1260–1268 Nucleic Acids Research, 1995, Vol. 23, No. 7 

# Structural characterization of d(CAACCCGTTG) and d(CAACGGGTTG) mini-hairpin loops by heteronuclear NMR: the effects of purines versus pyrimidines in DNA hairpins

**Daina Z. Avizonis and David R. Kearns***

Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92037-0343, USA

Received September 12, 1994; Revised and Accepted February 10, 1995

## ABSTRACT

**The DNA decamers, d(CAACCCGTTG) and d(CAACGGGTTG) were studied in solution by proton and heteronuclear NMR. Under appropriate conditions of pH, temperature, salt concentration and DNA concentration, both decamers form hairpin conformations with similar stabilities [Avizonis and Kearns (1995) *Biopolymers*, 35, 187–200]. Both decamers adopt mini-hairpin loops, where the first and last four nucleotides are involved in Watson–Crick hydrogen bonding and the central two nucleotides, CC or GG respectively, form the loop. Through the use of proton–proton, proton–phosphorus and natural abundance proton–carbon NMR experiments, backbone torsion angles (β, γ and ε), sugar puckers and interproton distances were measured. The nucleotides forming the loops of these decamers were found to stack upon one another in an L1 type of loop conformation. Both show $\gamma^{tr}$ and unusual β torsion angles in the loop-closing nucleotide $G_7$, as expected for mini-hairpin loop formation. Our results indicate that the β and ε torsion angles of the fifth and sixth nucleotides that form the loop and the loop-closing nucleotide $G_7$ are not in the standard *trans* conformation as found in B-DNA. Although the loop structures calculated from NMR-derived constraints are not well defined, the stacking of the bases in the two different hairpins is different. This difference in the base stacking of the loop may provide an explanation as to why the cytosine-containing hairpin is thermodynamically more stable than the guanine-containing hairpin.**

## INTRODUCTION

Early analyses of hairpin loop structures concluded that DNA hairpins containing 4 or 5 nt in the loop would be the most stable and that DNA hairpin loops containing only 2 nt were believed to be sterically unlikely (2–11). However, more recent NMR analyses of hairpin loop structures have revealed that loops containing 2 nt can form stable mini-hairpin loops (12–15). Stable mini-hairpin loops have also been found to exist in RNA (14,16,17).

To date, only a few sequences containing 2 nt in the DNA hairpin loop have been studied (12–15,18–21). From this limited amount of information some general rules of hairpin folding have been proposed. There are two main structural families of mini-hairpin loops. The first, **L1**, is characterized by the stacking of the two loop nucleotides upon the 5′-end of the stem. This family of hairpins is thought to be preferred by loop nucleotide sequences when the first nucleotide is a purine. The second structural family, **L2**, is characterized by the first nucleotide of the loop folding into the minor groove of the stem and the second nucleotide stacking upon the 5′-end of the stem. **L2** appears to be preferred when the first base of the loop is a pyrimidine. Both families of structures share the common features of $\gamma^{tr}$ and $\beta^{+}$ torsion angles at the 5′–3′ loop junction (15) and that a C–G loop-closing base pair encourages formation of mini-hairpin loops (15,22).

The first purine-containing mini-hairpin loop was analyzed by Orbons *et al.* (18). This was an NMR study of the octamer, d($^{m5}$CG$^{m5}$CGAG$^{m5}$CG), where the cytosines are methylated at the $C_5$ position. The hairpin loop is formed by two purines, GA, with a C–G loop-closing base pair. Molecular dynamics studies indicate that this octamer adopts an **L1** structure. A more recent thermodynamic study of two near-palindromic dodecamers containing central GA and GI mispairs indicated that mini-hairpin loops were formed at low salt and DNA concentrations (20). This motivated a molecular dynamics study of their conformational feasibility (21). This conformational study showed that stable purine-containing mini-hairpin loops may be formed without a significant buckle in the final C–G loop-closing base pair. These theoretical models fell into four distinct structural groups, all of which appear to fall into the more general **L1** family.

Structures $\mathbf{L1anti_1}$ and $\mathbf{L1anti_2}$ are both characterized by an *anti*-X (180°) torsion angle of the second purine in the loop. They are distinguished from each other by the position of the second purine. $\mathbf{L1anti_1}$ has both purines stacked on top of each other; $\mathbf{L1anti_2}$ has the first purine of the loop stacked upon the 5′-end of the stem, while the second purine is unstacked. The other two models, $\mathbf{L1syn_1}$ and $\mathbf{L1syn_2}$, are characterized by a *syn*-X (0°)

* To whom correspondence should be addressed

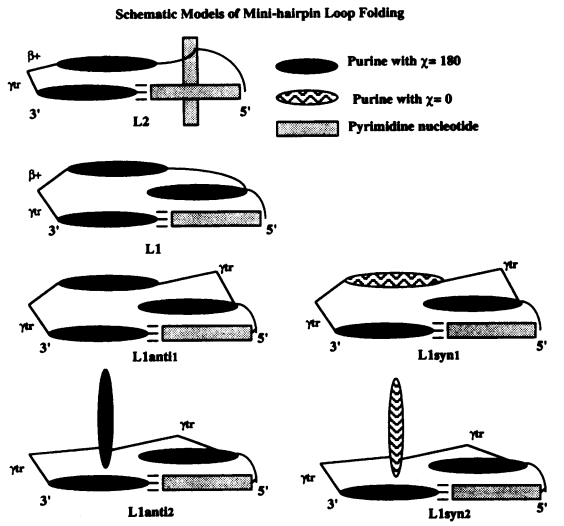


**Figure 1.** A schematic diagram showing the six possible mini-hairpin loop structures that have been published viewed from the major groove (15,21).

orientation of the second base in the loop. These two models are distinguished in the same way as **L1anti$_1$** and **L1anti$_2$**. All four of these theoretical models are distinguished from the original **L1** loop by an additional $\gamma^{tr}$ torsion angle in the loop and the lack of a $\beta^+$ torsion angle. All of the structures discussed here are represented schematically in Figure 1.

In a previous study (1) the interconversion and thermodynamics of hairpin and duplex conformers were analyzed for two DNA decamers: (i) d($C_1A_2A_3C_4C_5C_6G_7T_8T_9G_{10}$), C-mer; (ii) d($C_1A_2A_3C_4G_5G_6G_7T_8T_9G_{10}$), G-mer. Together in a 1:1 solution these two oligonucleotides form a normal self-complementary B-DNA duplex (in progress). Separately, each strand may form either a hairpin or a duplex containing a double tandem mispair in the center of the duplex. At low salt and DNA concentrations a stable hairpin form is favored for the G-mer. In contrast, the C-mer may form a stable monomeric hairpin under high salt (100 mM NaCl) and high DNA concentrations (8 mM), provided that the pH value is maintained above 6.4. According to UV melting studies, the C-mer hairpin melts at 3°C higher than the G-mer hairpin. These studies also indicate that the hairpins may have up to 4 bp in the stem, which supports the formation of a loop containing only 2 nt. Here we characterize the conformations of the C-mer and G-mer hairpins.

