

The trinucleotide repeat sequence d(GTC)₁₅ adopts a hairpin conformation

Adong Yu, Jeffrey Dill, Sara S. Wirth, George Huang¹, Vincent H. Lee², Ian S. Haworth^{1,2} and Michael Mitas*

Department of Biochemistry and Molecular Biology, Oklahoma State University, 246 Noble Research Center, Stillwater, OK 74078, USA, ¹Department of Pharmaceutical Sciences and ²Department of Biochemistry and Molecular Biology, University of Southern California, 1985 Zonal Avenue, Los Angeles, CA 90033, USA

Received March 23, 1995; Revised and Accepted June 7, 1995

ABSTRACT

The structure of a single-stranded (ss) oligonucleotide containing (GTC)₁₅ [ss(GTC)₁₅] was examined. As a control, parallel studies were performed with ss(CTG)₁₅, an oligonucleotide that forms a hairpin. Electrophoretic mobility, KMnO₄ oxidation and P1 nuclease studies demonstrate that, similar to ss(CTG)₁₅, ss(GTC)₁₅ forms a hairpin containing base paired and/or stacked thymines in the stem. Electrophoretic mobility melting profiles performed in ~1 mM Na⁺ revealed that the melting temperatures of ss(GTC)₁₅ and ss(CTG)₁₅ were 38 and 48°C respectively. The loop regions of ss(GTC)₁₅ and ss(CTG)₁₅ were cleaved by single-strand-specific P1 nuclease at the T25–C29 and G26–C27 phosphodiester bonds respectively (where the loop apex of the DNAs is T28). Molecular dynamics simulations suggested that in ss(GTC)₁₅ the loop was bent towards the major groove of the stem, apparently causing an increased exposure of the T25–C29 region to solvent. In ss(CTG)₁₅ guanine–guanine stacking caused a separation of the G26 and C27 bases, resulting in exposure of the intervening phosphodiester to solvent. The results suggest that ss(GTC)₁₅ and ss(CTG)₁₅ form similar, but distinguishable, hairpin structures.

INTRODUCTION

To aid in correlating potential structures and functions of triplet DNA repeats, we described a sequence-based classification system for them (1). 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 (2) and were uniquely associated with triplet repeat expansion diseases (TREDs). All six complementary single strands of Class I triplet repeats potentially formed stable hairpin structures. In support of this possibility we presented evidence that a single-stranded (ss) oligonucleotide containing 15 prototypic CTG repeats [ss(CTG)₁₅] formed a hairpin containing base paired and/or base stacked thymines in the stem (1).

Here we report the results of studies conducted on the second member of ss Class I triplet repeats: ss(GTC)₁₅. The results presented below demonstrate that, apart from modest differences in thermal stabilities, ss(GTC)₁₅ forms a hairpin similar to ss(CTG)₁₅.

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, UK). Sequences of oligonucleotides were: (GTC)₁₅, GATCC(GTC)₁₅GGTACCA; (GAC)₁₅, AGCTTGGTACC(CAG)₁₅G.

Plasmid DNA preparation

Preparation of plasmid pCTG15 has been described elsewhere (1). Plasmids bearing (GTC)₁₅ were prepared in a manner similar to that described for pCTG15.

DNA sequencing

Cesium chloride-purified plasmid DNAs containing GTC₁₅ were linearized with *Xho*I. For sequence analysis 8 pmol of a synthetic oligonucleotide complementary to a portion of the thymidine kinase promoter of pBLCAT2 (sequence GTTCGAATTCGC-CAATGACAA) was added to 1 µg linearized plasmid DNA. DNA sequence analysis was performed by the OSU Protein DNA Resource Facility on an ABI 373 automated sequencer. The triplet region within two plasmid DNAs were sequenced. One plasmid contained 15 perfect CTG repeats and was named pCTG15. Owing to the deletion of a thymine in the second triplet repeat the other plasmid was named pΔT_{II}CTG15.

KMnO₄ oxidation

KMnO₄ oxidation of ss(GTC)₁₅, ΔT_{II}GTC₁₅ and ss(CTG)₁₅ were as previously described (1). Briefly, the pyrimidine-rich triplet repeat strand was liberated from pGTC₁₅, pΔT_{II}GTC₁₅ or pCTG15 by digestion with *Hind*III followed by *Bam*HI. Oligo-

* To whom correspondence should be addressed

nucleotides were purified from the plasmid by agarose gel electrophoresis. DNAs were incubated with [³²P]ATP and polynucleotide kinase after digestion with *Hind*III. Unlabeled synthetic oligonucleotide (~1.4 pmol) of the same sequence as the labeled strand was added to 4×10^3 d.p.m. (~0.7 fmol) of 5'-end-labeled DNA. The DNAs were placed in a boiling water bath for 5 min and then cooled at room temperature for 5 min. KMnO₄ oxidation reactions were performed with the DNAs according to the method of McCarthy and Rich (3). For size markers dimethylsulfate (DMS) reactions were performed at 21 mM with the DNAs, essentially according to the method of Maxam and Gilbert (4). Electrophoresis was performed in a Hoeffer (San Francisco, CA) SE 600 series electrophoresis unit in 45 mM Tris–borate, 1 mM EDTA (TBE). The gel contained 20% polyacrylamide and 8 M urea. Tap water at 55°C was circulated through the unit during electrophoresis.

Electrophoretic mobility shift analysis

Fifteen micrograms of pGTC15 or pCTG15 were digested with 50 U *Hind*III for 1 h at 37°C in a volume of 70 µl. Recessed ends were labeled at the 3'-termini by adding 5 µl [α -³²P]dCTP, 5 µl [α -³²P]dATP (each 3000 Ci/mmol; ICN, Irvine, CA), 2.5 µl 5 mM dTTP and dGTP and 25 U Klenow enzyme (New England Biolabs, Beverly, MA). Reactions were incubated for 1 h at room temperature. Plasmid DNAs were extracted with 25:24:1 phenol:chloroform:isoamyl alcohol (PCI) and precipitated with ethanol. Resuspended DNAs were digested with 50 U *Bam*HI in a volume of 70 µl and applied directly to a Nucletrap column (Stratagene, La Jolla, CA) for further removal of unincorporated ³²P-dNTPs. Labeled oligonucleotides were not purified from labeled vector. Probes were diluted to 2×10^4 d.p.m./µl with H₂O. Oligonucleotides liberated from the plasmid at 0.7 nM were incubated with various amounts of *Escherichia coli* single-stranded DNA binding protein (5) (*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, 1 mM dithiothreitol at 37°C for 20 min. For electrophoretic analysis DNAs (4×10^4 d.p.m.) were diluted to 10 µl in buffer containing 8% glycerol, 10 mM HEPES, pH 8.5, 1 mM EDTA. One microliter of loading dye (50% glycerol, 0.4% bromophenol blue) was added to the DNA samples prior to gel electrophoresis. Electrophoresis was performed in a Hoeffer SE 600 series electrophoresis unit in TBE. Polyacrylamide concentration of the gel was 8%.

