The Trinucleotide Repeat Sequence d(CGG)$_{15}$ Forms a Heat-Stable Hairpin Containing G^{syn}•G^{anti} Base Pairs†

Michael Mitas,*,‡ Adong Yu,‡ Jeffrey Dill,‡ and Ian S. Haworth§

Department of Biochemistry and Molecular Biology, 246 Noble Research Center, Oklahoma State University, Stillwater, Oklahoma 74078, and Department of Pharmaceutical Sciences, University of Southern California, 1985 Zonal Avenue, Los Angeles, California 90033

Received May 22, 1995; Revised Manuscript Received July 26, 1995®

ABSTRACT: To investigate potential structures of d(CGG/CCG)$_n$ that might relate to their biological function and association with triplet repeat expansion diseases (TREDs), electrophoretic mobility, chemical modification, and P1 nuclease studies were performed with a single-stranded (ss) oligonucleotide containing (CGG)$_{15}$ [ss(CGG)$_{15}$]. The results suggest that ss(CGG)$_{15}$ forms a hairpin with the following features: (i) a stem containing G^{syn}•G^{anti} base pairs; (ii) at ≥200 mM K$^+$, CGG repeats on the 5′ portion of the stem base-paired to GCG repeats on the 3′ side (referred to as the (*b*) alignment); and (iii) heat stability (T_m = 75 °C in low ionic strength). At ≤100 mM K$^+$, dimethyl sulfate reactions indicated that the hairpin in the (*b*) alignment was in equilibrium with another structure, presumably a hairpin in the alternative (*a*) alignment (CGG repeats on the 5′ portion of the stem base-paired to CGG repeats on the 3′ portion of the stem). Molecular dynamics simulations suggested that the loop region of the (*a*) alignment contained two guanines stacked on top of one another. The same guanines in the (*b*) alignment were base-paired in a *syn*−*anti* arrangement. We propose that the stability of the loop partially determines the stem alignment.

To aid in the correlation of potential structures of triplet repeat nucleic acids with their function and their propensity to undergo expansion events, we described a sequence-based classification system for them (Mitas et al., 1995). Class I repeats, which were defined by the presence of a GC or CG palindrome, had the lowest base-stacking energies, exhibited the lowest rates of slippage synthesis (Schlötterer & Tautz, 1992), and were uniquely associated with triplet repeat expansion diseases (TREDs).[1] All six complementary single strands of Class I triplet repeats potentially formed stable hairpin structures. In support of this possibility, evidence was presented that single-stranded (ss) oligonucleotides containing 15 or more CTG (Mitas et al., 1995; Gacy et al., 1995) or GTC (Yu et al., 1995) repeats formed hairpins containing base-paired and/or base-stacked thymines in the stem.

Here, we report the results of studies conducted on the third member of ss Class I triplet repeats, ss(CGG)$_{15}$. The results provide evidence that ss(CGG)$_{15}$ forms an extremely heat-stable "slipped" hairpin containing G^{syn}•G^{anti} base pairs.

MATERIALS AND METHODS

Oligonucleotides. All oligonucleotides were synthesized on an Applied Biosystems 381A oligonucleotide synthesizer (Foster City, CA) with the trityl group on and purified with oligonucleotide purification cartridges (Cruachem, Glasgow, U.K.). Sequences of oligonucleotides were (CCG)$_{15}$, GATCC-(CCG)$_{15}$GGTACCA; (CGG)$_{15}$, AGCTTGGTACC(CGG)$_{15}$G; and (CTG)$_{15}$, GATCC(CTG)$_{15}$GGTACCA.

Plasmid DNA Preparation. Plasmid pCTG15 was described elsewhere.[1] The plasmid bearing (CGG)$_{15}$ (named pCCG15) was prepared in a manner similar to that described for pCTG15.

DNA Sequencing. To confirm that (CGG)$_{15}$ was correctly inserted into pCCG15 without nucleotide sequence alteration, the triplet repeat region was sequenced as previously described for pGTC15 (Yu et al., 1995). No deletions or mutations were detected in the triplet repeat region of pCCG15.

Dimethyl Sulfate Reactions. pCCG15 (15 μg) was first digested with *Hind*III and dephosphorylated with calf intestinal phosphatase. DNA was labeled at the 5′ end by incubation in buffer containing 1 mM DTT, 1 μL of 7000 Ci/mmol ATP [γ-^{32}P] (ICN, Irvine, CA), 50 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 0.1 mM spermidine, 0.1 mM EDTA, and 20 units of T4 polynucleotide kinase (New England Biolabs, Beverly, MA) for 1 h at 37 °C. Labeled DNA was extracted with 25:24:1 phenol:chloroform:isoamyl alcohol (PCI) and precipitated with ethanol. Resuspended DNA was digested with 50 units of *Bam*HI in a volume of 70 μL and applied directly to a Nuctrap column (Stratagene, La Jolla, CA) for removal of unincorporated ^{32}P-dNTPs. To separate labeled vector DNA from labeled oligonucleotide, DNAs were subjected to electrophoresis in a 2% agarose gel. Oligonucleotide-containing CGG$_{15}$ was excised from

† This work and M.M., A.Y., and J.D. were supported by the Oklahoma Agricultural Experimental Station at Oklahoma State University.

* Corresponding author: Michael Mitas, Ph.D., Department of Biochemistry and Molecular Biology, 246 Noble Research Center, Oklahoma State University, Stillwater, OK 74078. Telephone: 405-744-6199. Fax: 405-744-7799. E-mail address: mmitas@bmb-fs1.biochem.okstate.edu.

‡ Oklahoma State University.

§ University of Southern California.

® Abstract published in *Advance ACS Abstracts,* September 15, 1995.

[1] Abbreviations: TREDs, triplet repeat expansion diseases; ss, single-stranded; ds, double-stranded; DMS, dimethyl sulfate; *Eco*SSB, *Escherichia coli* single-stranded DNA binding protein; EMMP, electrophoretic mobility melting profile; T_m, melting temperature; T_i, isomobility temperature.

