H., Ramsing, N. B., Germann, M. W., Elhorst, W., Kalisch, B. W., Kitzing, E. V., Pon, R. T., Clegg, R. C., & Jovin, T. M. (1988) *Science 241*, 551–557.

Wang, Y., de los Santos, C., Gao, X., Gaffney, B., Jones, R., Breslauer, K., Live, D., & Patel, D. (1991) *Biophys. J. 59*, 493a.

Weintraub, H. (1985) *Cell 42*, 705–711.

Williams, L. D., Chawla, B., & Shaw, B. R. (1987) *Biopolymers 26*, 591–603.

Williamson, J. R., Raghuraman, M. K., & Cech, T. R. (1989) *Cell 59*, 871–880.

Zahler, A. M., Williamson, J. R., Cech, T. R., & Prescott, D. M. (1991) *Nature 350*, 718–720.

Zimmerman, S. B., Cohen, G. H., & Davies, D. R. (1975) *J. Mol. Biol. 92*, 181–192.