Electrophoretic mobility melting profiles

Oligonucleotides containing (GTC)₁₅ or (CTG)₁₅ were prepared as described for the electrophoretic mobility shift analysis. Even though oligonucleotides were cleaved from the plasmid, a significant amount of ssDNA was usually present, as determined by native polyacrylamide gel electrophoresis. To obtain a homogeneous population of labeled dsDNAs, oligonucleotides liberated from the plasmid at 0.7 nM were incubated with 0.5 nM *Eco*SSB in 8% glycerol, 0.2 M NaCl, 18 mM HEPES, pH 7.40, 1 mM EDTA, 1 mM dithiothreitol at 37°C for 20 min. To obtain a homogeneous population of labeled ss DNAs 1 µM unlabeled synthetic oligonucleotide of the same sequence as the labeled strand was added, placed in a 90°C water bath for 5 min and then placed at 25°C for 5 min. For electrophoretic analysis DNAs (4×10^4 d.p.m.) were diluted to 10 µl in buffer containing 8%

glycerol, 10 mM HEPES, pH 8.5, 1 mM EDTA. One microliter of loading dye (50% glycerol, 0.4% bromophenol blue) was added to the DNA samples prior to gel electrophoresis. Electrophoresis was performed in a Hoeffer SE 600 series electrophoresis unit in TBE. Polyacrylamide concentration of the gel was 8%. Temperature of the gel was determined by a Fisher Scientific thermometer equipped with a Type K beaded probe (catalogue no. 15-077-10). The temperature probe was inserted through the top sealing gasket and 4 cm into the glass plates prior to casting the polyacrylamide gel. While casting the gel the probe was positioned so that the lead wire inserted into lane number 14 (of 15). For a typical electrophoresis experiment the chambers were pre-chilled or pre-heated for 2 h and pre-run for 2 h at 28 mA/gel (2 mA/cm) at the appropriate temperature. DNA samples were applied and subjected to electrophoresis for 150 min at the appropriate temperature. The temperature of the gel was monitored throughout electrophoresis and usually did not vary by >1°C. To arrive at a gel temperatures of 10–16°C electrophoresis was performed in a 6°C room. A H₂O/Xerex antifreeze mixture (1:1) was circulated through the electrophoresis unit at 1–10°C. Temperature of the H₂O/antifreeze mixture was controlled with a Brinkman RC6 temperature bath. To arrive at gel temperatures between 25 and 66°C the H₂O/antifreeze mixture was circulated through the electrophoresis unit at 20–72°C. To avoid loss of heat at higher temperatures the electrophoresis unit was encased in a styrofoam container.

P1 nuclease digestion

Oligonucleotides containing (GTC)₁₅ or (CTG)₁₅ were labeled with the use of Klenow enzyme as described for the electrophoretic mobility melting profiles. Labeled oligonucleotides were separated from labeled vector by electrophoresis in a 2% agarose gel. Oligonucleotides containing triplet repeats were excised from gels 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 d.p.m. (0.7 fmol) 3'-end-labeled DNA and placed in a boiling water bath for 3 min and then placed on ice for 5 min. P1 nuclease digestions were performed in 50 mM Tris–HCl, pH 7.5, 10 mM MgCl₂, 50 mM NaCl at 37°C essentially according to the method of Wohlrab (6) as previously described (1).

Molecular dynamics simulations

Preliminary gas phase energy minimization was used to provide input structures for molecular dynamics simulations that incorporated explicit water molecules. All calculations were performed using the AMBER4.0.1 all-atom force field program (7,8) running on an Indigo Workstation. Full-length hairpins of ss(CTG)₁₅ and ss(GTC)₁₅ were first constructed assuming a fully base paired stem (with T·T mismatches) and a single base (thymine) loop. The stem initially had a canonical B-DNA conformation built using the QUANTA4.0 package (Molecular Simulations, Burlington, MA). The single base loop was manually placed and then minimized into an appropriate position. Following this, the entire hairpin was subjected to 4000 cycles of energy minimization. The resulting structure was cleaved at the G_V–C_{XI} base pair [ss(CTG)₁₅] and the C_V–G_{XI} base pair [ss(GTC)₁₅] (see Figs 8 and 9 for base numbering). The resulting structures (nominally containing six G·C base pairs, two T·T pairs and the thymine loop in each case) were solvated in a C5 sphere

of TIP3P water molecules (9), with the center of mass of the sphere centered upon the mass center of the loop region (designated as the two G-C pairs and the thymine loop). The water sphere fully solvated the loop region and extended to the terminal pair of the stem. The final two base pairs of the stem were restrained to their minimized positions in the simulation, with force constants of 2 kcal/mol/C5. Sixteen sodium counterions were added at the points of most negative potential (10). The molecular dynamics simulation was performed for 100 ps at 298 K (following a brief period in which the temperature was increased from 0 K) for each hairpin, using a step length of 0.002 ps and SHAKE bond length constraint (11). Coordinates were saved every 0.4 ps.

RESULTS

To analyze DNA containing GTC repeats, a ds oligonucleotide containing 15 GTC triplet repeats was cloned into a plasmid as described in Materials and Methods. Oligonucleotides liberated from the plasmid were utilized for studies, since they were unequivocally full-length. Non-repetitive sequences were also included in the termini of the oligonucleotides to help prevent 'slippage' of the hairpin structure and to provide restriction sites for release of the oligonucleotide from the plasmids.

To determine whether ss(GTC)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(GTC)₁₅ was investigated by performing electrophoretic studies with labeled ss(GTC)₁₅ mixed with various amounts of unlabeled ss synthetic oligonucleotide containing ss(GTC)₁₅ (Fig. 1A). If ss(GTC)₁₅ formed a stable intramolecular hairpin structure, increasing the concentration of unlabeled synthetic ss(GTC)₁₅ should not result in the appearance of a form of DNA that migrated with slow electrophoretic mobility.

In the absence of added unlabeled ss synthetic oligonucleotide a predominant species of DNA with a relatively fast electrophoretic mobility was observed, corresponding to ssDNA (Fig. 1A). Addition of a 10⁵-fold molar excess (final DNA concentration 7 μM) of unlabeled ss synthetic oligonucleotide of the same sequence as ss(GTC)₁₅ did not result in the formation of a slow migrating complex (as anticipated), indicating that ss(GTC)₁₅ formed a stable unimolecular structure.

To demonstrate that the unlabeled ss synthetic oligonucleotide containing (GTC)₁₅ was not degraded and contained GTC repetitive sequences, a control experiment was performed with labeled ss(GAC)₁₅ (Fig. 1B). Addition of increasing amounts of unlabeled ss synthetic oligonucleotide containing (GTC)₁₅ to the labeled ss(GAC)₁₅ probe resulted in complete conversion of the fast migrating ss form to the slow migrating ds form.