0006-2960/95/0434-12803$09.00/0

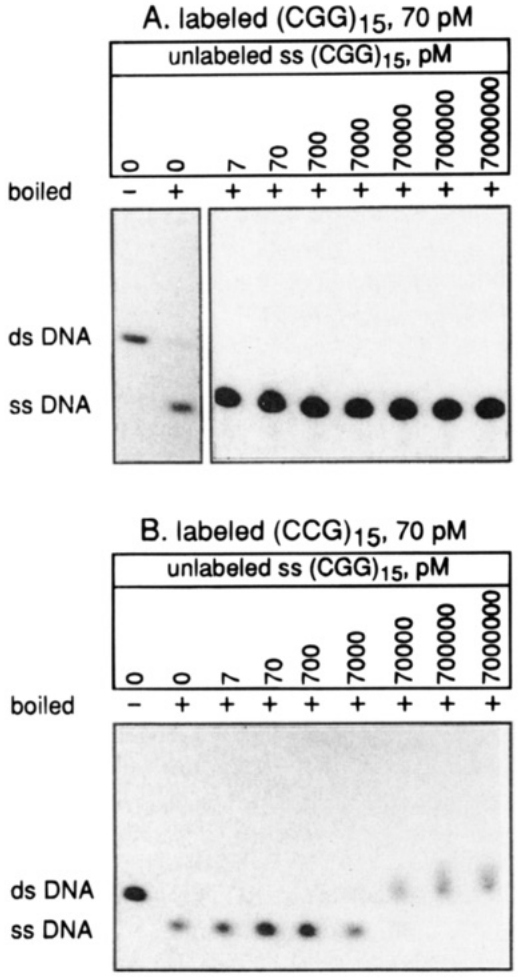


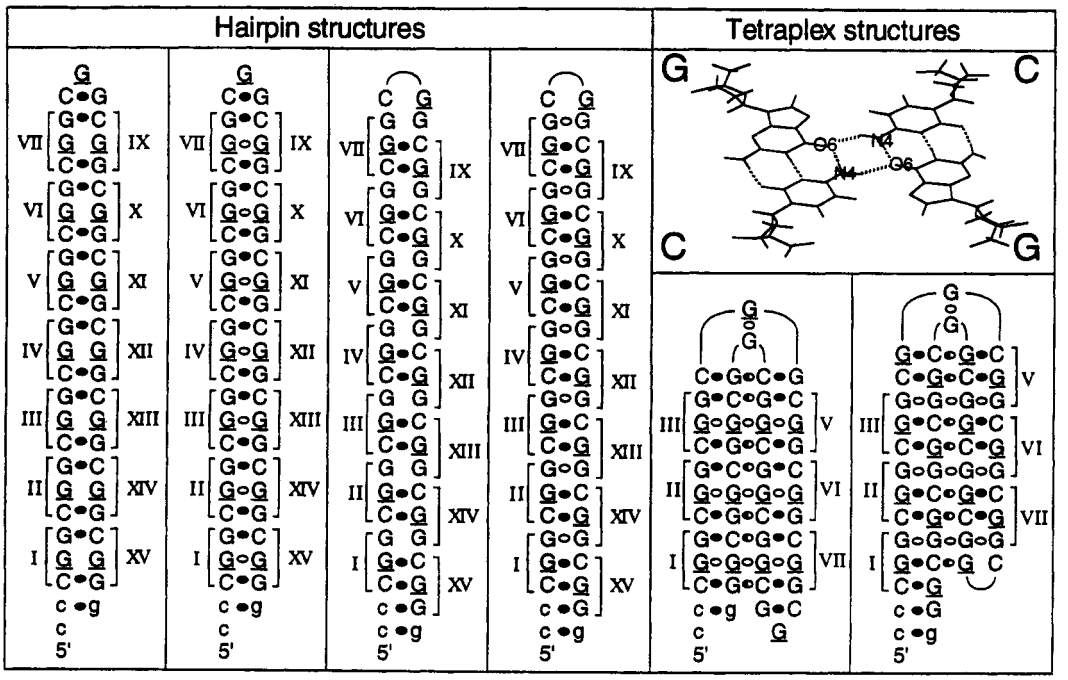
FIGURE 1: Structure of ss$(CGG)_{15}$ is concentration-independent. A ds oligonucleotide containing $(CGG)_{15}$ was excised from plasmid pCCG15 as described in Materials and Methods. Strands end-labeled with the use of Klenow enzyme were (A) ss$(CGG)_{15}$ and (B) ss$(CCG)_{15}$. Prior to gel electrophoresis, unlabeled synthetic oligonucleotide containing $(CGG)_{15}$ at the above concentration (in picomolar) was added to the indicated labeled strand, and the mixture was placed in a boiling H_2O bath for 10 min and cooled at room temperature for 15 min. Electrophoretic conditions of the gels were equivalent. Electrophoresis was performed at 16 °C.

the agarose gel and purified with glass beads (Mermaid Kit, Bio101, La Jolla, CA). Unlabeled synthetic oligonucleotide (1.4 pmol) of the same sequence as the labeled strand was added to 4×10^3 dpm (~0.7 fmol) of labeled DNA and the mixture placed in a boiling H_2O bath for 3 min and placed on ice for 5 min. This procedure resulted in complete conversion of all labeled DNA to the ss form (see Figure 1). Dimethyl sulfate (DMS) reactions were performed at 37 °C for 3 min essentially according to the method of Maxam and Gilbert (1980) as previously described (Mitas et al., 1995). Reaction products were analyzed on a sequencing gel containing 20% polyacrylamide and 8 M urea.

P1 Nuclease Digestion. pCCG15 (15 μg) was digested with 50 units of *Bam*HI for 1 h at 37 °C in a volume of 70 μL. Recessed ends were labeled at the 3′ termini by addition of 5 μL of α-^{32}P-dCTP, 5 μL of α-^{32}P-dATP (each 3000 Ci/mmol, ICN, Irvine, CA), 2.5 μL of 5 mM dTTP and dGTP, and 25 units of Klenow enzyme (New England Biolabs). Reaction mixtures were incubated for 1 h at room temperature. Plasmid DNAs were extracted with PCI and precipitated with ethanol. Resuspended DNA was digested with 50 units of *Hind*III in a volume of 70 μL and applied directly to a Nuctrap column. Purification of the oligonucleotide-containing CGG15 was as described above for the 5′ end-labeled fragment. Unlabeled synthetic oligonucleotide (1.4 pmol) of the same sequence as the labeled strand was added to 4×10^3 dpm (~0.7 fmol) of 3′ end-labeled DNA and the mixture placed in a boiling H_2O bath for 3 min and placed on ice for 5 min. P1 nuclease digestions were performed in 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, and 50 mM NaCl at 37 °C essentially according to the method of Wohlrab (1992) as previously described (Mitas et al., 1995). Reaction products were analyzed on a sequencing gel containing 20% polyacrylamide and 8 M urea.

Electrophoretic Mobility Melting Profiles. Oligonucleotide-containing $(CGG)_{15}$ was labeled at the 3′ end with the use of Klenow enzyme as described above, except that the oligonucleotide was not purified from vector DNA. Oligonucleotide-containing $(CTG)_{15}$ was prepared in a manner similar to that described for $(CGG)_{15}$ except that the starting plasmid was pCTG15. Also, the order of restriction enzymes was reversed. The electrophoretic mobility melting profile (EMMP) technique was described previously (Yu et al., 1995). Briefly, to obtain a homogeneous population of labeled double-stranded (ds) DNAs, oligonucleotides liberated from plasmid at 0.7 nM were incubated with 0.5 nM *Escherichia coli* single-stranded DNA binding protein (Lohman & Ferrari, 1994) (*Eco*SSB; a generous gift of Dr. Timothy Lohman, Washington University School of Medicine) in 8% glycerol, 0.2 M NaCl, 18 mM HEPES (pH 7.40), 1 mM EDTA, and 1 mM DTT at 37 °C for 20 min. To obtain a homogeneous population of labeled ss$(CTG)_{15}$, 1 μM unlabeled synthetic oligonucleotide of the same sequence as the labeled strand was added and the mixture placed in a 90 °C H_2O bath for 5 min and kept at 25 °C for 5 min. To obtain a homogeneous population of labeled ss$(CGG)_{15}$, unlabeled synthetic oligonucleotide of the same sequence as the labeled strand was added to yield a final concentration of 1 μM. The DNAs were placed in a 100 °C H_2O bath for 5 min and kept at 25 °C for 5 min. For electrophoretic analysis, DNAs (4×10^4 dpm) were diluted to 10 μL in buffer containing 8% glycerol, 10 mM HEPES (pH 8.5), and 1 mM EDTA. Loading dye (1 μL) (50% glycerol, 0.4% bromophenol blue) was added to the DNA samples prior to gel electrophoresis. Electrophoresis was performed in a Hoeffer (San Francisco, CA) SE600 series unit at various temperatures at 25 mA/gel in 45 mM Tris-borate (pH 8.5) and 1 mM EDTA (TBE). Gel plates were 14 cm (length) × 16 cm (width) × 1.5 mm (thickness). Electrophoresis was stopped when the bromophenol blue marker migrated 10 cm. Dried gels were placed between two intensifying screens (Dupont) and exposed to Fuji RX film 3 h to overnight at −80 °C.

Annealing Studies. Single-stranded $(CGG)_{15}$ was labeled as described in the EMMP studies. Labeled ss$(CGG)_{15}$ (0.07 nM) in 10 mM Tris-HCl (pH 7.9) and 1 mM EDTA was heated to 100 °C for 5 min and immediately placed on ice. Aliquots (8 μL) were added to 400 μL Eppendorf tubes and placed on ice. Unlabeled synthetic ss oligonucleotide containing ss$(CCG)_{15}$ was added to each tube (final concentration = 7 μM). A 2 L beaker containing H_2O was heated to 95 °C and cooled to ~50 °C over a period of 2 h. At various temperatures, the 400 μL Eppendorf tubes were placed in the 2 L beaker. The DNA within each tube was



	1	2	3	4	5	6
structure number						
hairpin alignment	*(a)*	*(a)*	*(b)*	*(b)*	*(a)*	*(b)*
G-G base pairs?	no	yes	no	yes	yes	yes
# of potential H bonds	48	62	48	62	88	88
DMS reactivity: 5′G/3′G	y/y = 1.0	0.5y/y = 0.5	y/y = 1.0	y/0.5y = 2.0	0/y = 0	y/0 ~ infinity
triplets susceptible to P1	VIII	VIII	VIII	VIII	IV, VIII, XII	IV, VIII, XII

