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

Adong Yu, Jeffrey Dill, Sara S. Wirth, George Huang¹, Jeffrey H. Lee², Ian S. Haworth¹,² and Michael Mitas*¹

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), GGTACCA; (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 XhoI. For sequence analysis 8 pmol of a synthetic oligonucleotide complementary to a portion of the thymidine kinase promoter of pBLCAT2 (sequence GTTCGAAATTCCGCAATGACAA) was added to 1 µg linearized plasmid DNA. DNA sequence analysis was performed by the OSU Protein DNA Resource Facility on an ABI373 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 pΔT₁₁CTG15.

KMnO₄ oxidation

KMnO₄ oxidation of ss(GTC)₁₅, ΔT₁₁GTC₁₅ and ss(CTG)₁₅ were as previously described (1). Briefly, the pyrimidine-rich triplet repeat strand was liberated from pGTC15, pΔT₁₁GTC15 or pCTG15 by digestion with HindIII followed by BamHI. Oligo-
nucleotides were purified from the plasmid by agarose gel electrophoresis. DNAs were incubated with \[^{32}P\]ATP and polynucleotide kinase after digestion with HindIII. Unlabeled synthetic oligonucleotide (~1.4 pmol) of the same sequence as the labeled strand was added to 4 x 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. KMnO\(_4\) 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 HindIII for 1 h at 37°C in a volume of 70 μl in a 3% agarose gel. Recessed ends were labeled at the 3'-termini by adding 5 μl [α-\(^{32}P\)]dCTP, 5 μl [α-\(^{32}P\)]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 BamHI in a volume of 70 μl and applied directly to a Nuclacp column (Stratagene, La Jolla, CA) for further removal of unincorporated \(^{32}P\)-dNTPs. Labeled oligonucleotides were not purified from labeled vector. Probes were diluted to 2 x 10^6 d.p.m./μl with H\(_2\)O.

Oligonucleotides liberated from the plasmid at 0.7 nM were incubated with various amounts of *Escherichia coli* single-stranded DNA binding protein (5) (EcoSSB, 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 x 10^6 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 H\(_2\)O/Xerex antifreeze mixture (1:1) was circulated through the electrophoresis unit at 1–10°C. Temperature of the H\(_2\)O/antifreeze mixture was controlled with a Brinkman RC6 temperature bath. To arrive at gel temperatures between 25 and 66°C the H\(_2\)O/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 x 10^6 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 MgCl\(_2\), 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\(_{v}\) base pair [ss(CTG)_15] and the G\(_{v}-C\(_{v}\) 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/Å2. 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 10-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 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 ΔTII(GTC)15]. The predicted structure of ΔTII(GTC)15 is shown on the right in Figure 2B. Due to the deletion in ΔTII(GTC)15, the thymine in triplet XIV does not have a nucleotide with which it can pair. Therefore, if ΔTII(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 1](image-url)
The results from Figure 1 showed that ss(GTC) is migrated with rapid mobility at 16°C. The rapid electrophoretic mobility of ss(GTC) is and ss(CTG) i5 are熔解 profiles of temperature of 70°C (Fig. 5B).

In the presence of KC1 significant oxidation of the thymines in the stem of ss(GTC) i5 were more stable than the T-T interactions in ss(CTG) i5. In the presence of KC1 significant oxidation of the thymines in the stem region of ss(GTC) i5 occurred at 50°C (Fig. 3A). In the presence of KC1 significant oxidation required an incubation temperature of 60-70°C (Figs 4B and 5B). Stabilization of the T-T interactions by added KC1 is consistent with a hairpin structure of ss(GTC) i5.

In the presence of KC1 significant oxidation of the thymines in the stem region of ss(CTG) i5 required an incubation temperature of 60-70°C (Figs 4B and 5B). Stabilization of the T-T interactions by added KC1 is consistent with a hairpin structure of ss(GTC) i5.

The results from Figure 1 showed that ss(GTC) i5 migrated with rapid mobility at 16°C. The rapid electrophoretic mobility of ss(GTC) i5 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 (GTC) i5 and (CTG) i5 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 mM EcoSSB. At this concentration EcoSSB anneals ss(GTC) i5 to its complementary sequence (Fig. 6) and ss(CTG) i5 to its complementary sequence (data not shown).