KMnO₄ oxidizes T3 and T_{VIII} of ss(GTC)₁₅

The results shown in Figure 1 indicated that ss(GTC)₁₅ formed a stable unimolecular structure. Similar results were obtained with ss(CTG)₁₅, which we have shown forms a hairpin (1). The predicted hairpin structure of ss(GTC)₁₅ is shown in Figure 2A. The 15 sets of triplet repeats within ss(GTC)₁₅ are identified in Figure 2A and in the subsequent text by roman numerals. For clarity, nucleotides not part of the triplet repeat region are

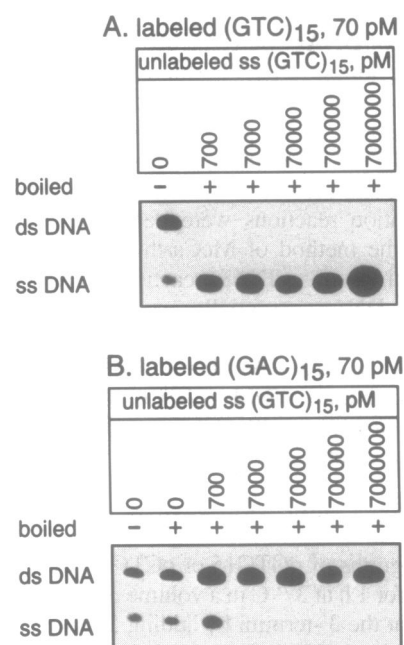


Figure 1. The structure of ss(GTC)₁₅ is concentration-independent. A ds oligonucleotide containing (GTC)₁₅ was excised from plasmid pGTC15 as described in Materials and Methods. Strands labeled with ³²P were (A) ss(GTC)₁₅ and (B) ss(GAC)₁₅. Prior to gel electrophoresis, unlabeled synthetic oligonucleotide containing (GTC)₁₅ at the above concentration was added to the indicated labeled strand and placed in a boiling water bath for 10 min and then cooled at room temperature for 15 min. DNA samples were applied to a native 8% polyacrylamide gel.

identified by their one letter symbol followed by the respective arabic numeral. Nucleotides forming part of the triplet repeat region are identified by their one letter symbol followed by their respective triplet number in subscript roman. For example, T_{VIII} is the nucleotide at the apex in the presumed loop of ss(GTC)₁₅.

To investigate the possibility of a hairpin conformation of ss(GTC)₁₅, experiments were performed with KMnO₄, which preferentially oxidizes unpaired or unstacked thymines, resulting in strand cleavage upon subsequent treatment with piperidine (12,13). Treatment of ss(GTC)₁₅ with KMnO₄/piperidine at 30°C resulted in cleavage at T3 and T_{VIII} (Fig. 2A, left side), a result consistent with a hairpin structure of ss(GTC)₁₅. Minor cleavage products were also observed at T_{VII} and T_{IX}. The results suggest that ss(GTC)₁₅, like ss(CTG)₁₅, forms a hairpin containing base paired and/or base stacked thymines in the stem.

To further investigate the possibility of a hairpin conformation of ss(GTC)₁₅, KMnO₄ oxidation experiments were performed with a ss oligonucleotide of the same sequence, but lacking a single thymine within triplet II [the sequence referred to as ΔT_{II}(GTC)₁₅]. The predicted structure of ΔT_{II}(GTC)₁₅ is shown on the right in Figure 2B. Due to the deletion in ΔT_{II}(GTC)₁₅, the thymine in triplet XIV does not have a nucleotide with which it can pair. Therefore, if ΔT_{II}(GTC)₁₅ formed the hairpin as shown in Figure 2B, the thymine in triplet XIV, as well as the thymine in triplet VIII (the loop apex), should be sensitive to KMnO₄ oxidation. The KMnO₄ oxidation results (Fig. 2B) are in agreement with those expected; the thymines in triplets VIII and

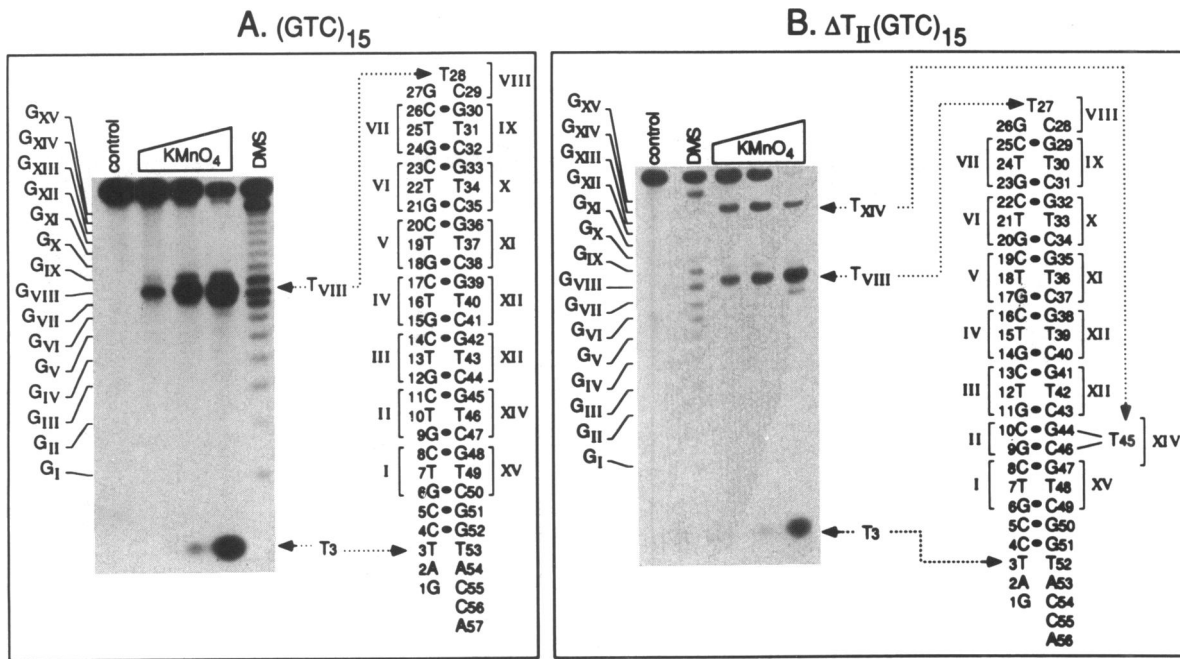


Figure 2. KMnO_4 oxidation of $\text{ss}(\text{GTC})_{15}$ and $\text{ss} \Delta\text{T}_{\text{II}}(\text{GTC})_{15}$. KMnO_4 oxidations were performed at 30°C in 50 mM Na^+ with (A) $\text{ss}(\text{GTC})_{15}$ and (B) $\Delta\text{T}_{\text{II}}(\text{GTC})_{15}$ as described in Materials and Methods. From left to right the amounts of KMnO_4 in the reaction mixtures were 0, 0.13, 0.25 and 0.75 mM. The marker lane contained DNAs reacted with DMS. The positions of the G residues in the triplet repeats are indicated to the left in each figure. The deduced hairpin structures of the DNA sequences are shown to the right in the figures. Roman numerals depict a particular triplet repeat. Conventional arabic numerals depict the position of a nucleotide with respect to the 5'-end. KMnO_4 oxidation sites of particular thymine residues are indicated by arrows.

XIV were preferentially oxidized, providing further evidence for a hairpin conformation.

KMnO_4 oxidation at various temperatures reveals reduced thermal stability of the $(\text{GTC})_{15}$ hairpin

To further investigate the possibility of a hairpin conformation of $\text{ss}(\text{GTC})_{15}$, KMnO_4 oxidation experiments were performed in 50 mM Na^+ with and without 150 mM KCl at 40 , 50 , 60 and 70°C . For comparison, parallel experiments were performed with $\text{ss}(\text{CTG})_{15}$. In the presence or absence of KCl the thymines in the stems of $\text{ss}(\text{GTC})_{15}$ and $\text{ss}(\text{CTG})_{15}$ were not oxidized by KMnO_4 at 40°C (data not shown). In the absence of added KCl significant oxidation of the thymines in the presumed stem region of $\text{ss}(\text{GTC})_{15}$ occurred at 50°C (Fig. 3A). In the presence of KCl significant oxidation required an incubation temperature of 60 – 70°C (Figs 4B and 5B). Stabilization of the T–T interactions by added KCl is consistent with a hairpin structure of $\text{ss}(\text{GTC})_{15}$.