FIGURE 2: Potential hairpin and tetraplex structures of ss(CGG)$_{15}$. Four potential hairpin and two potential tetraplex structures of ss(CGG)$_{15}$ are shown. Nontriplet repeat sequences of ss(CGG)$_{15}$ anticipated to form base pairs are in lowercase letters. Triplet repeat numbers are indicated by roman numerals. Watson–Crick C•G base pairs are indicated by filled ovals. Non-Watson–Crick G•G base pairs are indicated by open ovals. To distinguish the 5′G in a triplet repeat from the 3′G, the 5′G is underlined. In hairpin alignment (*a*), the 5′Gs are opposite one another. In hairpin alignment (*b*), the 3′Gs are opposite one another. Non-Watson–Crick C•G base pairs (described below) in the tetraplex structures are indicated by half-filled ovals. In the tetraplex structures, the non-Watson–Crick base pairs that form between the left-most strand (triplets I–III) and the right-most strand (triplets V–VII) are not shown. Tetraplex structure 5 was constructed by the folding in half of the hairpin in alignment (*a*). Tetraplex structure 6 was constructed in a similar manner from the hairpin in the (*b*) alignment. Molecular-modeling studies revealed that structures 5 and 6 contained favorable G_4 quartets and CGCG quartets. The CGCG quartet, obtained from computer modeling studies of structures 5 and 6, is shown in the upper right corner. Dotted lines in the CGCG quartet indicate H bonds. Vertically oriented H bonds are from Watson–Crick base pairs derived from the hairpin. Horizontally oriented H bonds arise from hairpin–hairpin interactions. Energy minimization indicated favorable G•G base pairs in the tetraplex loops. The CGCG quartet shown above differs from the CGCG quartet arrangement crystallized by Leonard et al. (1995). First, we found that the CGCG quartets more appropriately fit into the tetraplex structure with interactions between the functionality (O6 of guanine and the 4-amino group of cytosine) from the major groove side of the two G•C base pairs. This is in contrast with the Leonard et al. structure, in which the minor groove functions (the 2-amino group of guanine and the O2 of cytosine) interact in the center of the quartet. Second, we were able to model a structure having an essentially coplanar arrangement of the two Watson–Crick base pairs within a CGCG quartet, in contrast to the 30° rotation angle between the pairs in the Leonard structure. This planarity is important in increasing the stacking interaction with the planar guanine quartet. Hairpin alignments other than those shown contained fewer Watson–Crick base pairs and were not considered thermodynamically favorable. Expected features of structures 1–6 are listed below each structure and are discussed in the subsequent text. The variable *y* refers to an arbitrary measurement of dimethyl sulfate (DMS) reactivity of a given G residue within a triplet repeat. For simplicity, it was assumed that a G residue within a G•G base pair in a hairpin would alternate between the *syn* and *anti* conformations. Therefore, the DMS reactivity of a G residue that alternated between the two conformations should be 0.5, relative to the DMS reactivity of a G residue within a C•G base pair.

subsequently analyzed in an 8% polyacrylamide gel as described for the EMMP studies. The temperature of the gel was 20 °C. Autoradiographs of each gel were scanned three times with a PDI densitometer (Huntington Station, NY) to determine percent conversion of ss(CGG)$_{15}$ to ds(CGG)$_{15}$. Background subtracted from the scanned images was 0.15 optical density unit.

Computer Modeling. The ssDNA sequence 5′-AGCTTGGTACC(CGG)$_{15}$G can potentially fold into two different hairpin alignments which maximize base pairing in the hairpin stem. These are designated (*a*) and (*b*) and illustrated in Figure 2. Using QUANTA4.0 (Molecular Simulations, Burlington, MA), each hairpin was constructed with an initial assumption of Watson–Crick pairing for all G•C and C•G base pairs, including the C•G "pairs" in triplet VIII that are in the loop regions of the two alignments. The single guanine in the loop region of alignment (*a*) was initially included as an overhanging end. Hence, the loops of each hairpin were initially in an "unclosed" state and were minimized into a favorable conformation as described below. The overhanging 5′ end of the sequence was included in a random conformation.

A series of energy minimizations was performed on a Silicon Graphics Indigo workstation using the AMBER4.0 force field (Pearlman et al., 1994). First, 2000 cycles of minimization were performed on the loop structure only, followed by 1000 cycles on the G•G base pairs only and finally a 4000-cycle relaxation of the whole hairpin. All-

atom force field parameters and AMBER4.0 charges were assigned to the DNA. A distance-dependent dielectric of the form $\epsilon = 4r_{ij}$ was used with a 12 Å residue-based nonbonded cutoff.

In each alignment, the G•G mismatch pairs were included as either an *anti*−*anti* or an *anti*−*syn* base pair, with the latter arrangement being constructed from the former by a 180° rotation about the glycosidic torsion angle of one of the guanines, followed by a constrained minimization of the G•G pairs only to form the appropriate hydrogen bonds. Distance constraints [of 2.0 Å with a force constant of 1000 kcal (mol^{-1} $Å^{-1}$)] were used to pair the guanine bases (N7(*syn*)−NH_2(*anti*) and O6(*syn*)−H1(*anti*)). These constraints were then removed in the minimization of the whole hairpin.

Molecular dynamics simulations of the two alignments solvated in a periodic waterbox were also performed. The length of the hairpin stem was reduced in these calculations. Hence, triplets III−VI and X−XIII were removed from the structures, as was the 5′-AGCTTGGTA end. This leaves a single, overhanging 5′ cytosine end in the (*a*) alignment and a fully base-paired stem in the (*b*) alignment (although the total number of pairs is still formally equivalent, assuming the CG loop of the (*b*) alignment is not base-paired). These structures were subjected to a minimization process similar to that described above and then solvated in a periodic box of TIP3P water molecules (Jorgensen et al., 1983). Only alignments containing *syn*−*anti* G•G pairs were considered. Twenty-three sodium counterions were added at the points of most negative potential (van Gunsteren et al., 1986). The boxes contained a total of 2289 water molecules for alignment (*a*) and 2175 water molecules for alignment (*b*), with a minimum distance of about 8 Å between the DNA and the edge of the box. Molecular dynamics simulations using the nPT ensemble were performed for 100 ps at 298 K (including linear heating from 0 to 298 K in the first 10 ps) for each hairpin, using a dielectric constant of 1, a step length of 0.002 ps, and SHAKE bond length constraint.

RESULTS

To analyze DNA containing CGG repeats, a double-stranded (ds) oligonucleotide containing 15 CGG triplet repeats was cloned into plasmid as described in Materials and Methods. Oligonucleotides liberated from plasmid were utilized for studies since they were unequivocally full length. For clarity, the triplet repeats within the ss oligonucleotide containing $(CGG)_{15}$ [ss$(CGG)_{15}$] are indicated in the subsequent text by roman numerals.

To determine whether ss$(CGG)_{15}$ exhibited properties of an intramolecular hairpin or some type of intermolecular structure, various studies were performed. First, the molecular composition of the structure(s) formed with ss$(CGG)_{15}$ was investigated by the performing of electrophoretic studies with labeled ss$(CGG)_{15}$ mixed with various amounts of unlabeled ss synthetic oligonucleotide of the same sequence (Figure 1A). If ss$(CGG)_{15}$ formed a stable intramolecular hairpin structure, increasing the concentration of unlabeled synthetic oligonucleotide should not result in the appearance of a form of DNA that migrated with slow electrophoretic mobility.

To determine the electrophoretic properties of ds$(CGG)_{15}$ and ss$(CGG)_{15}$, the ds oligonucleotide-containing labeled $(CGG)_{15}$ [and an equimolar amount of unlabeled $(CCG)_{15}$] was subjected to electrophoresis before and after being heated to 100 °C. Heating resulted in conversion to a species of DNA that migrated with relatively fast electrophoretic mobility (Figure 1A), indicating that the electrophoretic mobility of ss$(CGG)_{15}$ was faster than that of ds$(CGG)_{15}$. Addition of a 10^5-fold molar excess (final DNA concentration = 7 μM) of unlabeled ss synthetic ss$(CGG)_{15}$ did not result in the formation of the slow-migrating complex, indicating that ss$(CGG)_{15}$ formed a stable unimolecular structure.