## EXPERIMENTAL

d(CAACGGGTTG) and d(CAACCCGTTG) were purchased from Pharmacia LKB Biotechnology Inc. (NJ). The DNA samples were treated with Chelex to remove any divalent and/or paramagnetic ions and used without further purification. All NMR measurements were recorded on a Bruker AMX-500 operating system with quadrature detection at 288 K unless otherwise noted. NMR data were collected in phase-sensitive mode using time proportional phase incrementation (23). Proton chemical shifts were referenced against TMSP-$D_6$ as an internal standard. Phosphorus chemical shifts were referenced against external phosphoric acid. Carbon chemical shifts were referenced against 3-trimethylsilyl-1-propionate externally.

d(CAACGGGTTG) (63 $OD_{260\ nm}$) was dissolved in 400 µl $D_2O$ and the pH (uncorrected) was adjusted to 6.2 with HCl/NaOH. No additional salt was added to the sample. The resulting solution was dried under a stream of nitrogen and redissolved four times in 99.98% $D_2O$. The final concentration of d(CAACGGGTTG) was 1.5 mM DNA.

d(CAACCCGTTG) (173 $OD_{260\ nm}$) was prepared in the same way, with the pH (uncorrected) adjusted to 6.5 with HCl/NaOH. No additional salt was added to the sample. The resulting solution was dried under a stream of nitrogen and redissolved four times in 99.98% $D_2O$. The final concentration of d(CAACCCGTTG) was 7.8 mM in strand.

### Homonuclear NMR spectroscopy

Three NOESY spectra with mixing times of 150, 300 and 500 ms were collected. For all three experiments, 4096 complex points in $t_2$ and 512 points in $t_1$ were collected with a relaxation delay of 1.8 s and a sweep width of 5555.56 Hz. Fifty six transients were averaged for each block. Spectra in water were obtained using the jump return sequence (24) at 274 K (278 K for the C-mer). A modified 250 ms mixing time NOESY with a jump return sequence in place of the read pulse was used to collect spectra in 90% $H_2O$ (24) at 274 K (278 K for the C-mer). These spectra were collected with 4096 complex points in $t_2$,, 512 blocks in $t_1$, a sweep width of 11111.11 Hz and 64 transients. The partially-exclusive COSY (PE-COSY) spectrum (25,26) was recorded with 4096 complex points in $t_2$ and 512 points in $t_1$. The double quantum filtered COSY (DQF-COSY) (27,28) was collected with 4096 complex points in $t_2$ and 512 points in $t_1$, with a sweep width of 5555.56 Hz (5055 Hz for the C-mer) and a relaxation delay of 2 s. Two TOCSY spectra with mixing times of 30 and 70 ms were collected (29). Both TOCSY experiments had 2048 complex points in $t_2$, 400 points in $t_1$ and a sweep width of 5555.56 Hz.

### Heteronuclear NMR spectroscopy

Phosphorus experiments were collected at 202.46 MHz. One-dimensional spectra were recorded with proton decoupling during acquisition. The non-selective $^1H$-$^{31}P$ COSY (30) was collected with 2048 points in the proton dimension ($t_2$) and 512 points in the phosphorus dimension ($t_1$), with a sweep width of 2501 Hz (5024 Hz for the C-mer) in the proton dimension and 607 Hz in the phosphorus dimension. Selective $^1H$-$^{31}P$ COSY (31) were collected with a 180° Gaussian pulse exciting only the 3′-sugar protons. The proton sweep width was 1500 Hz, with 2048 complex data points, and the phosphorus dimension sweep width was 607 Hz, with 400 points in $t_1$.

Natural abundance $^{13}C$ spectra were recorded at 125.76 MHz. The folded HSQC (32–34) was collected with 500 $t_1$ points in the carbon dimension, with a sweep width of 2515 Hz (20 p.p.m.) and 256 scans (64 for the C-mer) for each $t_1$ block; 4096 points were collected in the $t_2$ proton dimension, with a sweep width of 4505 Hz.

### Data analysis

All NMR data were transferred to a Silicon Graphics computer (either an IRIS 4D or Indigo) and processed using FELIX (Hare Research Inc., Woodinville, WA). NOESY data were apodized with an 80° phase-shifted sine-squared function and zero-filled to 2048 points in both dimensions. Distance constraints were

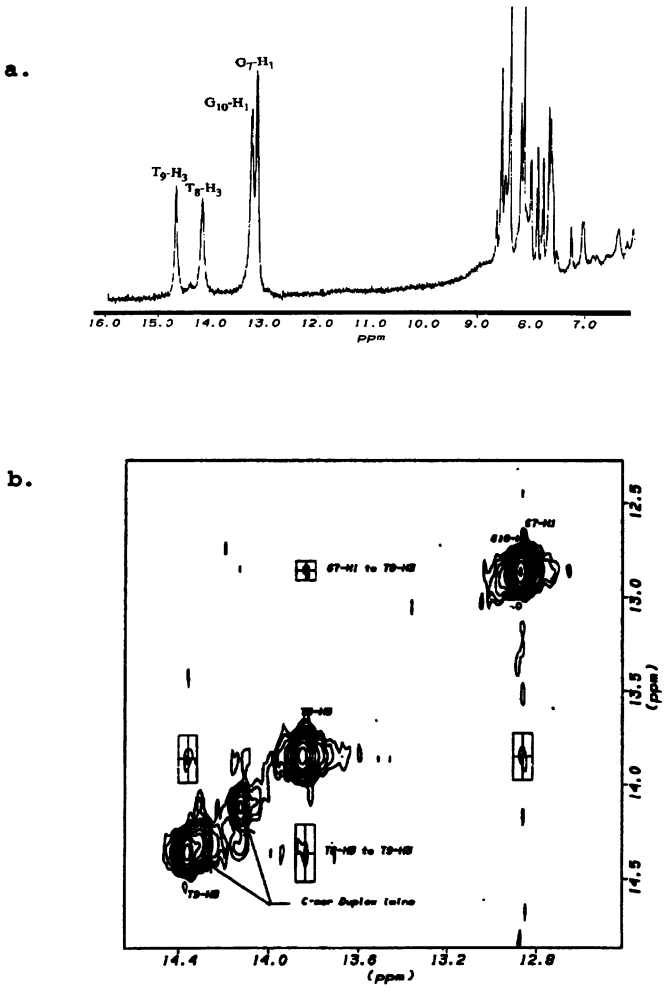


**Figure 2.** (**a**) The imino proton region of d(CAACCCGTTG) showing four resolved imino protons at 274 K, 5 mM DNA, 100 mM NaCl, pH 6.7. (**b**) A portion of the 11NOESY showing inter-imino cross-peaks.

estimated using very strong (1.8–2.2 Å), strong (2.2–2.5 Å), medium-strong (2.5–2.9 Å), medium (2.9–3.3 Å), weak-medium (3.3–3.7 Å), weak (3.7–4.1 Å) and very weak (4.1–6.0 Å) cross-peak intensities from the 150 and 300 ms (checked against 500 ms) NOESY spectra. These constraint limits were derived from integrated cross-peak intensities with known distances, such as the $C_4H_5$–$H_6$ cross-peak. The cross-peak intensities were categorized by contour counting. The NOESY spectra in water required a different processing scheme. The water resonance was subtracted from the spectrum as described by Marion and Bax (26). These data were then apodized with a Gaussian window with a line narrowing of 8 Hz in the $t_2$ dimension. Before the Fourier transform of the $t_1$ dimension was calculated, the first block of the $t_2$ dimension was multiplied by 1/2. The $t_1$ dimension was then processed with an 80° phase-shifted sine-squared function.