In the absence of KCl significant oxidation of the thymines in the stem region of $\text{ss}(\text{CTG})_{15}$ required an incubation temperature of 60°C (Fig. 4A). This result indicated that the T–T interactions in $\text{ss}(\text{CTG})_{15}$ were more stable than the T–T interactions in $\text{ss}(\text{GTC})_{15}$. In the presence of KCl significant oxidation of the thymines in the stem of $\text{ss}(\text{CTG})_{15}$ required an incubation temperature of 70°C (Fig. 5B).

Electrophoretic mobility melting profiles of $\text{ss}(\text{GTC})_{15}$ and $\text{ss}(\text{CTG})_{15}$

The results from Figure 1 showed that $\text{ss}(\text{GTC})_{15}$ migrated with rapid mobility at 16°C . The rapid electrophoretic mobility of

$\text{ss}(\text{GTC})_{15}$ 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 test this hypothesis, polyacrylamide gel electrophoresis was performed with $(\text{GTC})_{15}$ and $(\text{CTG})_{15}$ at a range of temperatures. To obtain homogeneous preparations of dsDNAs, labeled ssDNAs and an equimolar amount of their unlabeled complementary sequences were incubated with 0.5 nM EcoSSB . At this concentration *EcoSSB* anneals $\text{ss}(\text{GTC})_{15}$ to its complementary sequence (Fig. 6) and $\text{ss}(\text{CTG})_{15}$ to its complementary sequence (data not shown).

In agreement with the expected results, a decrease in the relative electrophoretic mobility of $\text{ss}(\text{GTC})_{15}$ was observed at higher temperatures (Fig. 7). The relative electrophoretic mobility of $\text{ss}(\text{GTC})_{15}$ decreased at a sharp rate between 33 and 45°C . In contrast, the relative electrophoretic mobility of $\text{ss}(\text{CTG})_{15}$ decreased at a sharp rate between 41 and 55°C . At 55°C the relative electrophoretic mobility of $\text{ss}(\text{GTC})_{15}$ was 0.890 , the lowest value observed. Between 10 and 55°C the electrophoretic mobility of $\text{ss}(\text{GTC})_{15}$ decreased 22.3% . Half maximal electrophoretic mobility of $\text{ss}(\text{GTC})_{15}$ was 1.027 , corresponding to a temperature of 38°C . By analogy to data derived from a UV absorbance melting profile, we refer to the temperature corresponding to half maximal electrophoretic mobility as the melting temperature (T_m) of the DNA. Note that the isomobility temperature, or T_i [defined as the temperature at which the electrophoretic mobilities of $\text{ss}(\text{GTC})_{15}$ and $\text{ds}(\text{GTC})_{15}$ were identical] was 39°C , a temperature comparable to the T_m of $\text{ss}(\text{GTC})_{15}$. The temperature corresponding to an electrophoretic mobility of 1.027 for $\text{ss}(\text{CTG})_{15}$ was 48°C . This result indicates

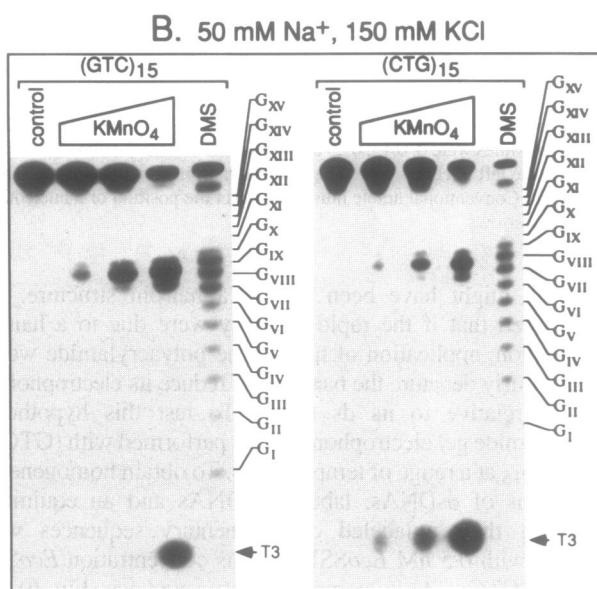
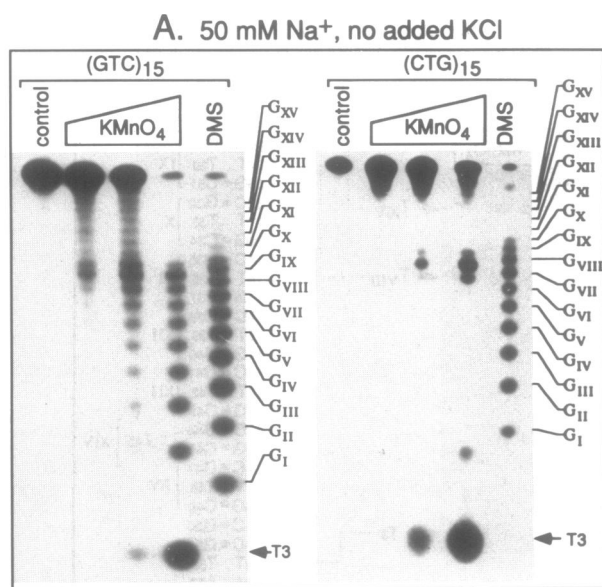


Figure 3. KMnO_4 oxidation of ss(GTC)₁₅ and ss(CTG)₁₅ at 50°C. KMnO_4 oxidations were performed with ss(GTC)₁₅ or ss(CTG)₁₅ as described in Materials and Methods at 40°C in (A) 50 mM Na^+ or (B) 50 mM Na^+ , 150 mM KCl. From left to right the amount of KMnO_4 in the reaction mixtures were 0, 0.13, 0.25 and 0.75 mM. The marker lane contained DNAs reacted with DMS. The positions of the G residues within a triplet repeat are indicated to the right of each autoradiograph. The signal at the bottom of each autoradiograph corresponds to T3, a nucleotide not part of the triplet repeat region.

that the T_m of ss(GTC)₁₅ was 10°C lower than the T_m of ss(CTG)₁₅.

P1 nuclease digestions of ss(GTC)₁₅ and ss(CTG)₁₅

To further compare the structures of ss(GTC)₁₅ and ss(CTG)₁₅, each was labeled near the 3'-terminus with Klenow enzyme and reacted with single-strand-specific P1 nuclease. Reaction products were applied to a 20% polyacrylamide–8 M urea sequencing gel. To generate size markers, the DNAs were separately reacted with