To demonstrate that the unlabeled ss synthetic oligonucleotide-containing $(CGG)_{15}$ was not degraded and contained CGG repetitive sequences, a control experiment was performed with a ds oligonucleotide-containing labeled $(CCG)_{15}$ and unlabeled $(CGG)_{15}$ (Figure 1B). Addition of increasing amounts of unlabeled ss synthetic oligonucleotide to labeled ss$(CCG)_{15}$ resulted in complete conversion of the fast-migrating ss form to the slow-migrating ds form, indicating that the unlabeled ss synthetic oligonucleotide-containing $(CGG)_{15}$ was not degraded and contained CGG repetitive sequences.

Dimethyl Sulfate Reactions with ss(CGG)$_{15}$. Six potential intramolecular structures of ss$(CGG)_{15}$ are shown in Figure 2. Analysis of the DMS reactivities of the two Gs within a given triplet (the 5′G of a given triplet is herein designated G, while the 3′G is designated G) should provide clues regarding which of the six intramolecular structures was adopted by ss$(CGG)_{15}$ (Figure 2). For example, if ss$(CGG)_{15}$ formed a tetraplex in alignment (*a*) (structure 5 in Figure 2), the DMS reactivity of the 5′G within a given triplet should approximate 0 since the N7 of a G in a G quartet will not react with DMS. In contrast, the DMS reactivity of the 3′G within the same triplet should be some measurable quantity (*y*). Therefore, the 5′G/3′G DMS reactivity ratio within a given triplet for structure 5 is $0/y = 0$. Structures 2 and 4−6 are characterized by unique 5′G/3′G DMS reactivity ratios. The 5′G/3′G DMS reactivity ratios of structures 1 and 3 are identical to one another, but different from the rest (Figure 2).

DMS reactions were performed with ss$(CGG)_{15}$ at 10 °C in 50 mM Na^+ and various concentrations of KCl (Figure 3). With no added KCl, the G residues in the middle of the triplet region (G31−G37) were hypersensitive to DMS, indicating that this region formed a loop. DMS hypersensitivity was not observed at triplets IV or XII. These results are consistent with a hairpin structure of ss$(CGG)_{15}$. As the concentration of KCl was increased, the relative reactivities of the Gs in the loop region decreased, indicating that the loop region was significantly stabilized by salt.

The autoradiograph was scanned with a densitometer to determine the relative intensities of the adjacent G residues within a triplet repeat. At 0.75 M KCl, the reactivity of the 5′G, relative to that of the 3′G within a CGG triplet, was 1.99 ± 0.16. DMS reactivities similar to those in 0.75 M KCl were observed at 0.4 or 0.2 M KCl (Table 1). These results are inconsistent with a tetraplex structure of ss$(CGG)_{15}$ and suggest that, within the stem region, the 3′Gs of the CGG triplets formed G^{syn}•G^{anti} base pairs characteristic of a hairpin in the (*b*) alignment (structure 4 in Figure 2).

In the absence of added KCl, the 5′G/3′G ratio of DMS reactivities was 1.40 ± 0.13. These results suggest that only ~40% of the G−G mismatches were base-paired or, alternatively, that ss$(CGG)_{15}$ was in alignment (*b*) 70% of the time and alignment (*a*) 30% of the time.

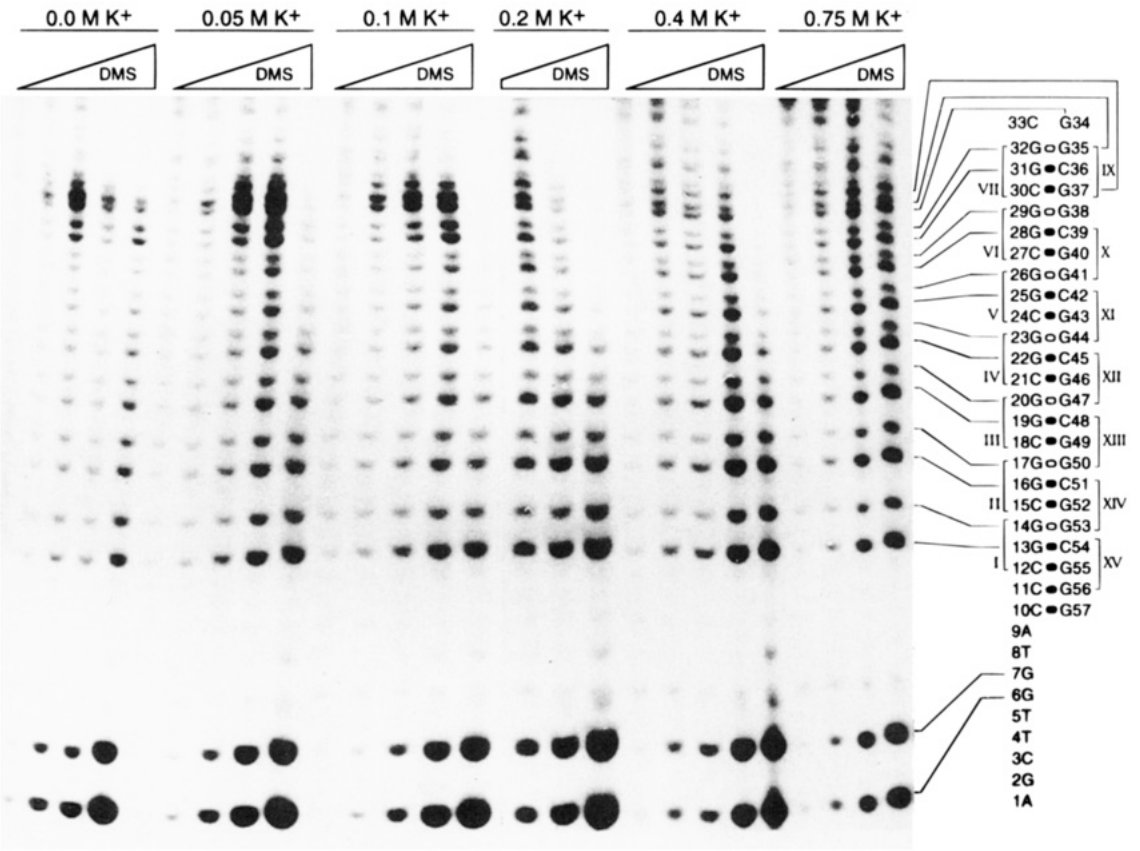


FIGURE 3: Dimethyl sulfate reactions with ss(CGG)$_{15}$ at various K^+ concentrations. DMS methylation was performed at 10 °C with ss(CGG)$_{15}$ as described in Materials and Methods. Reaction mixtures contained 50 mM Na^+ and various concentrations of K^+ (as indicated). For a given concentration of K^+, the concentration of DMS reacted with ss(CGG)$_{15}$ was (from left to right) 0 mM, 5.3 mM, 21 mM, 84 mM, and 0.42 M (with the exception of 0.2 M K^+, where the concentrations of DMS were 21 mM, 84 mM, and 0.42 M). Reaction mixtures were applied to a sequencing gel containing 20% polyacrylamide and 8 M urea. The deduced structure of ss(CGG)$_{15}$ in 0.75 M K^+ is shown to the right of the gel. Open ovals denote G•G base pairs. Filled ovals denote C•G base pairs.

Table 1: Dimethyl Sulfate Reactivities of Adjacent G Residues in the Stem of ss(CGG)$_{15}$[a]

KCl, M	DMS reactivities 5′G/3′G
0	1.40 ± 0.13
0.05	1.56 ± 0.15
0.10	1.82 ± 0.22
0.20	2.03 ± 0.21
0.40	1.96 ± 0.14
0.75	1.99 ± 0.16

[a] The autoradiograph shown in Figure 3 was scanned with a PDI densitometer (×3). For reaction products generated at the specified KCl concentration, the optical density of a band derived from cleavage of the 5′G within a given triplet repeat was determined and divided by the optical density of the 3′G within the same triplet repeat. Optical density values obtained from triplet repeats I−III were used for data analysis. Values in the table represent the mean ± standard deviation.