In agreement with the expected results, a decrease in the relative electrophoretic mobility of ss(GTC) i5 was observed at higher temperatures (Fig. 7). The relative electrophoretic mobility of ss(GTC) i5 decreased at a sharp rate between 33 and 45°C. In contrast, the relative electrophoretic mobility of ss(CTG) i5 decreased at a sharp rate between 41 and 55°C. At 55°C the relative electrophoretic mobility of ss(GTC) i5 was 0.890, the lowest value observed. Between 10 and 55°C the electrophoretic mobility of ss(GTC) i5 decreased 22.3%. Half maximal electrophoretic mobility of ss(GTC) i5 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 (Tm) of the DNA. Note that the isomobility temperature, or Tm [defined as the temperature at which the electrophoretic mobilities of ss(GTC) i5 and ds(GTC) i5 were identical] was 39°C, a temperature comparable to the Tm of ss(GTC) i5. The temperature corresponding to an electrophoretic mobility of 1.027 for ss(CTG) i5 was 48°C. This result indicates

**Figure 2. KMnO4 oxidation of ss(GTC) i5 and ss ATn(GTC) i5. KMnO4 oxidations were performed at 30°C in 50 mM Na+ with (A) ss(GTC) i5 and (B) ATn(GTC) i5 with** ss(GTC) i5. **KMnC>4 oxidation sites of particular thymine residues are indicated by arrows.** XIV were preferentially oxidized, providing further evidence for a hairpin conformation.

**KMnO4 oxidation at various temperatures reveals reduced thermal stability of the (GTC) i5 hairpin**

To further investigate the possibility of a hairpin conformation of ss(GTC) i5, KMnO4 oxidation experiments were performed in 50 mM Na+ with and without 150 mM KC1 at 40, 50, 60 and 70°C. For comparison, parallel experiments were performed with ss(CTG) i5. In the presence or absence of KC1 the thymines in the stems of ss(GTC) i5 and ss(CTG) i5 were not oxidized by KMnO4 at 40°C (data not shown). In the absence of added KC1 significant oxidation of the thymines in the presumed stem region of ss(GTC) i5 occurred at 50°C (Fig. 3A). In the presence of KC1 significant oxidation required an incubation temperature of 60-70°C (Figs 4B and 5B). Stabilization of the T-T interactions by added KC1 is consistent with a hairpin structure of ss(GTC) i5.

In the absence of KC1 significant oxidation of the thymines in the stem region of ss(CTG) i5 required an incubation temperature of 60°C (Fig. 4A). This result indicated that the T-T interactions in ss(CTG) i5 were more stable than the T-T interactions in ss(GTC) i5. In the presence of KC1 significant oxidation of the thymines in the stem of ss(CTG) i5 required an incubation temperature of 70°C (Fig. 5B).

**Electrophoretic mobility melting profiles of ss(GTC) i5 and ss(CTG) i5**

The results from Figure 1 showed that ss(GTC) i5 migrated with rapid mobility at 16°C. The rapid electrophoretic mobility of ss(GTC) i5 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 (GTC) i5 and (CTG) i5 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 mM EcoSSB. At this concentration EcoSSB anneals ss(GTC) i5 to its complementary sequence (Fig. 6) and ss(CTG) i5 to its complementary sequence (data not shown).

In agreement with the expected results, a decrease in the relative electrophoretic mobility of ss(GTC) i5 was observed at higher temperatures (Fig. 7). The relative electrophoretic mobility of ss(GTC) i5 decreased at a sharp rate between 33 and 45°C. In contrast, the relative electrophoretic mobility of ss(CTG) i5 decreased at a sharp rate between 41 and 55°C. At 55°C the relative electrophoretic mobility of ss(GTC) i5 was 0.890, the lowest value observed. Between 10 and 55°C the electrophoretic mobility of ss(GTC) i5 decreased 22.3%. Half maximal electrophoretic mobility of ss(GTC) i5 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 (Tm) of the DNA. Note that the isomobility temperature, or Tm [defined as the temperature at which the electrophoretic mobilities of ss(GTC) i5 and ds(GTC) i5 were identical] was 39°C, a temperature comparable to the Tm of ss(GTC) i5. The temperature corresponding to an electrophoretic mobility of 1.027 for ss(CTG) i5 was 48°C. This result indicates...
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 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$_{5\text{VIII}}$ and C$_{55}$, indicating that these nucleotides participated in base pairing and/or base stacking interactions (Fig. 8). At the 3'-terminus the C$_{55}$–C$_{56}$ 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$_{1\text{VII}}$. Therefore, the major site of P1 cleavage within the triplet repeat region was the T$_{1\text{VII}}$–C$_{1\text{VII}}$ phosphodiester bond. Minor sites...
of P1 cleavage in the loop region included the CVH–GVIII, GVIII–TVIII and TVIII–CVIII phosphodiester bonds. Surprisingly, the TVIII–CVIII phosphodiester bond was not the major site of P1 cleavage. This result was surprising since TVIII was readily oxidized by KMnO₄. The structure of ss(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 T₅₃–A₅₄ and A₅₄–C₅₅ phosphodiester bonds of ss(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 KMnO₄. In the hairpin structure of ss(GTC)₁₅ T₅₃ is opposite T₃. Hence, since T₃ is not base paired to T₅₃, the 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 C₄–G₅₂ and G₁–C₅₅ (as shown in Fig. 8).