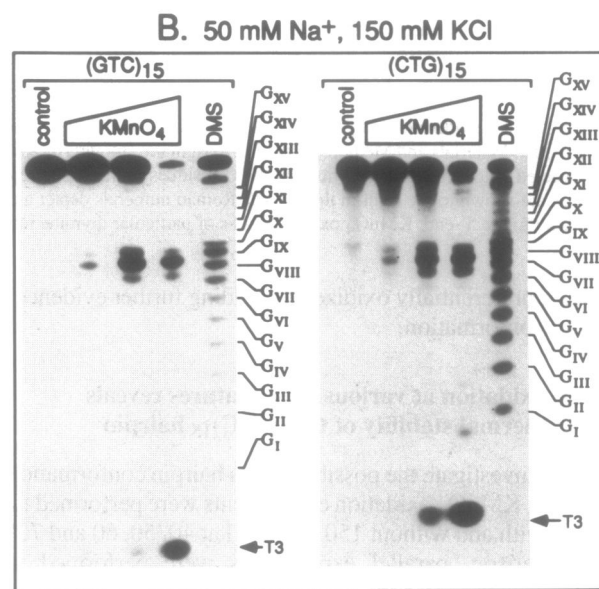
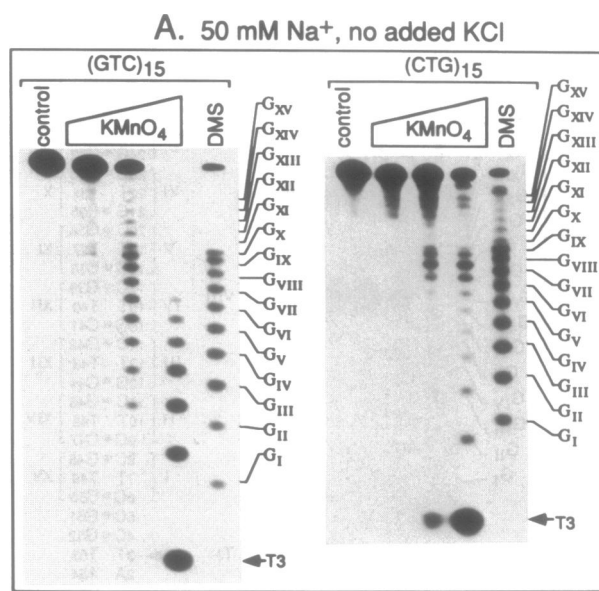


Figure 4. KMnO_4 oxidation of ss(GTC)₁₅ and ss(CTG)₁₅ at 60°C. KMnO_4 oxidations were performed at 50°C with ss(GTC)₁₅ or ss(CTG)₁₅ as described in the legend to Figure 3. (A) 50 mM Na^+ . (B) 50 mM Na^+ , 150 mM KCl.

DMS or KMnO_4 and cleaved with piperidine. In comparing the size of the P1 products to the markers it is important to note that piperidine completely removes the 5' base oxidized by KMnO_4 or methylated by DMS. Hence, the signals in the marker lanes do not contain a thymine or guanine respectively at the 5'-terminus.

Incubation of ss(GTC)₁₅ with P1 nuclease did not result in cleavage of the phosphodiester bond between C_{VIII} and C₅₅, indicating that these nucleotides participated in base pairing and/or base stacking interactions (Fig. 8). At the 3'-terminus the C₅₅–C₅₆ phosphodiester bond was extensively degraded. In the vicinity of the loop the major product of P1 nuclease digestion co-migrated with the KMnO_4 product generated from cleavage of T_{VII}. Therefore, the major site of P1 cleavage within the triplet repeat region was the T_{VII}–C_{VII} phosphodiester bond. Minor sites

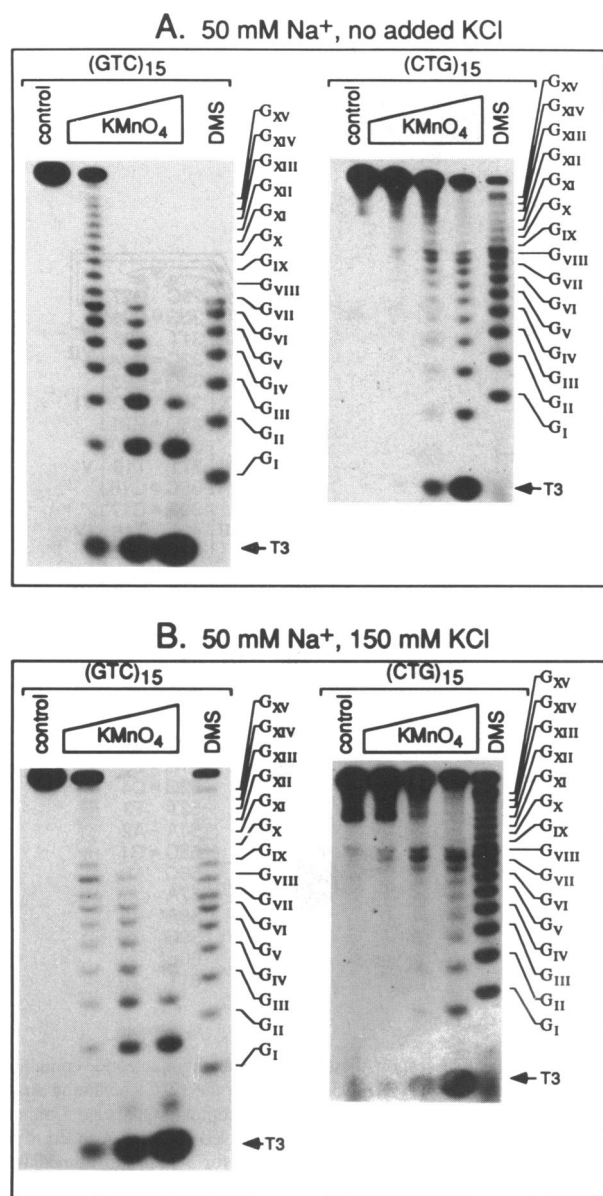


Figure 5. KMnO_4 oxidation of $\text{ss}(\text{GTC})_{15}$ and $\text{ss}(\text{CTG})_{15}$ at 70°C . KMnO_4 oxidations were performed at 60°C with $\text{ss}(\text{GTC})_{15}$ or $\text{ss}(\text{CTG})_{15}$ as described in the legend to Figure 3. (A) 50 mM Na^+ . (B) 50 mM Na^+ , 150 mM KCl.

of P1 cleavage in the loop region included the $\text{C}_{\text{VII}}-\text{G}_{\text{VIII}}$, $\text{G}_{\text{VIII}}-\text{T}_{\text{VIII}}$ and $\text{T}_{\text{VIII}}-\text{C}_{\text{VIII}}$ phosphodiester bonds. Surprisingly, the $\text{T}_{\text{VIII}}-\text{C}_{\text{VIII}}$ phosphodiester bond was not the major site of P1 cleavage. This result was surprising since T_{VIII} was readily oxidized by KMnO_4 . The structure of $\text{ss}(\text{GTC})_{15}$ most consistent with these results is shown in Figure 8. Triplet I is base paired to triplet XV, triplet II to triplet XIV, etc. A hairpin whereby triplet I is base paired to triplet XIV, triplet II to triplet XIII, etc. is inconsistent with the results.

The $\text{T}_{53}-\text{A}_{54}$ and $\text{A}_{54}-\text{C}_{55}$ phosphodiester bonds of $\text{ss}(\text{GTC})_{15}$ were not hydrolyzed by P1 nuclease. This result was not expected, since, as shown in the experiments above, T3 was readily oxidized by KMnO_4 . In the hairpin structure of $\text{ss}(\text{GTC})_{15}$ T53 is opposite T3. Hence, since T3 is not base paired to T53, the