The Hairpin Structure of ss(CGG)$_{15}$ Is Observed at Temperatures between 1 and 37 °C. The DMS experiments described above were performed at 10 °C. To determine the effect of temperature on the deduced hairpin structure of ss(CGG)$_{15}$, DMS experiments were performed with ss(CGG)$_{15}$ at 1, 5, 10, 23, and 37 °C in 50 mM Na^+ and 100 mM K^+. We were particularly interested in results at lower temperatures, since these conditions might favor tetraplex formation. Reactions performed at lower or higher temperatures yielded results similar to those reactions conducted at 10 °C (Figure 4), suggesting that the hairpin structure of ss(CGG)$_{15}$ was stable over a range of temperatures.

Single-Strand Specific P1 Nuclease Cleaves the Middle of the Triplet Repeat Region in ss(CGG)$_{15}$. To further characterize the secondary structure of ss(CGG)$_{15}$, digestions were performed with single-strand specific P1 nuclease in 50 mM Na^+ at pH 7.5 (Figure 5). The results show significant cleavage of the G32−C33 and G34−G35 phosphodiester bonds. Cleavage of phosphodiester bonds in triplets IV or XII was not observed. These results are consistent with a hairpin structure of ss(CGG)$_{15}$ and are not consistent with a tetraplex structure (Figure 2).

In the nontriplet repeat region of the DNA, only minor cleavages of the G58−A59 and A59−T60 phosphodiester bonds were observed. These results indicate extensive base-pairing and/or base-stacking interactions within the nontriplet repeat region. Due to the specific alignment of the hairpin structure, A59 and C61 can form base pairs with T8 and G6, respectively (as shown in Figure 5).

Melting Profile of ss(CGG)$_{15}$. The results from Figure 1 showed that ss(CGG)$_{15}$ migrated with rapid mobility at 16 °C. The rapid electrophoretic mobility of ss(CGG)$_{15}$ might have been due to a hairpin structure. We hypothesized that, if the rapid mobility were due to a hairpin conformation, application of heat to the polyacrylamide would simultaneously denature the hairpin and reduce its electrophoretic mobility (relative to its ds form). To investigate this possibility, an electrophoretic mobility melting profile (EMMP) of ss(CGG)$_{15}$ was generated. EMMPs, like temperature gradient gel electrophoresis (Ke & Wartell, 1993; Wartell

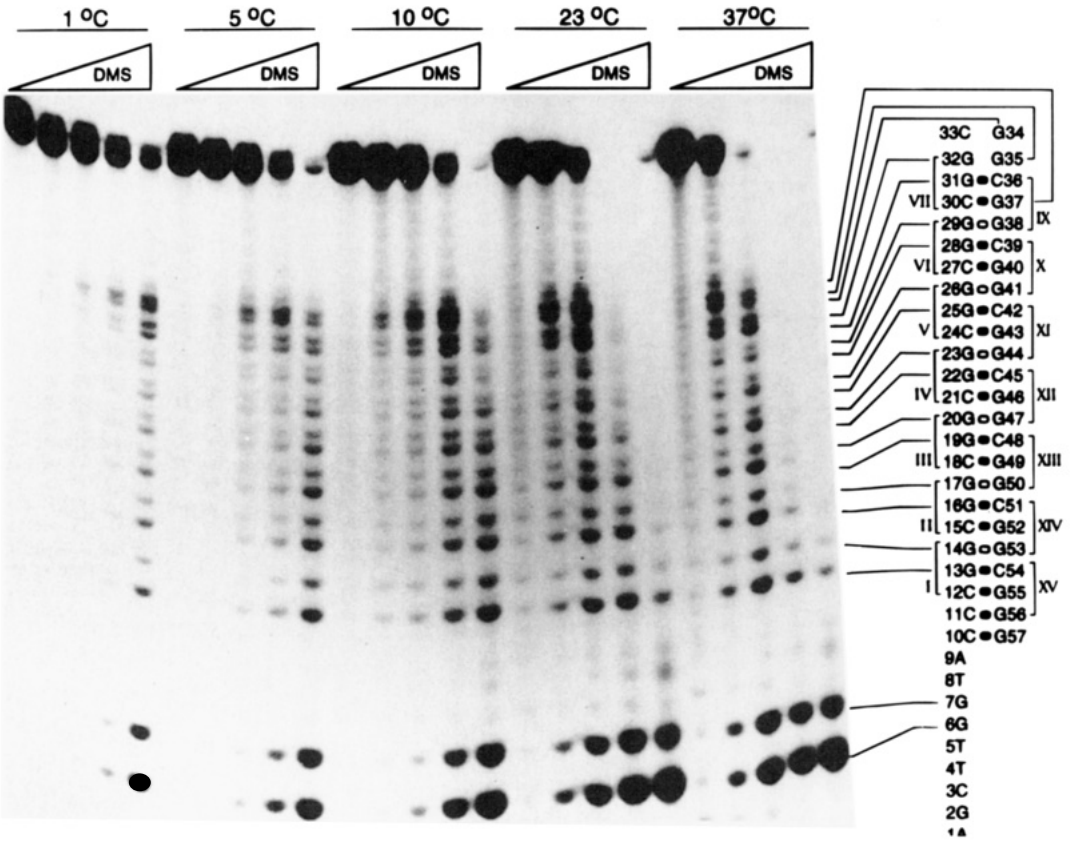


FIGURE 4: Dimethyl sulfate reactions with ss(CGG)$_{15}$ at various temperatures. DMS methylation of ss(CGG)$_{15}$ was performed in 50 mM Na^+ and 100 mM KCl at the indicated temperature as described in Materials and Methods. For a given temperature, the concentration of DMS reacted with ss(CGG)$_{15}$ was (from left to right) 0 mM, 5.3 mM, 21 mM, 84 mM, and 0.42 M. Reaction mixtures were applied to a sequencing gel containing 20% polyacrylamide and 8 M urea.

et al., 1990), can be used to estimate the melting temperature (T_m) of a DNA by determination of the midpoint of its electrophoretic phase transition (Yu et al., 1995). For control in the EMMP analysis, ss(CTG)$_{15}$, a hairpin that contains base-paired and/or base-stacked thymines in the stem, was analyzed. Previous EMMP studies have shown that the T_m of ss(CTG)$_{15}$ is approximated by the temperature at which the electrophoretic mobilities of ss(CTG)$_{15}$ and ds(CTG)$_{15}$ are equal (referred to as the isomobility temperature, or T_i). At low ionic strength, the T_m of ss(CTG)$_{15}$ is 48 °C, whereas the T_i is 50 °C (Yu et al., 1995). For reference in the EMMP studies described below, the ds forms of the DNAs are used. The dsDNAs contain more than twice as many Watson–Crick C•G base pairs compared to their respective ss hairpin forms and hence are far more stable.

The EMMPs of ss(CTG)$_{15}$ and ss(CGG)$_{15}$ are shown in Figure 6A and plotted in Figure 6C. In agreement with previous data (Yu et al., 1995), the T_m of ss(CTG)$_{15}$ was 48 °C. At the highest temperature (66 °C), the electrophoretic mobility of ss(CGG)$_{15}$ was not reduced to the same extent as that of ss(CTG)$_{15}$. This result suggests that the structure of ss(CGG)$_{15}$ was more heat-stable than the hairpin structure of ss(CTG)$_{15}$. Since ss(CGG)$_{15}$ migrated faster than its ds form at 66 °C, the results also suggest that the T_m of ss(CGG)$_{15}$ was above 66 °C.

To obtain an estimate of T_m for ss(CGG)$_{15}$, annealing experiments were performed with labeled ss(CGG)$_{15}$ in the hairpin conformation and a 10^6-fold molar excess of unlabeled ss(CCG)$_{15}$ (Figure 6B). The T_m of ss(CGG)$_{15}$ was estimated by determination of the amount of heat required to anneal ss(CGG)$_{15}$ to ss(CCG)$_{15}$. Extent of annealing was determined by the measuring of the conversion of the hairpin form of (CGG)$_{15}$ to its Watson–Crick ds form. The results of three independent annealing experiments are superimposed upon the EMMP results (Figure 6C). Incubation of labeled ss(CGG)$_{15}$ with unlabeled ss(CCG)$_{15}$ at or below 58 °C did not result in conversion of labeled ss(CGG)$_{15}$ to its ds form. This result was consistent with the EMMP studies which indicated that the electrophoretic mobility of ss(CGG)$_{15}$ was rapid at ≤58 °C (Figure 6A). At 75 °C, one-half of the ss form of (CGG)$_{15}$ was converted to its ds form. These results indicate that the T_m of ss(CGG)$_{15}$ was approximately 75 °C, 27 °C higher than the T_m of ss(CTG)$_{15}$. The results indicate that G•G base pairs contribute a significant amount of stability to the hairpin structure of ss(CGG)$_{15}$.