All J-correlated spectra were apodized with a 45° phase-shifted, sine-squared function and zero-filled to 4096 points in both dimensions to enhance spectral resolution.

### Restrained molecular dynamics calculations

Molecular dynamics and energy minimization were calculated on a Silicon Graphics Indigo workstation using the XPLOR program

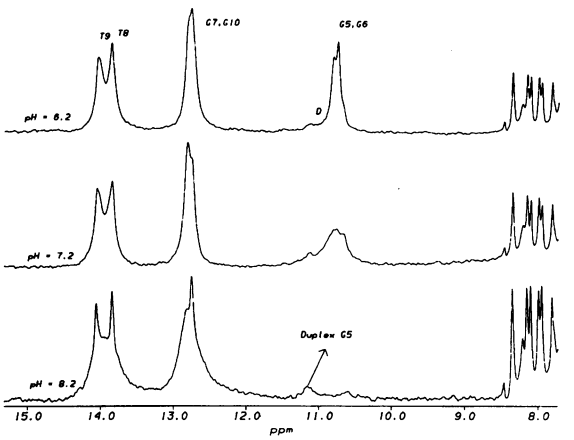


**Figure 3.** The low-field region of the $^1$H-NMR spectrum of d(CAACGGGTTG) at 293 K in 90% $H_2O$ at three different pH values.

(Brünger, 1992). For these structure calculations both NOESY distance data and NMR-derived dihedral constraints were used. NOEs involving intra-sugar interactions were not used due to the possibility that these NOEs may be significantly effected by spin diffusion. Starting structures satisfying loop closure were generated by varying torsional angles and base stacking schemes in the loop region using InsightII (Biosym Technologies Inc.). Structures were determined using a simulated annealing protocol. The system was cooled from 1500 K to a final temperature of 300 K in 25 K steps with 0.250 ps dynamics/step. The scaling factor for the NOE energy term was increased at each temperature during the molecular dynamics simulation. The weighting factor for the non-symmetry-related van der Waal's energy term was also gradually increased to prevent further changes in the global conformation as the temperature of the system was lowered. The final part of the calculation involved all-atom Powell conjugate gradient energy minimization. Hydrogen bond constraints were used throughout the calculation to maintain both DNA base pairing and the double helical nature of the hairpin stem. Further dihedral constraints between base pairs were added in order to avoid serious buckle or stagger displacement.

## RESULTS

### Exchangeable protons

Spectra for analysis of exchangeable protons were measured in water at a temperature of 278 K. The low-field portion of the $^1$H-NMR spectrum of the C-mer is displayed in Figure 2a. Four clearly resolved imino protons are observed in the spectral region of 15.0–13.0 p.p.m., which corresponds to the positions of imino protons that are hydrogen bonded in Watson–Crick type C–G or A–T base pairs. The imino spectrum does not change with DNA or salt concentration (data not shown). It does, however, change as the pH is lowered below 6.5. The NOESY spectrum collected in water was used to assign the imino resonances. NOEs between the imino protons of $G_7$–$T_8$ and $T_9$–$T_8$ are observed, indicating that the stem of the hairpin is made up of four Watson–Crick base pairs, leaving $C_5$ and $C_6$ to form a mini-hairpin loop (Fig. 2b). The NOE between $T_9$ and $G_{10}$ is not observed due to the expected fraying of the terminal base pair.

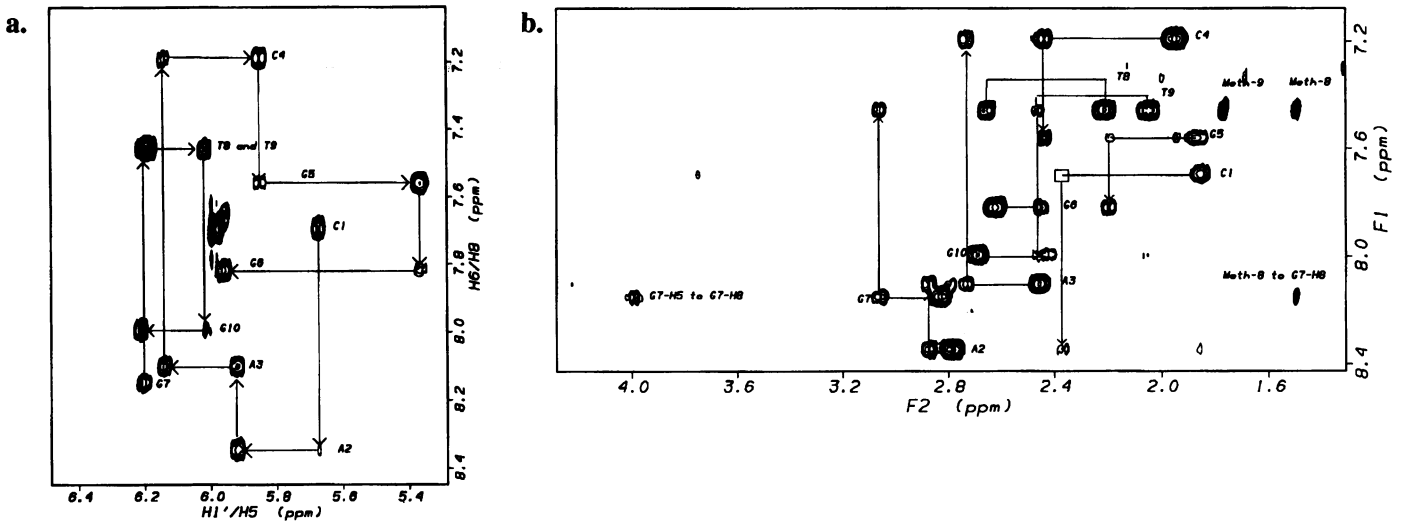


**Figure 4.** The $H_{1'}$–$H_5$, $H_6$ and $H_8$ base proton region of the C-mer (**a**) and the $H_{2'}/H_{2''}$–$H_6$ and $H_8$ base proton region of the G-mer (**b**) showing the sequential connectivity from nucleotides 1–6 and 7–10.

Figure 3 shows the low-field region of the $^1$H-NMR spectrum of the G-mer in 90% $H_2O$ and 10% $D_2O$ at three different pH values. The chemical shift of these protons closely resembles those reported for the d(CAACCCGTTG) mini-hairpin loop. These resonances correspond to four Watson–Crick type base pairs in the hairpin stem. The two overlapping signals at 10.9 p.p.m. were assigned to the imino protons of $G_5$ and $G_6$, which form the mini-hairpin loop. The resonance positions of these imino protons agree well with previous studies of DNA hairpin studies (18,35). Haasnoot demonstrated that one major characteristic of DNA hairpin loops is that the imino protons residing in the loop undergo base catalyzed exchange, whereas the imino protons that are base paired in the stem do not (35). This is observed by the broadening and disappearance of the unpaired $G_5$ and $G_6$ imino protons with increasing pH, while the resonances corresponding to the imino protons of the stem remain sharp. From these data we conclude that the hairpin form of the G-mer consists of four Watson–Crick base pairs in the stem.