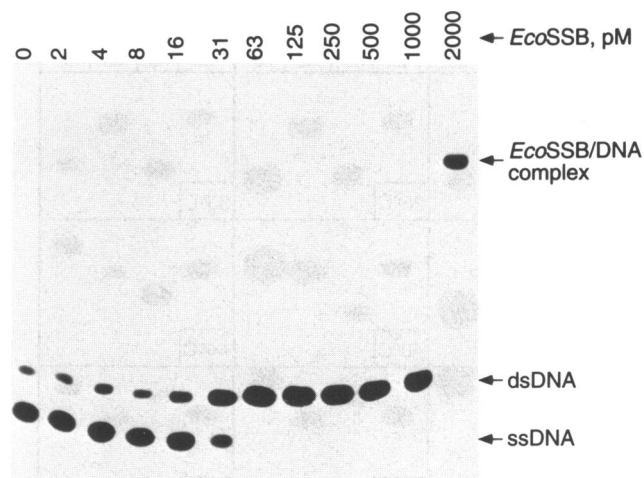


Figure 6. *EcoSSB* anneals labeled $\text{ss}(\text{GTC})_{15}$ to its unlabeled complementary sequence. Oligonucleotide liberated from pGTC_{15} at 0.7 nM was incubated with varying amounts of *EcoSSB* in 8% glycerol, 0.2 M NaCl, 18 mM HEPES, pH 7.40, 1 mM EDTA, 1 mM dithiothreitol and varying amounts of protein, as indicated. The final volume of the reaction mixture was 10 μl . The reaction mixtures were incubated at 37°C for 20 min. One microliter of loading dye (50% glycerol, 0.4% bromophenol blue) was added to the DNA samples prior to gel electrophoresis. During electrophoresis 15°C tap water was circulated through a Hoefer SE 600 series unit (San Francisco, CA). Electrophoresis was performed in $1\times$ TBE buffer (45 mM Tris-borate, 1 mM EDTA) through a 6% polyacrylamide gel at 2 mA/cm.

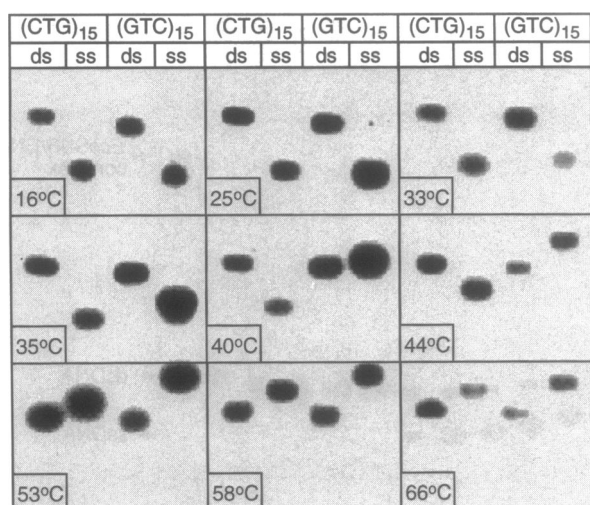
phosphodiester bonds of T53 should be cleaved by P1. The inability of P1 to cleave the phosphodiester bonds of T53 suggests that base stacking interactions and not base pairing interactions inhibit cleavage. The base stacking interactions may arise from neighboring C-G base pairs that form between C_4-G_{52} and G_1-C_{55} (as shown in Fig. 8).

Similar to $\text{ss}(\text{GTC})_{15}$, incubation of $\text{ss}(\text{CTG})_{15}$ with P1 nuclease did not result in cleavage of phosphodiester bonds between the 3' portion of the loop and C_{55} (Fig. 9). However, in contrast to multiple sites of P1 cleavage in the loop of $\text{ss}(\text{GTC})_{15}$, a single digestion product in the loop of $\text{ss}(\text{CTG})_{15}$ was observed. The product co-migrated with the DMS product generated from cleavage of G_{VII} (the G_{VII} band is faint, but clearly detectable on the autoradiograph). Therefore, the single site of cleavage within the triplet repeat region was the $\text{G}_{\text{VII}}-\text{C}_{\text{VIII}}$ phosphodiester bond. These results suggest an unexpected asymmetric structure of the loop region, as described in further detail below.

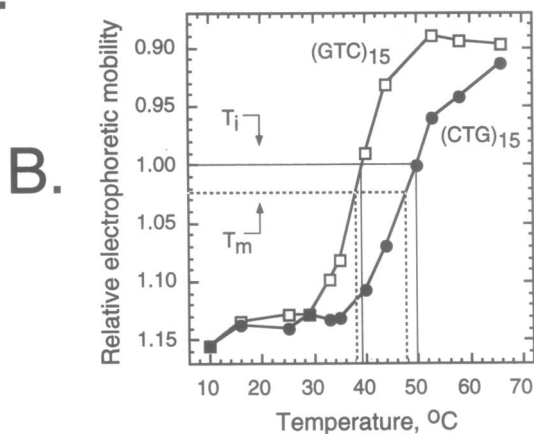
The results of nuclease cleavage of $\text{ss}(\text{CTG})_{15}$ and $\text{ss}(\text{GTC})_{15}$ are consistent with hairpin structures and suggest extensive base pairing and/or base stacking interactions between the T-T mismatches.

Molecular dynamics simulations

To provide a structural basis for the unexpected P1 nuclease digestion results, we examined potential loop structures in the $\text{ss}(\text{CTG})_{15}$ and $\text{ss}(\text{GTC})_{15}$ hairpins using molecular dynamics simulations. The starting conformation for these calculations was a fragment of an energy minimized structure which in turn started from a 15 triplet repeat hairpin having a fully base paired stem (including T-T mismatches) and a single base thymine loop [as shown in Figs 8 and 9 for $\text{ss}(\text{GTC})_{15}$ and $\text{ss}(\text{CTG})_{15}$ respectively]. During the energy minimization of $\text{ss}(\text{CTG})_{15}$ the $\text{C}_{\text{VIII}}-\text{G}_{\text{VIII}}$



A.



B.

Figure 7. Electrophoretic mobility melting profiles of ss(GTC)₁₅ and ss(CTG)₁₅. (A) Electrophoretic analysis of (GTC)₁₅ and (CTG)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) Relative mobilities of the ss DNAs plotted as a function of temperature. Relative mobility of ssDNA = distance ssDNA migrated from origin/distance dsDNA migrated from origin.

and C_{IX}-G_{VII} base pairs (those closest to the single base loop) were essentially retained. For ss(GTC)₁₅ the base pairing of G_{VIII}-C_{VIII} was compromised in the minimization.

The minimized hairpins were cleaved below the N20-N36 base pair and the resulting fragment was fully solvated. During the subsequent 100 ps of molecular dynamics considerable changes occurred in the loop conformations of both hairpins. Conformations exemplified by those shown in Figure 10 developed within ~50 ps and were essentially retained over the remainder of the trajectory. In each case the nominal GC pairing within trinucleotide repeat VIII was lost and replaced by a stacking interaction between G_{VII} and G_{VIII} for ss(CTG)₁₅ and between G_{VIII} and G_{IX} for ss(GTC)₁₅. In each case C_{VIII} was displaced and formed a weak stacking interaction with T_{VIII}.