Computer Modeling of the ss(CGG)$_{15}$ Hairpin Structure. The experimental data above are consistent with ss(CGG)$_{15}$ folding into a hairpin conformation in alignment (*b*) at or above 200 mM K^+. In an attempt to understand the preferential formation of (*b*) over (*a*), each alignment was examined theoretically, using minimization and molecular dynamics simulations. Minimized energies of the G^{anti}•G^{anti} paired hairpins (structures 1 and 3 in Figure 2) were very similar. Inclusion of G^{syn}•G^{anti} pairs shown in Figure 7 (structures 2 and 4 in Figure 2) with the guanines in trinucleotides I–VII in *syn* conformations stabilized the hairpins substantially, as expected, but provided no indication of a significantly greater stability associated with one alignment or the other. Minimization of a hairpin in alignment (*b*) with G^{syn}•G^{anti} pairing such that the guanosines of trinucleotide repeats I, III, V, and VII were *syn* and II, IV, and VI *anti* had no effect on the minimized energy,

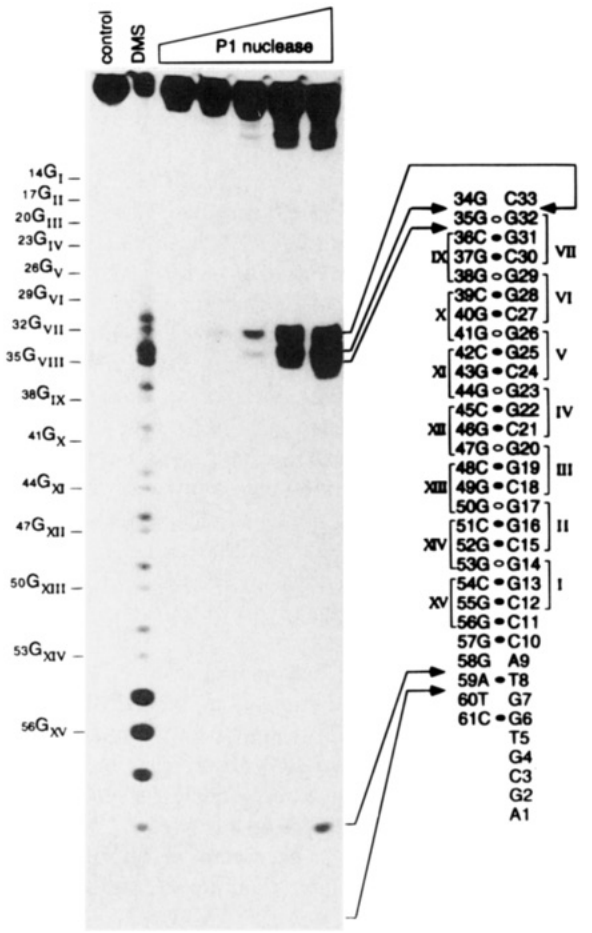


FIGURE 5: P1 nuclease digestion of ss(CGG)$_{15}$. The oligonucleotide containing (CCG)$_{15}$ purified from pCCG15 was labeled on the CGG-containing strand only with ^{32}P. The unlabeled synthetic oligonucleotide (1 μmol) of the same sequence as the labeled strand was added to the oligonucleotide, and the mixture was placed in a 100 °C H_2O bath for 3 min and immediately chilled on ice. The amounts of P1 nuclease used to digest ss(CGG)$_{15}$ (from left to right) at 37 °C in 50 mM Na^+ were 1.15×10^{-2} and 3.46×10^{-2} unit, respectively. Dimethyl sulfate (21 mM) reactions were performed as described in Materials and Methods. Roman numerals represent triplet repeat numbers. Arrows indicate sites of P1 nuclease cleavage.

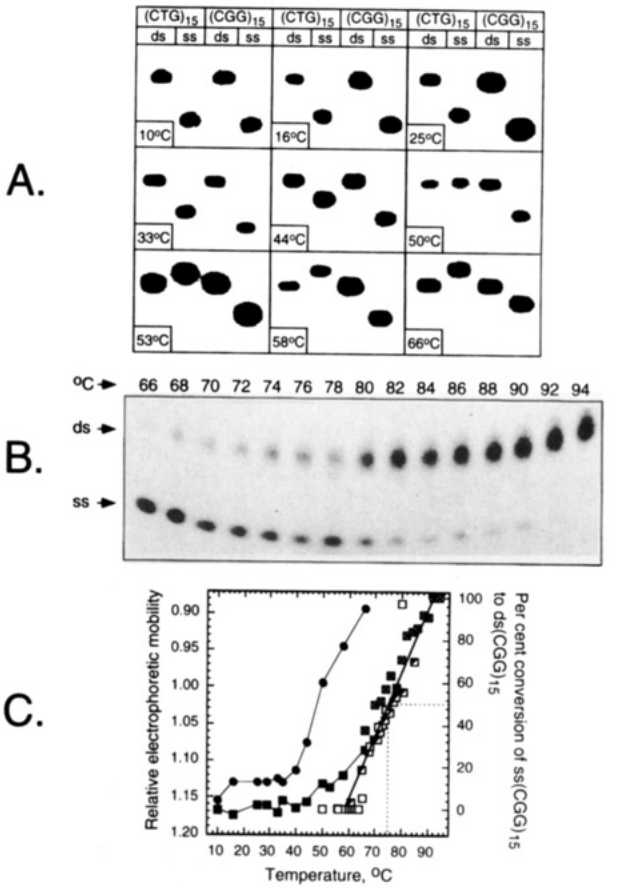


FIGURE 6: Melting profiles of ss(CGG)$_{15}$. (A) Electrophoretic mobility melting profiles (EMMPs) of (CTG)$_{15}$ and (CGG)$_{15}$ at various temperatures. The double-stranded (ds) and single-stranded (ss) forms of the DNAs were prepared as described in Materials and Methods. Electrophoretic conditions were as described in Materials and Methods. The DNA samples shown within a given square were applied to the same polyacrylamide gel. The temperature indicated in each square refers to the temperature of the polyacrylamide gel. The figure represents a total of nine separate electrophoresis experiments. (B) Annealing of ss(CGG)$_{15}$ to ss(CCG)$_{15}$. Annealing experiments were performed as described in Materials and Methods. (C) Melting profile of ss(CGG)$_{15}$; superimposition of EMMP and annealing data. The first plot is the relative electrophoretic mobility of ss(CTG)$_{15}$ (filled circles) and ss(CGG)$_{15}$ (filled squares) (from A) versus temperature. The relative mobility of ssDNA = the distance ssDNA migrated from the origin/the distance dsDNA migrated from the origin. The second plot is percent conversion of ss(CGG)$_{15}$ to ds(CGG)$_{15}$ (three separate experiments; 1, 2, and 3 are open squares, half-filled squares, and squares with an x, respectively) versus temperature. Results from experiment 3 are shown in B. Data from experiments 1 and 2 are not shown. Percent conversion = dsDNA/(ssDNA + dsDNA). Quantification of the results is described in Materials and Methods. The line through the data points from the annealing experiment was derived from a linear regression analysis using Cricket Graph (Computer Associates International, Islandia, NY). The *y* axes in the two plots were arranged such that the relative mobility of 1.027 aligned with conversion of 50% of ss(CGG)$_{15}$ to ds(CGG)$_{15}$.

suggesting a probable random arrangement of *syn*−*anti* pairing, in line with the experimental data.

During the minimization of structure 2, the nominal C•G pair in trinucleotide repeat VIII was retained. However, this interaction did not persist in the molecular dynamics simulation. Instead, G_{VIII} and G_{VII} adopted a stacked orientation (Figure 8A). In structure 4, these same two guanine bases remain in a *syn*−*anti* base pair and the whole loop region was consequently much tighter (Figure 8B). It seems reasonable to assume that the close proximity of the phosphates to one another in this loop configuration would be favored in high, but not low, salt. Thus, the differential reactivities of the loop region to DMS as a function of KCl concentration (Figure 3) could be explained on the basis of equilibrium between two loop conformations, one of which is much more stable in high KCl. The stacking of $\underline{G}_{VIII}$ onto G_{VIII} provides additional stability to alignment (*b*) (Figure 8B).