## Non-exchangeable protons

The base and sugar protons of both decamers were assigned by means of NOESY and TOCSY experiments. The sequential assignments for the $H_{1'}$ sugar proton to base proton of the C-mer and $H_{2'}/H_{2''}$ sugar protons to base proton of the G-mer are illustrated in Figure 4a and b. The standard B-DNA assignment scheme of sugar proton to its own base proton and its $n-1$ base proton (36,37) may be followed from nucleotide $C_1$ to nucleotide $G_6$ for the G-mer or $C_6$ for the C-mer, at which point there is a break in the assignment pattern, suggesting the absence of stacking on the 3′-side of the loop. The standard B-form stacking pattern resumes at $G_7$ and continues to the final nucleotide, $G_{10}$. There is no base proton to sugar proton connectivity observed between nucleotides $G_6$ (or $C_6$ of the C-mer) and $G_7$, indicative of a sharp turn in the backbone of the helix caused by $\gamma^{tr}$ and unusual β torsion angle combinations (12,13,15,18,19). The proton chemical shift assignments are presented in Table 1.

The chemical shifts exhibit some interesting features. In the G-mer, the $H_8$ base proton of $G_5$ and $G_6$ are shifted up-field compared with other guanines. These up-field shifts may indicate a perturbation in the stacking of the loop with respect to the stem. The chemical shifts for the $H_{4'}$, $H_{5'}$ and $H_{5''}$ sugar protons of $G_6$ are shifted up-field. The $H_{4'}$ proton is shifted 1 p.p.m. up-field from where it is normally expected (4.0–4.5 p.p.m.). The $H_{5'}$ and $H_{5''}$ protons are slightly shifted up-field and show a large difference in their chemical shifts, whereas in B-form DNA the geminal $H_{5'}$ and $H_{5''}$ protons have nearly degenerate chemical shifts. Similar up-field shifts, though less significant, are observed for the C-mer sugar protons. This up-field shift of the $H_{4'}$, $H_{5'}$ and $H_{5''}$ sugar protons of the second nucleotide of the loop is a common feature of all the mini-hairpin loop structures determined by NMR to date. It has been suggested that these protons are located within a shielding zone of the 3′-neighboring

**Table 1.** Proton chemical shifts of d(CAACCCGTTG) and d(CAACGGGTTG) hairpins

| Base | | $T_8H_7H_5$ | $H_6H_8$ | $H_{1'}$ | $H_{2'}$ | $H_{2''}$ | $H_{3'}$ | $H_{4'}$ | $H_{5'}$ | $H_{5''}$ | N-H |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_1$ | C-mer | 5.99 | 7.69 | 5.65 | 1.85 | 2.35 | 4.71 | 4.05 | 3.73[a] | | |
| | G-mer | 5.98 | 7.69 | 5.67 | 1.86 | 2.37 | 4.71 | 4.07 | 3.75[a] | | |
| $A_2$ | C-mer | 7.65 | 8.34 | 5.93 | 2.82 | 2.89 | 5.07 | 4.40 | 4.13 | 4.06 | |
| | G-mer | 7.95 | 8.34 | 5.92 | 2.79 | 2.89 | 5.05 | 4.40 | 4.13 | 4.00 | |
| $A_3$ | C-mer | 7.92 | 8.21 | 6.16 | 2.59 | 2.82 | 5.04 | 4.43 | 4.23[a] | | |
| | G-mer | 7.60 | 8.10 | 6.14 | 2.45 | 2.73 | 4.97 | 4.40 | 4.23[a] | | |
| $G_4$ | C-mer | 5.39 | 7.44 | 5.96 | 2.10 | 2.38 | 4.79 | 4.33 | 4.11 | 4.24 | |
| | G-mer | 4.94 | 7.19 | 5.85 | 1.95 | 2.44 | 4.75 | 4.24 | 4.04 | 4.13 | |
| $C_5$ | C-mer | 6.06 | 7.90 | 6.25 | 2.06 | 2.33 | 4.62 | 4.28 | 3.97 | 4.18 | |
| $G_5$ | G-mer | | 7.56 | 5.37 | 1.86 | 2.19 | 4.61 | 4.12 | 3.95 | 4.08 | 10.80 |
| $C_6$ | C-mer | 5.44 | 7.49 | 5.51 | 1.72 | 2.23 | 4.58 | 3.73 | 3.77 | 3.86 | |
| $G_6$ | G-mer | | 7.82 | 5.96 | 2.62 | 2.45 | 4.72 | 3.04[b] | 3.31[b] | 3.88[b] | 10.74 |
| $G_7$ | C-mer | | 8.06 | 6.22 | 2.92 | 3.15 | 4.89 | 4.42 | 3.84 | 4.01 | 12.87 |
| | G-mer | | 8.14 | 6.20 | 2.83 | 3.07 | 4.89 | 4.47 | 3.99 | 4.18 | 12.75 |
| $T_8$ | C-mer | 1.52 | 7.52 | 6.27 | 2.22 | 2.66 | 4.91 | 4.32 | 4.21 | 4.32 | 13.75 |
| | G-mer | 1.50 | 7.46 | 6.19 | 2.21 | 2.66 | 4.90 | 4.31 | 4.21 | 4.33 | 13.85 |
| $T_9$ | C-mer | 1.79 | 7.49 | 5.96 | 2.08 | 2.45 | 4.91 | 4.21 | 4.14[a] | | 14.24 |
| | G-mer | 1.77 | 7.46 | 6.02 | 2.05 | 2.43 | 4.91 | 4.20 | 4.13[b] | 4.19[b] | 14.04 |
| $G_{10}$ | C-mer | | 8.00 | 6.22 | 2.71 | 2.42 | 4.75 | 4.26 | 4.16[a] | | 12.88 |
| | G-mer | | 7.99 | 6.21 | 2.68 | 2.41 | 4.74 | 4.23 | 4.12 | 4.22 | 12.80 |

[a]Degenerate 5′/5″ chemical shifts.
[b]Assigned from $^{13}$C HSQC data.