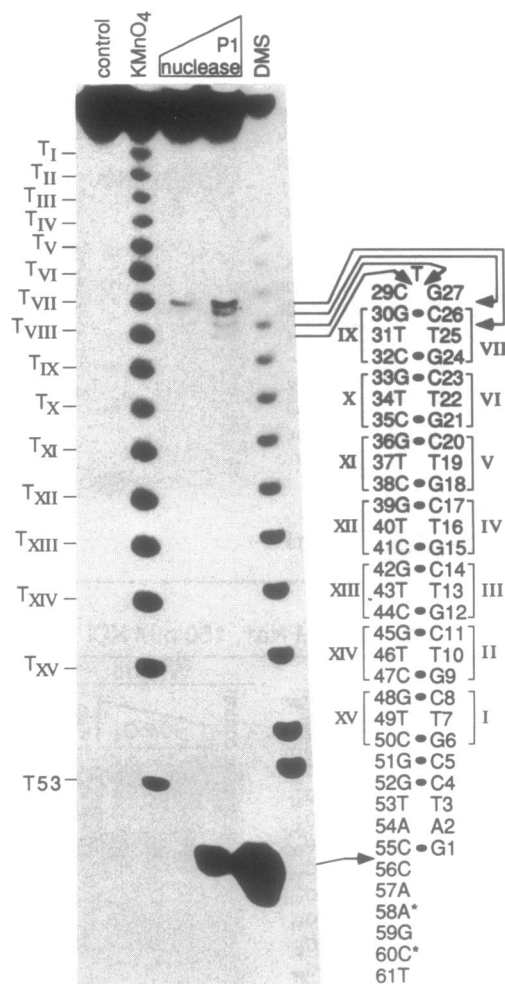


Figure 8. P1 nuclease digestion of ss(GTC)₁₅. The oligonucleotide containing (GTC)₁₅ purified from pGTC15 was labeled on the GTC-containing strand only with Klenow enzyme and ³²P. One micromolar unlabeled synthetic oligonucleotide of the same sequence as the labeled strand was added to the oligonucleotide, placed in a 100°C water bath for 3 min and immediately chilled on ice. The amounts of P1 nuclease used to digest ss(GTC)₁₅ (from left to right) at 37°C in 50 mM Na⁺ were 1.15 × 10⁻² and 3.46 × 10⁻² U respectively. KMnO₄ (250 μM) oxidation of ss(GTC)₁₅ was performed as described in Materials and Methods at 50°C. DMS (21 mM) reactions were performed as described in Materials and Methods. Roman numerals represent triplet repeat numbers. The nucleotides labeled with ³²P are marked with an asterisk. Arrows indicate sites of P1 nuclease cleavage.

In the ss(CTG)₁₅ hairpin loop (Fig. 10A) the stacking of G_{VIII} onto G_{VII} essentially 'intercalates' G_{VIII} between G_{VII} and C_{VIII}, resulting in a stretching of the phosphodiester linkage between G_{VII} and C_{VIII} and, as shown in Figure 10A, providing a solvent-exposed single-stranded site which would presumably be a good target for P1 nuclease digestion. For ss(GTC)₁₅ stacking of G_{VIII} onto G_{IX} could conceivably produce the same effect for the phosphodiester linkage between C_{VIII} and G_{IX}. However, this linkage was not cleaved by P1 nuclease and, further, did not develop the same level of solvent exposure as the equivalent site of ss(CTG)₁₅. Instead, the loop of ss(GTC)₁₅ developed a bend towards the major groove of the upper stem region. This can clearly be seen when compared with the ss(CTG)₁₅ conformation

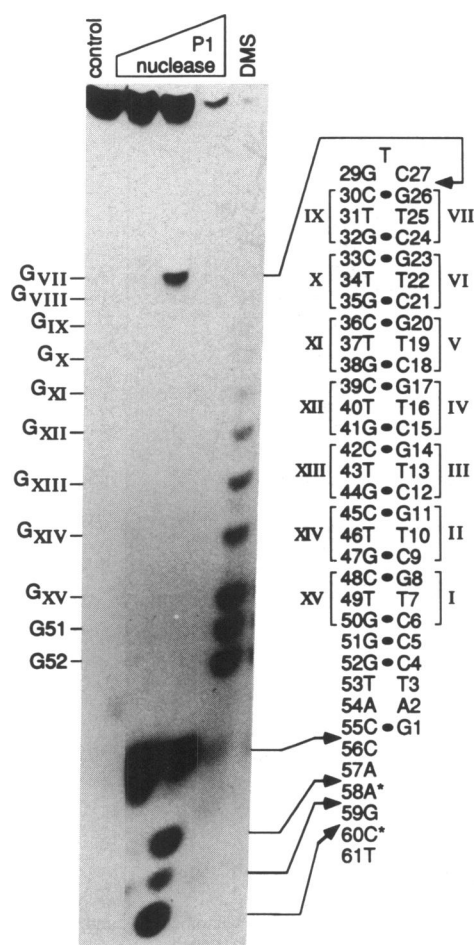


Figure 9. P1 nuclease digestion of ss(CTG)₁₅. The oligonucleotide containing (CTG)₁₅ purified from pCTG15 was labeled on the CTG-containing strand only with Klenow enzyme and ³²P. One micromolar unlabeled synthetic oligonucleotide of the same sequence as the labeled strand was added to the oligonucleotide liberated from the plasmid, placed in a 100°C water bath for 3 min and immediately chilled on ice. The amounts of P1 nuclease used to digest ss(CTG)₁₅ (from left to right) were 1.15×10^{-2} , 3.46×10^{-2} and 0.104 U respectively. DMS reactions were performed as described for ss(GTC)₁₅ (Fig. 8). Roman numerals represent triplet repeat numbers. The nucleotides labeled with ³²P are marked with an asterisk. Arrows indicate sites of P1 nuclease cleavage.

in Figure 10. These conformers are viewed from exactly the same angle and are taken at equivalent points of the dynamics trajectory (80 ps). This bend may result in protection of the phosphodiester linkages to the 3'-side of T_{VIII} and a corresponding increase in the solvent exposure of those linkages to the 5'-side of T_{VIII}. These latter phosphodiester are highlighted in Figure 10B and are found to be the major sites of P1 nuclease digestion experimentally.

DISCUSSION

In 50 mM Na⁺, 150 mM K⁺ significant oxidation of the thymines in the stem of ss(GTC)₁₅ required an incubation temperature of 60–70°C (Figs 4 and 5). These results indicate that under physiological conditions the T·T mismatches in a ss(GTC)₁₅ hairpin are extensively stacked. Apart from modest differences in temperature dependencies, the KMnO₄ reactivities of the thymines

in ss(GTC)₁₅ were similar to those of ss(CTG)₁₅ (Figs 3–5). Since ¹H NMR studies have shown that the thymines in the homoduplex d(CTG/CTG)₃ are base paired (X. Gao, personal communication) we suspect that the thymines in ss(GTC)₁₅ are also base paired, presumably through a hydrogen bond formed between H3 and O4.

Electrophoretic mobility melting profiles (EMMPs): a method for assessing hairpin stability

At 16°C ss(GTC)₁₅ and ss(CTG)₁₅ migrated rapidly in polyacrylamide gels relative to their ds forms. To determine whether the rapid electrophoretic mobilities of these DNAs were due to hairpin conformations, electrophoresis experiments were performed at various temperatures. We refer to the set of data generated from these experiments (Fig. 7) as an electrophoretic mobility melting profile, or EMMP. The EMMP technique described in this paper is a variation of temperature gradient gel electrophoresis (TGGE) developed by Wartell and colleagues (14,15). In TGGE a temperature gradient, either perpendicular or parallel to the direction of electrophoresis, is applied to a homogeneous polyacrylamide gel. When the temperature of the gel is sufficient to partially melt duplex (and presumably hairpin) DNA, its electrophoretic mobility is reduced. TGGE, like hypochromicity assays (16), can be used to estimate the *T*_m of a DNA by determining the midpoint of its electrophoretic phase transition.