There is one further contrast between alignment (*a*) and alignment (*b*) which provides us with a possible mechanism for slippage of (*a*) to (*b*), should (*a*) be initially formed. In alignment (*a*), formation of bifurcated hydrogen bonds occurs between a hydrogen atom of a given cytosine 4-NH_2 group and guanine O6 atoms in the complementary base (G) and the guanine base 3′ to the complement (data not shown). These bifurcated hydrogen bonds form during energy minimization and occur transiently in the molecular dynamics simulation of alignment (*a*).

This structural feature could provide a route to the adoption of alignment (*b*) in the stem region, once the appropriate loop conformation has been adopted. It is apparent from Figure 8 that slippage of the 3′ side of the hairpin stem such

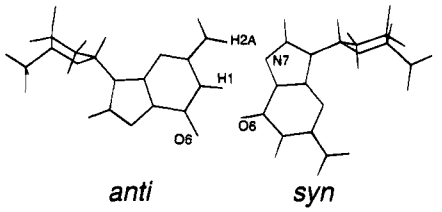


FIGURE 7: *Syn−anti* G•G base pair. The G^{syn}•G^{anti} base pair showing the N7 atom of the G^{syn} involved in hydrogen bonding and hence DMS unreactive.

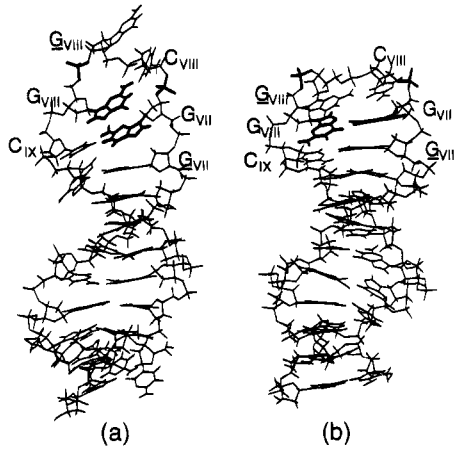


FIGURE 8: Conformations generated in molecular dynamics simulations of 5′-CC(CGG)$_7$G in the (*a*) and (*b*) alignments. Triplet numbering refers to the larger, 15-trinucleotide repeat structure (see Figure 2). The essential difference between these structures lies in the orientation of G_{VIII} and G_{VII} (bases shown in bold) which stack in the loop of alignment (*a*) but retain a *syn−anti* base pair in the loop of alignment (*b*). The phosphodiesters cleaved by P1 nuclease (Figure 5) are shown in bold.

that each base moves "down" one base pair is required for movement from the (*a*) to the (*b*) alignment. We note that this slippage mechanism is apparently independent of the organization (that is all *syn* on one strand, etc.) of the $\underline{G}^{syn}$•$\underline{G}^{anti}$ pairs in alignment (*a*), since the C•G inter-base pair hydrogen bond forms readily to both *syn* and *anti* guanines. The slippage mechanism does require further *syn−anti* rotation of the appropriate guanines to complete the conversion to the (*b*) alignment.

DISCUSSION

To determine potential structures of CGG repeats that might relate to their biological function and association with TREDs, studies were performed with ss(CGG)$_{15}$. The results presented in this paper suggest that ss(CGG)$_{15}$ formed a stable intramolecular hairpin that contained G^{syn}•G^{anti} base pairs in the stem. CGG triplets on the 5′ side of the stem preferentially base-paired with GCG triplets on the 3′ side of the stem, giving rise to an alignment referred to as (*b*). The results of chemical modification with DMS (Figures 3 and 4) suggested that a given G residue within a G•G base pair alternated between *anti* and *syn* conformations. The conclusion that G^{syn}•G^{anti} base pairs are contained within the stem region of the ss(CGG)$_{15}$ hairpin is supported by recent ^{1}H NMR studies on stable duplex structures of d(GGC)$_n$ (where n = 4, 5, or 6) (Chen et al., 1995). The conclusion that a given G residue with a G•G pair alternated between *syn* and *anti* is consistent with the rapid (~14 000 s^{-1}) transition of G^{syn}•G^{anti} base pairs in duplex DNA (Lane & Peck, 1995).

At ≥200 mM KCl, the DMS reactivities of the 5′Gs/3′Gs were ~2.0 (Table 1), indicating that ss(CGG)$_{15}$ exclusively formed a hairpin in the (*b*) alignment. At KCl concentrations ≤100 mM, the DMS reactivities of the 5′Gs/3′Gs were ≤1.82, indicating that the hairpin in alignment (*b*) (structure 4 in Figure 2) was in equilibrium with another structure. Presumably, this other structure was 2 (a hairpin in the (*a*) alignment containing G^{syn}•G^{anti} base pairs) or 3 (a hairpin in the (*b*) alignment lacking G^{syn}•G^{anti} base pairs). Owing to the stability of the G^{syn}•G^{anti} base pair, we think it is more likely that, at KCl ≤100 mM, structure 4 was in equilibrium with structure 2. Molecular dynamics simulations suggested that structure 4 contained a *syn−anti* base pair ($35G_{VIII}$ and $32G_{VII}$) in the loop region. In contrast, simulations of the loop in structure 2 predicted that $35G_{VIII}$ was stacked on top of $32G_{VII}$. We propose that, at higher salt concentrations, the $32G_{VII}$−$35G_{VIII}$ *syn−anti* pair is favored while, at lower salt concentrations, the $35G_{VIII}$−$32G_{VII}$ stack is favored. Hence, at lower salt concentrations, the $35G_{VIII}$−$32G_{VII}$ stack may tend to drive the duplex toward the (*a*) alignment.

Hairpins Formed from Repeats of CGG Are More Stable Than Alternative Tetraplex Structures. The T_m of the ss(CGG)$_{15}$ and ss(CTG)$_{15}$ hairpins, determined in low ionic strength (~1 mM Na^+) by annealing or EMMP data, were 75 and 48 °C, respectively (Figure 6). Similar T_m values, obtained by conventional UV absorbance measurements, were reported by Gacy et al. (1995) for d(CGG)$_{25}$ and d(CTG)$_{25}$ (76.1 and 52.0 °C, respectively). These results indicate that the G^{syn}•G^{anti} base pairs in d(CGG)$_n$ hairpins are substantially more stable than the T•T base pairs in ss(CTG)$_n$ hairpins. The T_m of ss(^{5m}CCG)$_{15}$ is ~8 °C higher than that of the ss(CGG)$_{15}$ hairpin (M. Mitas and J. Dill, unpublished results), indicating that the stacking energy afforded by cytosine methylation further stabilizes the hairpin. Repeats of CTG (or CAG) lack the CpG dinucleotide that is necessary for methylation by human (cytosine-5) methyltransferase (Smith, 1994; Bestor & Verdine, 1994; Kumar et al., 1993). These and other (Yu et al., 1995; M. Mitas, A. Yu, and J. Dill, unpublished data) results demonstrate that, among hairpins formed from Class I triplet repeats, those formed from CGG repeats are the most stable. We suspect that hairpins of similar stability are formed at the lagging strand of the replication fork, where expansion events are probably initiated.

The stability of the CGG hairpin may, in part, explain our failure to detect an intramolecular tetraplex structure. We modeled tetraplex structures containing (*a*) and (*b*) hairpin alignments (Figure 2) and found that structurally and energetically reasonable structures were generated from both alignments. These structures have their diagonally opposite strands parallel to each other, and antiparallel to the other diagonal pair, forming an intramolecular "head-to-tail" arrangement [for example, see Williamson (1994)]. All Watson−Crick G•C and G^{syn}•G^{anti} base pairs present in the hairpin structures are retained in the intramolecular tetraplexes, including the G•G pair that forms the "top" loop of the tetraplex. A 2:1 ratio of CGCG to G_4 quartets is present in the structures, with the CGCG quartets stabilized by hydrogen bonding between the formal G•C base pairs, as well as within them. On the basis of a simple minimized energy (excluding entropic and solvent effects), these structures seem favorable compared to the hairpins. We note that, in postulating the "triad DNA" conformation for ss(CGG)$_n$,

Kuryavyi and Jovin (1995a,b) found that simple minimization also predicted that a hairpin was less stable than a triad conformation. Our failure to detect either tetraplex or triad structures experimentally might reflect the experimental conditions. Indeed, we suspect that, under the conditions of Fry and Loeb (1994) (incubation at 4 °C for 18 h), an intermolecular tetraplex containing CGCG quartets may have formed by dimerization of two d(^{m5}CGG)$_n$ hairpins. However, such higher order conformations would probably be in equilibrium with hairpin or duplex structures (in contrast with, for example, a four-stranded tetraplex in equilibrium with four single strands), and the stability of the (*b*) alignment hairpin may strongly influence this equilibrium position for ss(CGG)$_{15}$. Hence, while high salt should favor an intermolecular tetraplex, compared to four random coil single strands, this may not necessarily be true for a single-stranded hairpin−tetraplex equilibrium.