**Table 2.** Important NOE constraints and calculated distances for the loops of each hairpin

| d(CAACCCGTTG) | | | | | d(CAACGGGTTG) | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Proton | Proton to | NOESY Intensities | Constraint (Å) | Distance (Å) | Proton | Proton to | NOESY Intensities | Constraint (Å) | Distance (Å) |
| $C_4$ $H_{1'}$ | $C_5$ $H_{5'}$ | w w | 3.7-4.3 | 3.6 | $C_4$ $H_{1'}$ | $G_5$ $H_{5'}$ | m s | 2.5-3.1 | 2.5 |
| $C_4$ $H_{1'}$ | $C_5$ $H_{5''}$ | a w | 4.1-6.0 | 5.3 | $C_4$ $H_{1'}$ | $G_5$ $H_{5''}$ | w m | 3.3-3.9 | 4.0 |
| $C_4$ $H_{1'}$ | $C_5$ $H_6$ | m m | 2.9-3.5 | 2.9 | $C_4$ $H_{1'}$ | $G_5$ $H_{4'}$ | m s | 2.5-3.1 | 3.1 |
| $C_4$ $H_{2'}$ | $C_5$ $H_6$ | w w | 4.1-6.0 | 4.6 | $C_4$ $H_{1'}$ | $G_5$ $H_{3'}$ | a w | 4.1-6.0 | 4.9 |
| $C_4$ $H_{2''}$ | $C_5$ $H_6$ | m m | 2.9-3.5 | 3.0 | $C_4$ $H_{1'}$ | $G_5$ $H_8$ | w m | 3.3-3.9 | 4.1 |
| $C_4$ $H_{3'}$ | $C_5$ $H_6$ | a w | 4.1-6.0 | 5.1 | $C_4$ $H_{2'}$ | $G_5$ $H_8$ | w m | 3.3-3.9 | 3.6 |
| $C_4$ $H_{4'}$ | $C_5$ $H_5$ | w w | 4.1-6.0 | 6.4 | $C_4$ $H_{2''}$ | $G_5$ $H_8$ | s s | 2.2-2.8 | 2.2 |
| $C_4$ $H_{2''}$ | $C_5$ $H_5$ | w m | 3.3-3.9 | 3.1 | $C_4$ $H_{3'}$ | $G_5$ $H_8$ | w w | 3.7-4.3 | 4.3 |
| $C_4$ $H_{2'}$ | $C_5$ $H_5$ | a w | 4.1-6.0 | 4.2 | | | | | |
| $C_4$ $H_{1'}$ | $C_5$ $H_{3'}$ | a w | 4.1-6.0 | 6.0 | $G_5$ $H_{1'}$ | $G_6$ $H_{5'}$ | m m | 2.9-3.5 | 3.6 |
| | | | | | $G_5$ $H_{1'}$ | $G_6$ $H_{5''}$ | a w | 4.1-6.0 | 5.2 |
| $C_5$ $H_{1'}$ | $C_6$ $H_{4'}$ | w w | 3.7-4.2 | 4.6 | $G_5$ $H_{1'}$ | $G_6$ $H_{4'}$ | a w | 4.1-6.0 | 5.2 |
| $C_5$ $H_{1'}$ | $C_6$ $H_{5'}$ | w w | 3.7-4.2 | 3.5 | $G_5$ $H_{4'}$ | $G_6$ $H_{5'}$ | a w | 4.1-6.0 | 4.5 |
| $C_5$ $H_{1'}$ | $C_6$ $H_{5''}$ | a w | 4.1-6.0 | 5.1 | $G_5$ $H_{1'}$ | $G_6$ $H_8$ | w w | 3.7-4.3 | 4.4 |
| $C_5$ $H_{2'}$ | $C_6$ $H_{5'}$ | a w | 4.1-6.0 | 5.6 | $G_5$ $H_{2'}$ | $G_6$ $H_8$ | w m | 3.3-3.6 | 3.9 |
| $C_5$ $H_{2''}$ | $C_6$ $H_{5'}$ | a w | 4.1-6.0 | 4.6 | $G_5$ $H_{2''}$ | $G_6$ $H_8$ | s s | 2.2-2.8 | 2.5 |
| $C_5$ $H_{2''}$ | $C_6$ $H_{5''}$ | a w | 4.1-6.0 | 5.7 | $G_5$ $H_{3'}$ | $G_6$ $H_8$ | w w | 3.7-4.3 | 4.3 |
| $C_5$ $H_{1'}$ | $C_6$ $H_6$ | w m | 3.3-3.9 | 3.3 | | | | | |
| $C_5$ $H_{2'}$ | $C_6$ $H_6$ | w m | 3.3-3.9 | 4.6 | $G_6$ $H_{5'}$ | $G_7$ $H_{1'}$ | a w | 4.1-6.0 | 4.3 |
| $C_5$ $H_{2''}$ | $C_6$ $H_6$ | m m | 2.9-3.5 | 2.9 | $G_6$ $H_{5'}$ | $G_7$ $H_{4'}$ | a w | 4.1-6.0 | 5.9 |
| $C_5$ $H_{1'}$ | $C_6$ $H_5$ | m w | 3.3-3.9 | 3.7 | $G_6$ $H_{5''}$ | $G_7$ $H_{4'}$ | w w | 3.7-4.3 | 4.1 |
| $C_5$ $H_{2'}$ | $C_6$ $H_5$ | m w | 3.3-3.9 | 3.7 | $G_6$ $H_{4'}$ | $G_7$ $H_{4'}$ | a w | 4.1-6.0 | 5.2 |
| $C_5$ $H_{2''}$ | $C_6$ $H_5$ | m m | 2.9-3.5 | 2.5 | | | | | |
| $C_5$ $H_{3'}$ | $C_6$ $H_5$ | a w | 4.1-6.0 | 4.9 | $G_6$ $H_{5'}$ | $T_8$ $H_{4'}$ | a w | 4.1-6.0 | 6.1 |
| | | | | | $G_6$ $H_{5''}$ | $T_8$ $H_{4'}$ | a w | 4.1-6.0 | 5.4 |
| $C_6$ $H_{4'}$ | $G_7$ $H_{4'}$ | a w | 4.1-6.0 | 4.1 | | | | | |

a=absent, vw=very weak, m=medium, vs=very strong

base ($G_7$ in this case). These up-field shifts may serve as an excellent marker, among others, for the sharp backbone turn at the 3′–5′ loop–stem junction of mini-hairpin loop structures (12,13,15,18,19).

The NOEs also show features of unusual torsion angles and base stacking. The important inter-residue NOEs and their intensities are summarized in Table 2. There are no NOEs between nucleotides $G_6$ and $C_4$ in the G-mer or between $C_6$ and $C_4$ of the C-mer, even at long mixing times (500 ms), which would support the existence of an **L2** structure (13,15).

## χ Torsion angles ($H_{1'}$–$C_{1'}$–$N_{\{9\ or\ 1\}}$–$C_{\{8\ or\ 6\}}$)

The glycosidic torsion angle χ is estimated from the internuclear distance from the $H_5$, $H_6$, or $H_8$ base proton to the $H_{1'}$ sugar proton. The other base proton to sugar proton NOEs may be significantly affected by spin diffusion. For χ torsion angles of *anti* rotamers (100–180°), weak NOE intensities between the $H_{1'}$ sugar proton and base proton are expected at moderate mixing times (75–150 ms). Strong NOEs are expected even at very short mixing times (50 ms) for χ in the *syn* orientation (0 ± 20°). For all of the bases of the G-mer and the C-mer, weak NOEs were observed between the base protons and the $H_{1'}$ sugar proton, showing that the χ torsion angles are in the *anti* range (Fig. 4a for the C-mer). These results rule out the **L1syn$_1$** and **L1syn$_2$** models for mini-hairpin loop folding for both decamers.