For electrophoretic standards in the EMMP studies ds(GTC)₁₅ and ds(CTG)₁₅ were used. There were several reasons for using these dsDNAs as standards. First, the ds and ss forms of the DNAs migrated to similar positions in the polyacrylamide gel. Hence, local changes in temperature, pH, polyacrylamide pore size, etc. should effect ss and dsDNAs in a similar manner. Secondly, the ds forms of the DNAs contained more than twice as many Watson–Crick C-G base pairs compared with their respective ss forms. Hence, the ds forms were much more heat stable. Thirdly, since the ssDNAs in the present study were obtained from a plasmid, the ds forms DNAs were readily available. The fourth and final reason for using the dsDNAs as electrophoretic standards was that preliminary data indicated that as the temperature of the polyacrylamide gel was increased, the ss forms of the DNAs started to migrate more slowly than their ds forms. Hence, a qualitative change, as well as a quantitative change, could be documented by EMMPs. If a longer dsDNA sequence had been used as a standard the qualitative effect would not have been observed.

From the EMMP data the *T*_m of ss(GTC)₁₅ was determined to be 38°C, 10°C lower than the *T*_m of ss(CTG)₁₅. The difference in stabilities of the DNAs (10°C) agrees with the KMnO₄ data and provides further evidence that the ss(GTC)₁₅ hairpin is not as heat stable as the ss(CTG)₁₅ hairpin (Figs 3–5). The EMMP data also revealed a single electrophoretic phase transition for ss(GTC)₁₅ and ss(CTG)₁₅ (Fig. 7). These results suggest that once the stacking interactions of the thymines are destabilized, the entire hairpin begins to melt. This interpretation is consistent with KMnO₄ experiments that showed high temperatures were required for oxidation of the thymines in the stems of the ss(GTC)₁₅ and ss(CTG)₁₅ hairpins.

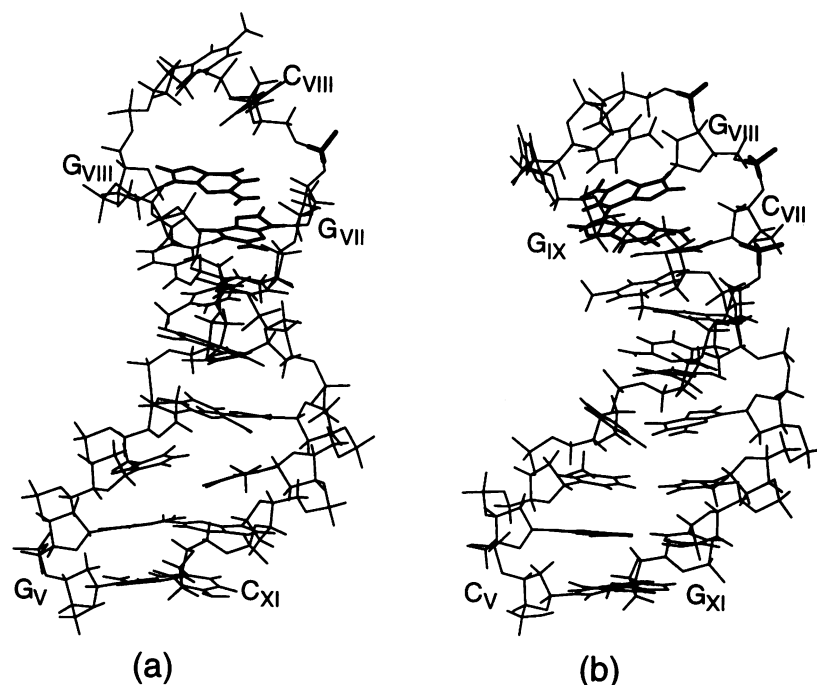


Figure 10. Conformations generated in molecular dynamics simulations of (a) ss(CTG)₁₅ and (b) ss(GTC)₁₅. Each conformation is that after 80 ps simulation time. Base numbering follows the scheme outlined in Figures 9 and 10. The three-dimensional structure of the loop is aligned in a similar manner to the schematic defining the primary sequence in Figures 9 and 10 (i.e. the bases to the 5'-side of TvIII are on the right-hand side of the loop region).

Guanine-guanine stacking appears to play an important role in the loop structures of ss(GTC)₁₅ and ss(CTG)₁₅

Results from P1 nuclease digestion revealed an unexpected asymmetry in phosphodiester cleavage patterns (Figs 8 and 9). The loop regions of ss(GTC)₁₅ and ss(CTG)₁₅ were preferentially cleaved at the 5'-sides. Molecular dynamics simulations suggested that the phosphodiester cleavage patterns could be explained on the basis of guanine-guanine stacking in the loops (Fig. 10). This finding raises interesting possibilities concerning loop and stem structures in hairpins formed from the ss Class I triplet repeat member (CGG)_n. With this sequence a hairpin can be formed with either a 'perfectly' aligned stem (i.e. CGG repeats on the 5'-side base paired to CGG repeats on the 3'-side) and a three-membered loop or a 'misaligned' stem (i.e. CGG repeats on the 5'-side base paired to GCG repeats on the 3'-side) and a two-membered loop. If the most stable loop structure of the ss(CG G)_n hairpin is not compatible with the most stable stem structure, a flexible hairpin might result.

ACKNOWLEDGEMENTS

MM, AY and JD were supported by the Oklahoma Agricultural Experiment Station at Oklahoma State University. We thank Dr Ulrich Melcher (OSU) for critical review of this manuscript.

REFERENCES

- Mitas, M., Yu, A., Dill, J., Kamp, T.J., Chambers, E.J. and Haworth, I.S. (1995) *Nucleic Acids Res.*, **23**, 1050–1059.
- Schlötterer, C. and Tautz, D. (1992) *Nucleic Acids Res.*, **20**, 211–215.
- McCarthy, J.G. and Rich, A. (1991) *Nucleic Acids Res.*, **19**, 3421–3429.
- Maxam, A.M. and Gilbert, W. (1980) *Methods Enzymol.*, **65**, 499–560.
- Lohman, T.M. and Ferrari, M.E. (1994) *Annu. Rev. Biochem.*, **63**, 527–570.
- Wohlrab, F. (1992) *Methods Enzymol.*, **212B**, 294–301.
- Pearlman, D.A., Case, D.A., Caldwell, J.C., Seibel, G.L., Singh, U.C., Weiner, P. and Kollman, P.A. (1991) University of California, San Francisco.
- Weiner, P., Kollman, P.A., Nguyen, D.T. and Case, D.A. (1986) *J. Comput. Chem.*, **7**, 230–252.
- Jorgensen, W.L., Chandrasekhar, J., Madura, J., Impey, R.W. and Klein, M.L. (1983) *J. Chem. Phys.*, **79**, 926–937.
- van Gunsteren, W.F., Berendsen, H.J., Guersten, R.G. and Zwinderman, H.R. (1986) *Annls NY Acad. Sci.*, **482**, 287–303.
- van Gunsteren, W.F. and Berendsen, H.J.C. (1977) *Mol. Phys.*, **34**, 1311–1327.
- Hayatsu, H. and Ukita, T. (1967) *Biochem. Biophys. Res. Commun.*, **29**, 556–561.
- Rubin, C.M. and Schmid, C.W. (1980) *Nucleic Acids Res.*, **8**, 4613–4619.
- Ke, S.-H. and Wartell, R.M. (1993) *Nucleic Acids Res.*, **21**, 5137–5143.
- Wartell, R.M., Hosseini, H.G. and Moran, C.P., Jr (1990) *Nucleic Acids Res.*, **18**, 2699–2705.
- Puglisi, J.D. and Tinoco, I., Jr (1989) *Methods Enzymol.*, **180**, 304–325.