In a previous study, Hardin et al. (1992) provided circular dichroism (CD) and ^{1}H NMR evidence that the sequence d(CGCG$_3$GCG) formed a parallel-stranded intermolecular tetraplex (different from structures 5 and 6 described in Figure 2) that contained six G_4 quartets and three C_4^{2+} quartets. We refer to this tetraplex in the subsequent text as a $(G_4)_2$−(C_4^{2+}) tetraplex. The $(G_4)_2$−(C_4^{2+}) tetraplex was stabilized best by K^+ and least by Li^+ and required a pH of ≤6.8 for protonation of the cytosines. Like d(CGCG$_3$GCG), the guanine to cytosine ratio in d(CGG)$_n$ is 2:1. Therefore, like d(CGCG$_3$GCG), ss oligonucleotides containing d(CGG)$_n$ could form inter- and possibly intramolecular tetraplexes containing $(G_4)_2$ and (C_4^{2+}) quartets. In support of this possibility, Chen (1995) has obtained CD and electrophoretic evidence that d(CGG)$_4$ is capable of forming a potassium- and acid-dependent $(G_4)_2$−$(C4^{2+})$ intermolecular tetraplex. Although we can unambiguously rule out the possibility that ss(CGG)$_{15}$ forms an intramolecular tetraplex of the $(G_4)_2$−(C_4^{2+}) type at 7.5 < pH < 8.5 (Figures 3−5), we cannot dismiss the possibility that such a tetraplex might form at pH values slightly below neutrality.

CGG Repeats within mRNAs Are Predicted To Form Stable Hairpins. Repeats of CGG are present in many genes (Riggins et al., 1992; Mitas et al., 1995; Gacy et al., 1995) and are often located in the 5′ untranslated regions. These repeats may form hairpins and play a role in gene regulation by the affecting of protein translation, splice site selection, or mRNA stability. We suspect that, of mRNA hairpins formed from Class I triplet repeats, those formed from r(CGG)$_n$ are the most stable.

Linear migration of the 40S ribosome along the 5′ untranslated region of the *FMR1* gene is impaired when the ribosome encounters more than 200 CGG repeats (Feng et al., 1995). The results presented in this paper suggest that a hairpin formed from r(CGG)$_{>200}$ would be extremely stable and would provide a formidable barrier to protein translation. A question was raised by Feng et al. (1995) whether similar impairment might occur during translation of expanded CAG repeats in genes associated with neurologic diseases (Sutherland & Richards, 1994). Electrophoretic mobility melting profiles (M. Mitas and J. Dill, unpublished results) and UV absorbance melting profiles (Gacy et al., 1995) indicate that hairpins formed from (CAG)$_n$ are far less stable than those formed from (CGG)$_n$. Thus, we predict that similar translation impairment through CAG repeats would occur only when the repeat number ≫200.

ACKNOWLEDGMENT

We thank Dr. Ulrich Melcher (Oklahoma State University) for critical review of this paper. We thank Dr. Timothy Lohman (Washington University School of Medicine) for the generous supply of *Eco*SSB. We thank the OSU Recombinant DNA/Protein resource facility for DNA sequence analysis and for use of the PDI densitometer. We thank Drs. Goutam Gupta (Los Alamos National Laboratory) and Charles Hardin (North Carolina State University) for helpful discussions.

REFERENCES

Bestor, T. H., & Verdine, G. L. (1994) *Curr. Opin. Cell Biol. 6*, 380−389.
Chen, F.-M. (1995) *J. Biol. Chem.* (in press).
Chen, X., Mariappan, S. V. S., Catasti, P., Ratliff, R., Moyzis, R. K., Laayoun, A., Smith, S. S., Bradbury, E. M., & Gupta, G. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 5199−5203.
Feng, Y., Zhang, F., Lokey, L. K., Chastain, J. L., Lakkis, L., Eberhart, D., & Warren, S. T. (1995) *Science 268*, 731−734.
Fry, M., & Loeb, L. A. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 4950−4954.
Gacy, A. M., Geoffrey, G., Juranic, N., Macura, S., & McMurray, C. T. (1995) *Cell 81*, 533−540.
Hardin, C. C., Watson, T., Corregan, M., & Bailey, C. (1992) *Biochemistry 31*, 833−841.
Jorgensen, W. L., Chandrasekhar, J., Madura, J., Impey, R. W., & Klein, M. L. (1983) *J. Chem. Phys. 79*, 926−937.
Ke, S.-H., & Wartell, R. M. (1993) *Nucleic Acids Res. 21*, 5137−5143.
Kumar, S., Cheng, X., Klimasauskas, S., Mi, S., Posfai, J., Roberts, R. J., & Wilson, G. G. (1993) *Nucleic Acids Res. 22*, 1−10.
Kuryavyi, V. V., & Jovin, T. M. (1995a) *Nat. Genet. 9*, 339−341.
Kuryavyi, V. V., & Jovin, T. M. (1995b) *J. Biomol. Struct. Dyn. 12*, a126.
Lane, A. N., & Peck, B. (1995) *Eur. J. Biochem. 230*, 1073−1087.
Leonard, G. A., et al. (1995) *Structure 3*, 335−340.
Lohman, T. M., & Ferrari, M. E. (1994) *Annu. Rev. Biochem. 63*, 527−570.
Maxam, A. M., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499−560.
Mitas, M., Yu, A., Dill, J., Kamp, T. J., Chambers, E. J., & Haworth, I. S. (1995) *Nucleic Acids Res. 23*, 1050−1059.
Pearlman, D. A., Case, D. A., Caldwell, J. C., Seibel, G. L., Singh, U. C., Weiner, P., & Kollman, P. A. (1991) *AMBER4.0*, University of California, San Francisco, CA.
Riggins, G. J., Lokey, L. K., Chastain, J. L., Leiner, H. A., Sherman, S. L., Wilkinson, K. D., & Warren, S. T. (1992) *Nat. Genet. 2*, 186−191.
Rohozinski, J., Hancock, J. M., & Keniry, M. A. (1994) *Nucleic Acids Res. 22*, 4653−4659.
Schlötterer, C., & Tautz, D. (1992) *Nucleic Acids Res. 20*, 211−215.
Sen, D., & Gilbert, W. (1991) *Curr. Opin. Struct. Biol. 1*, 435−438.
Smith, S. S. (1994) *Prog. Nucleic Acid Res. Mol. Biol. 49*, 65−111.
Sutherland, G. R., & Richards, R. I. (1994) *Am. Sci. 82*, 157−163.
van Gunsteren, W. F., & Berendsen, H. J. C. (1977) *Mol. Phys. 34*, 1311−1327.
van Gunsteren, W. F., Berendsen, H. J., Guersten, R. G., & Zwinderman, H. R. (1986) *Ann. N. Y. Acad. Sci. 482*, 287−303.
Wartell, R. M., Hosseini, H. G., & Moran, C. P., Jr. (1990) *Nucleic Acids Res. 18*, 2699−2705.
Williamson, J. R. (1993) *Curr. Opin. Struct. Biol. 3*, 357−362.
Williamson, J. R. (1994) *Annu. Rev. Biophys. Biomol. Struct. 23*, 703−730.
Williamson, J. R., Raghuraman, M. K., & Cech, T. R. (1989) *Cell 59*, 871−880.
Wohlrab, F. (1992) *Methods Enzymol. 212* (Part B), 294−301.
Yu, A., Dill, J., Wirth, S. S., Huang, G., Lee, V. H., Haworth, I. S., & Mitas, M. (1995) *Nucleic Acids Res. 23*, 2706−2714.

BI951148B