## Sugar pucker

Constraints for the sugar puckers were determined from both the cross-peak intensities of the DQF-COSY and the quantitatively measured coupling constants from a PE-COSY. This analysis only gives an estimate of the actual sugar pucker. It is well documented that the sugars of DNA are in fast equilibrium with other sugar conformations. Thus, only an average conformation may be determined by NMR (38,39). The $H_{2'}$–$H_{1'}$ and $H_{2''}$–$H_{1'}$ coupling constants were measured quantitatively from the PE-COSY spectrum (data not shown). All of the $^3J_{(H2'-H1')}$ constants (7–10 Hz) were greater than the $^3J_{(H2''-H1')}$ coupling constants (4.8–7 Hz), which limits the pseudorotational phase angle from 90° to 190° (39). The $H_{3'}$–$H_{2'}$, $H_{3'}$–$H_{2''}$ and $H_{3'}$–$H_{4'}$ coupling constants could not be directly measured due to cross-peak patterns complicated by small coupling constants, other passive couplings, poor resolution and overlap from neighboring protons. Since all of the $H_{3'}$–$H_{2'}$ and $H_{3'}$–$H_{2''}$ cross-peak intensities in the phase-sensitive DQF-COSY and PE-COSY were weak, indicating small coupling constants, all of the sugars were constrained to the 2′-*endo* conformational family (180–126°, $^2_3$T, $^2$E, $^2$T, $E_1$).

## γ Torsion angles (O–$C_{5'}$–$C_{4'}$–$C_{3'}$)

In principle the γ torsion angle can be determined from the heteronuclear J-couplings between the phosphorus and $H_{5'}$/$H_{5''}$ atoms and and between $H_{4'}$ and $H_{5'}$/$H_{5''}$(40). When γ is in the *gauche*$^+$ position, the $H_{4'}$–$H_{5'}$ and $H_{4'}$–$H_{5''}$ coupling constants are almost identical. The TOCSY cross-peaks between those protons are of the same intensity at moderate mixing times (25 ms) as well (15). Since this region of the spectrum is very crowded, one may need to rely on additional information from NOESY spectra. The NOE cross-peak to $H_{3'}$ and $H_{5'}$ is ~3.7 Å and the distance between $H_{3'}$ and $H_{5''}$ protons is ~2.5 Å. Normally only very weak NOEs are observed between the $H_{5'}$/$H_{5''}$ and the $H_8$/$H_6$ or the $H_{2'}$/$H_{2''}$ protons. In practice, however, the chemical shift differences between $H_{4'}$, $H_{5'}$ and $H_{5''}$ are very small and it is often not possible to assign them in moderately sized DNA oligomers. According to Kim *et al.* (41) the γ torsion angle could be loosely constrained by using the $\Sigma J_{\{H4'\}}$ coupling, where $\Sigma J_{\{H4'\}} = J_{(H3'-H4')} + J_{(H4'-H5')} + J_{(H4'-H5'')}$ and Figure 3 could then be used to approximate the torsion angle. $\Sigma J_{\{H4'\}}$can be measured from the $H_{3'}$–$H_{4'}$ cross-peak of the phase-sensitive DQF-COSY by calculating the peak-to-peak separation in the $H_{4'}$ dimension.

When γ is in the *trans* configuration, the coupling constant for $H_{4'}$–$H_{5'}$ is weaker than the coupling constant for $H_{4'}$–$H_{5''}$ (11 versus 2 Hz). Also, $\gamma^{tr}$ is associated with some characteristic NOE cross-peak intensities. The most important of these are NOEs from the $H_{5'}$ to its own base proton, which is approximately the same intensity as the base proton to its $H_{1'}$ sugar proton (about 3.3 Å). The $H_{5''}$ proton may also have an NOE to its base proton. Similar NOEs may also be observed for the γ *gauche*$^-$ configuration. Fortunately, $\gamma^-$ can be distinguished from $\gamma^{tr}$ in the $H_{5'}$/$H_{5''}$–$H_{2'}$/$H_{2''}$ region of the NOESY spectrum . For the $\gamma^{tr}$ torsion, the $H_{5'}$–$H_{2'}$ cross-peak is more intense than the $H_{5'}$–$H_{2''}$ cross-peak; the converse is true for the $\gamma^-$ torsion. A $\Sigma J_{\{H4'\}}$ value in the range 14–23 Hz can also be used to support a $\gamma^{tr}$ conformation (41).

In the C-mer, all of the nucleotides excluding $G_7$ have $\Sigma J_{\{H4'\}}$ in the range 6.1–7.4 Hz. No strong $H_{5'}$/$H_{5''}$ NOEs to base protons or $H_{2'}$ protons were observed at moderate mixing times, except those from $G_7$, which allows us to constrain all of the bases except $G_7$ to the $\gamma^+$ rotamer (60 ± 30°). As in the G-mer, the only base proton that shows an NOE to its $H_{5'}$/$H_{5''}$ are $G_7$ and $C_1$ (Fig. 4b). The $H_{5'}$–$H_8$ NOE intensity is of the same magnitude as the other $H_8$/$H_6$–$H_{1'}$ cross-peaks. The NOE pattern in $H_{2'}$/$H_{2''}$–$H_{5'}$/$H_{5''}$ is consistent with that of a $\gamma^{tr}$ torsion for $G_7$.

As in the C-mer hairpin, the loop-closing nucleotide of the G-mer, $G_7$, has the γ torsion angle in the *trans* conformation. Both

oligomers show an NOE between the $H_{4'}$ of the sixth nucleotide and $H_{4'}$ of $G_7$. Although weak, it lends further support to the unusual $\gamma^{tr}$ torsion for the $G_7$ nucleotide (Table 2). These results show that there is only one $\gamma^{tr}$ torsion angle, which excludes the theoretical models **L1anti$_1$** and **L1anti$_2$**, which have two $\gamma^{tr}$ torsion angles in the loop.

### β-(P–O5′–C5′–C4′) and ε-(P{*n* – 1}–O3′–C3′–C4′)

Heteronuclear experiments are necessary to derive constraints for the β and ε torsion angles. The coupling between phosphorus and $H_{3'}$ can be measured directly by a selective heteronuclear proton–phosphorus COSY where only the $H_{3'}$ protons are excited. This simplifies cross-peak patterns by not perturbing protons that are coupled to the $H_{3'}$ sugar protons. Even though this coupling constant may be measured directly, there are four possible solutions to the modified Karplus equation (31,42,43).

Ideally the β torsion angles can be constrained by measuring the $H_{5'}$–P, $H_{5''}$–P and $H_{4'}$–P coupling constants. However, unlike the $H_{3'}$ protons, the $H_{4'}$, $H_{5'}$ and $H_{5''}$ were not well resolved and could not be excited selectively. Structural information may, however, be estimated from the intensity of the $^1H$-$^{31}P$ cross-peaks. When the $H_{5'}$–P cross-peak is of equal intensity to the $H_{5''}$–P cross-peak then the β torsion angle is *trans*. When the $H_{5'}$–P cross-peak is much less intense than the $H_{5''}$–P cross-peak then β is in the *gauche*$^+$ range. When the $H_{5'}$–P cross-peak is much more intense than the $H_{5''}$–P cross-peak then β may be roughly constrained to the *gauche*$^-$ range. However, this requires stereospecific assignment of the geminal $H_{5'}/H_{5''}$ protons, as well as highly resolved proton–phosphorus spectra.