Similar to ss(GTC)₁₅, incubation of ss(CTG)₁₅ with P1 nuclease did not result in cleavage of phosphodiester bonds between the 3' portion of the loop and C₅₅ (Fig. 9). However, in contrast to multiple sites of P1 cleavage in the loop of ss(GTC)₁₅, a single digestion product in the loop of ss(CTG)₁₅ was observed. The product co-migrated with the DMS product generated from cleavage of G₉γ (the G₉γ band is faint, but clearly detectable on the autoradiograph). Therefore, the single site of cleavage within the triplet repeat region was the GVH–CVIII 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 ss(CTG)₁₅ and ss(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 ss(CTG)₁₅ and ss(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 ss(GTC)₁₅ and ss(CTG)₁₅ respectively]. During the energy minimization of ss(CTG)₁₅ the CVIII–GVIII...
and CIX–GVII base pairs (those closest to the single base loop) were essentially retained. For ss(GTC)15 the base pairing of GVIII–CVIII 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 GvII and GvIII for ss(CTG)15 and between GvIII and GvIX for ss(GTC)15. In each case CvIII was displaced and formed a weak stacking interaction with TvIII.

In the ss(CTG)15 hairpin loop (Fig. 10A) the stacking of GvIII onto GvII essentially 'intercalates' GvII between GvIII and CvIII, resulting in a stretching of the phosphodiester linkage between GvII and CvIII 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(CTG)15 stacking of GvIII onto GvX could conceivably produce the same effect for the phosphodiester linkage between CvIII and GvX. 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)\textsubscript{15}. The oligonucleotide containing (CTG)\textsubscript{15} purified from pCTGI5 was labeled on the CTG-containing strand only with Klenow enzyme and \textsuperscript{32}P. 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)\textsubscript{15} (from left to right) were 1.15 \times 10^{-2}, 3.46 \times 10^{-2} and 0.104 U respectively. DMS reactions were performed as described for ss(GTC)\textsubscript{15} (Fig. 8). Roman numerals represent triplet repeat numbers. The nucleotides labeled with \textsuperscript{32}P 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 tym and a corresponding increase in the solvent exposure of those linkages to the 5'-side of tym- These latter phosphodiesters are highlighted in Figure 10B and are found to be the major sites of P1 nuclease digestion experimentally.

DISCUSSION

In 50 mM Na\textsuperscript{+}, 150 mM K\textsuperscript{+} significant oxidation of the thymines in the stem of ss(GTC)\textsubscript{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)\textsubscript{15} hairpin are extensively stacked. Apart from modest differences in temperature dependencies, the KMnO\textsubscript{4} reactivities of the thymines in ss(GTC)\textsubscript{15} were similar to those of ss(CTG)\textsubscript{15} (Figs 3–5). Since \textsuperscript{1}H NMR studies have shown that the thymines in the homoduplex d(CTG/CTG)\textsubscript{3} are base paired (X. Gao, personal communication) we suspect that the thymines in ss(GTC)\textsubscript{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)\textsubscript{15} and ss(CTG)\textsubscript{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 \textit{Tm} of a DNA by determining the midpoint of its electrophoretic phase transition.

For electrophoretic standards in the EMMP studies ds(GTC)\textsubscript{15} and ds(CTG)\textsubscript{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 \textit{Tm} of ss(GTC)\textsubscript{15} was determined to be 38°C, 10°C lower than the \textit{Tm} of ss(CTG)\textsubscript{15}. The difference in stabilities of the DNAs (10°C) agrees with the KMnO\textsubscript{4} data and provides further evidence that the ss(GTC)\textsubscript{15} hairpin is not as heat stable as the ss(CTG)\textsubscript{15} hairpin (Figs 3–5). The EMMP data also revealed a single electrophoretic phase transition for ss(GTC)\textsubscript{15} and ss(CTG)\textsubscript{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 KMnO\textsubscript{4} experiments that showed high temperatures were required for oxidation of the thymines in the stems of the ss(GTC)\textsubscript{15} and ss(CTG)\textsubscript{15} hairpins.
Guanine–guanine stacking appears to play an important role in the loop structures of ss(GTC)$_{15}$ and ss(CTG)$_{15}$.

Results from PI 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(CGG)$_n$ 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.