From the selective $^1H$-$^{31}P$ COSY of the C-mer, the $H_{3'}$–phosphorus coupling constants were all determined to be 5–8 Hz. It was also observed that the $H_{5'}$–phosphorus cross-peaks are very weak for nucleotides $C_5$ and $G_7$ and absent for nucleotide $C_6$. The $H_{5''}$–phosphorus cross-peaks are very strong, especially that of $G_7$. These cross-peak patterns indicate that nucleotides $C_5$, $C_6$ and $G_7$ are in the $\beta^+$ torsional range.

From the selective $^1H$-$^{31}P$ experiment of the G-mer, the $H_{3'}$–P cross-peak coupling constants were determined to be 2–8 Hz. From the non-selective experiment, the $G_6$–$H_{5''}$ to $G_6$–P cross-peak is of medium intensity and the $G_6$–$H_{5'}$ to $G_6$–P cross-peak is absent. From these data, the β torsion angle for $G_6$ was estimated to be in the $0 \pm 60°$ range. Unfortunately, other nucleotides exhibited poor spectral resolution, including $G_7$, which was expected to show unusual features.

One may resolve some of the ambiguities of β and ε torsion angle constraints through the use of the $C_{4'}$–$H_{4'}$ cross-peak of the $^{13}C$ HSQC spectrum (27). This cross-peak displays the passive coupling of $C_{4'}$–$P_{(n+1)}$, reflecting the ε torsion angle, along with $C_{4'}$–$P_n$, reflecting the β torsion angle, as shown in Figure 5. In standard B-type DNA both of these angles are expected to be in the *trans* rotamer. This results in a $C_{4'}$–P coupling constant near 20 Hz.

The heteronuclear chemical shift assignments are summarized in Table 3. The nucleotides of the stem region in both hairpins gave $C_{4'}$–P coupling constants in the range 19–22 Hz. The terminal nucleotides, $C_1$ and $G_{10}$ gave $C_{4'}$–P coupling constants in the range 8–10 Hz. These data allow the backbone torsion angles ε and β to be constrained in the *trans* conformation ($180 \pm 30°$). In the loop region of both hairpins the P–$C_{4'}$ coupling constants are comparatively small (2–8 Hz), as shown in Table 4.

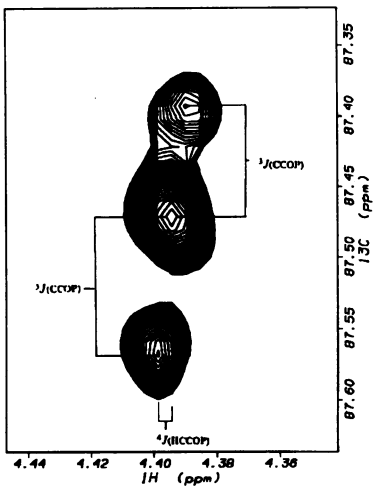


**Figure 5.** A representative $H_{4'}$–$^{13}C_{4'}$ cross-peak showing the useful passive coupling from phosphorus to carbon.

**Table 3.** Heteronuclear chemical shifts for d(CAACCCGTTG) and d(CAACGGGTTG) hairpins

| Base | | C1′ | C2′ | C3′ | C4′ | C5′ | C5 | P |
|---|---|---|---|---|---|---|---|---|
| C1 | C-mer | 87.78 | 80.08 | 77.79 | 88.23 | | 90.68 | |
| | G-mer | 87.80 | 80.08 | 77.64 | 88.18 | | 90.65 | |
| A2 | C-mer | 84.71 | 79.90 | 79.5 | 87.47 | 66.87 | | 2.86 |
| | G-mer | 84.66 | 80.04 | 79.54 | 87.49[d] | 66.91 | | 2.81 |
| A3 | C-mer | 84.47 | 80.62 | 77.47 | 86.64 | 66.90 | | 2.87 |
| | G-mer | 84.40 | 80.82 | c | 86.67[d] | 67.48 | | 2.79 |
| C4 | C-mer | 86.08[a] | 79.55 | 79.02 | 85.66[b] | 68.23 | 91.53 | 2.85 |
| | G-mer | 86.50 | 77.79 | 79.26 | 86.02 | 67.09 | c | 2.85 |
| C5 | C-mer | 88.89 | 8.80 | 79.49 | 87.58 | 67.38 | 91.03 | 3.07 |
| G5 | G-mer | 85.89 | absent | 79.62 | 86.97 | 67.12 | | 1.83 |
| C6 | C-mer | 89.23 | 83.34 | 80.56 | 86.89 | 66.49 | 92.67 | 2.22 |
| G6 | G-mer | 85.32 | 80.47 | c | 87.04 | 66.60 | | 1.90 |
| G7 | C-mer | 84.67[a] | 77.11 | 79.61 | 88.43 | 65.49 | | 2.67 |
| | G-mer | 84.44 | 78.72 | 79.74 | 88.73 | 64.14 | | 2.81 |
| T8 | C-mer | 85.64 | 88.84 | 77.22 | 85.89[b] | 67.75[b] | | 2.10 |
| | G-mer | 85.64 | 79.12 | c | 85.67 | 67.77 | | 2.29 |
| T9 | C-mer | 85.53[a] | 88.84 | 76.96 | 85.57 | 67.45 | | 2.57 |
| | G-mer | 85.37 | 79.12 | c | 85.37 | 67.62 | | 2.69 |
| G10 | C-mer | 84.75[a] | 81.57 | 76.64 | 87.88 | 67.38 | | 3.02 |
| | G-mer | 84.72 | 81.62 | 76.82 | 87.84 | 68.58 | | 2.92 |

[a]Degenerate 1′ proton shifts. Assignments are based on the d(CAACGGGTTG) $^{13}C$ data.
[b]Degenerate 4′ proton shifts.Assignments are based on the d(CAACGGGTTG) $^{13}C$ data.
[c]Under residual HOD peak.
[d]Degenerate chemical shifts; assignments are based on the d(CAACGGGTTG) $^{13}C$ HSQC results.

These couplings indicate unusual backbone conformations for both loops.

The G-mer HSQC spectrum shows that the coupling constants for $G_5$, $G_6$ and $G_7$ are 6.0, 2.0 and 5.0 Hz respectively, indicating unusual β and ε torsion angles. Comparing the ε torsion angles obtained from the $H_{3'}$–P coupling constants and the $C_{4'}$–$^{31}P$ coupling constant may narrow the range of possible ε torsion angles. In addition, if the β torsion angle is estimated from the $^1H$-$^{31}P$ COSY, the $C_{4'}$–P coupling constant range may be predicted and ε can be estimated. For a detailed account of how β and ε torsion angles were estimated the reader is referred to

**Table 4.** Heteronuclear coupling constraints for the d(CAACCCGTTG) and d(CAACGGGTTG) hairpins

| C- mer | | | | | | | G- mer[a] | | |
|---|---|---|---|---|---|---|---|---|---|
| Base | $C_{2'}$-P | $C_{3'}$-P | $H_{3'}$-P[b] | $C_{4'}$-P[i] | $H_{4'}$-P | $H_{3'}$-P[c] | Base | $C_{4'}$-P[i] | $H_{3'}$-P[c] |
| $C_1$ | <2 | 3.34 | 2.41 | 8.7[e] | <1 | 5.7 | $C_1$ | 8.0 | 5.4 |
| $A_2$ | <2 | 1.9 | 5.46 | 22.5 | 4.23 | 5.3 | $A_2$ | 19.0 | 4.0d |
| $A_3$ | c | c | c | 20.4 | 5.46 | 5.53 | $A_3$ | 21.0 | 5.6 |
| $C_4$ | <2 | 1.8 | 5.5 | 21.7[g] | 3.9f | 5.1 | $C_4$ | 20.0 | 6.2 |
| $C_5$ | <2 | 5.1 | 5.1 | 7.0[i] | d | 7.4 | $G_5$ | 6.0[i] | 5.0[b] |
| $C_6$ | <2 | 6.1 | 8.7 | 8.3[i] | 3.57 | 7.9 | $G_6$ | 2.0[hj] | 6.7 |
| $G_7$ | 2.6 | e | e | 3[hj] | 3.33[h] | 5.1 | $G_7$ | 5.0[i] | 2.0[b] |
| $T_8$ | <2 | e | e | 22.0[gi] | 1.3[h] | 5.3 | $T_8$ | 22.0 | 6.0 |
| $T_9$ | <2 | e | e | 21 | 2.2 | 5.3 | $T_9$ | 20.0 | 6.0 |
| $G_{10}$ | 3.0 | f | f | 9.8[e] | 3.53 | f | $G_{10}$ | 9.0 | f |

[a]Signal to noise was not sufficient to measure all coupling constants.
[b]From $^{13}C$ HSQC experiment.
[c]From selective $^1H_{3'}$–$^{31}P$ COSY.
[d]Could not be determined.
[e]Under HOD resonance.
[f]No phosphate at the 5′ and 3′ terminus.
[g]Isochronous chemical shifts. $C_4$ and $T_8$ couplings may be reversed.
[h]Coupling constant estimated from line width.
[i]Sum of coupling constants $H_{4'}(n)$ – P($n$) and $H_{4'}(n)$ – P ($n$+1).
[j]Only two peaks observed instead of three in the coupling pattern.

Avizonis (44). The β torsion angles for the loop were constrained to 0 ± 120° for $G_5$ and $G_7$, while $G_6$ was set to 0 ± 60°. The ε torsion angles for the loop were constrained to 0 ± 120° for $G_5$, –86 ± 20° for $G_6$ and 0 ± 100° for $G_7$.

The β and ε torsion angles of the C-mer were constrained using the method as used for the G-mer. The β torsion angles for nucleotides $C_5$, $C_6$ and $G_7$ were all constrained to 0 ± 120°. The ε torsion angles were constrained to 0 ± 120° for nucleotides $C_5$ and $C_6$, while $G_7$ was constrained to –76 ± 30°.

## MOLECULAR MODELING

Starting structures were constructed with the general features of a double helical stem consisting of four Watson–Crick base pairs, either two guanines or two cytosines forming a loop (with various conformations and orientations where the loop nucleotides are stacked or unstacked), γ torsion angle of nucleotide $G_7$ in the *trans* conformation and all sugar puckers in the 2′-*endo* conformation. Four starting structures for each decamer were built using InsightII (Biosym Technologies Inc.). From these four different starting structures, 40 hairpin structures were generated for both decamers by a simulated annealing protocol using the program XPLOR (Brünger, 1992). In both cases 30 of the 40 structures were within acceptable agreement with the experimental data. An average structure was calculated for each of the hairpins from the 30 superimposed structures. The superposition of these 30 structures are shown in Figure 6, along with plots of the RMS deviations from the average structure for all atoms excluding protons. The C-mer calculations included 200 NOE-derived distance constraints and 78 dihedral constraints, while the G-mer calculations included 183 NOE-derived distance constraints and 79 dihedral angle constraints. There were no NOE violations greater than 0.5 Å or dihedral angle constraint violations greater than 10° for either average hairpin structure. Intra-sugar NOE data were not used in the restrained molecular dynamics simulations, since their cross-peak intensities may be significantly affected by spin diffusion. Instead, data from dihedral angle constraints were used to calculate the sugar conformations. The important NOE constraints in both mini-hairpin loops are tabulated in Table 2. The RMS deviations are large within the loops of the hairpins (±16 to ±140° for the loop backbone dihedral angles and ±1 to ±7° for the stem). This feature is due to the fact that the dihedral angles could not be well constrained, as described above. Frequently these angles fell into two or more

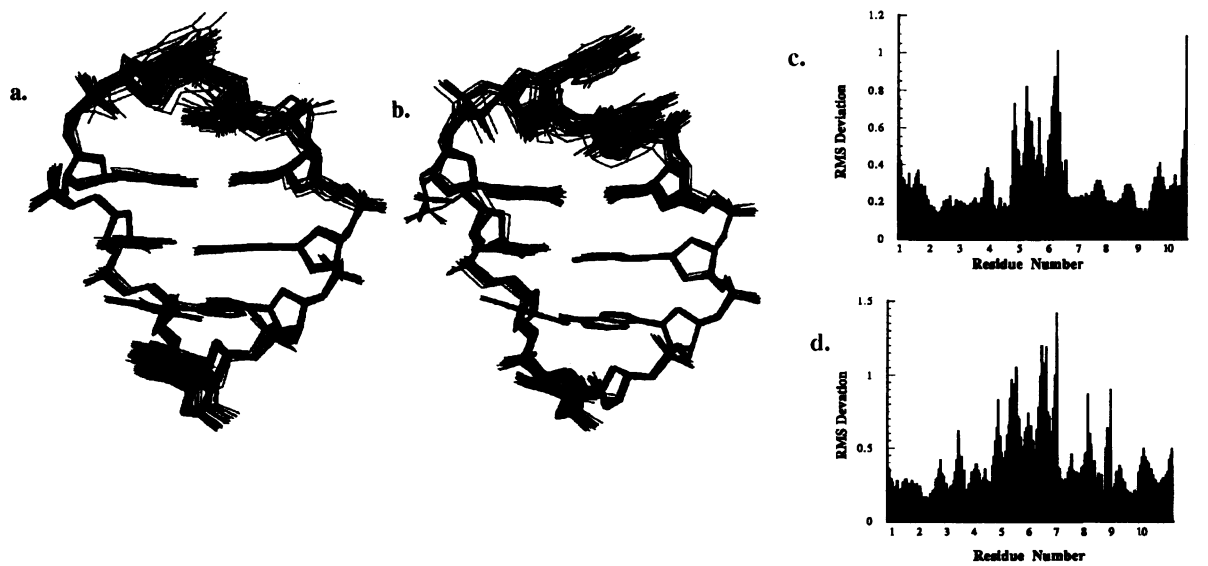


**Figure 6.** The superposition of 30 acceptable structures for (**a**) the C-mer and (**b**) the G-mer hairpins. The RMS deviations of the superimposed structures from the average structure for (**c**) the C-mer and (**d**) the G-mer are shown excluding the hydrogen atoms. The hydrogen atoms are not shown for clarity.

structural families with equivalent energies and equal agreement with experimental data. For example, in the G-mer, the β torsion angle of $G_6$ adopts both *gauche*$^+$ and *gauche*$^-$ orientations with no difference in structural energy. For this reason it has a very high RMS value. In most instances, one structural family could not be chosen over another. Due to the small number of NOE distance constraints and large range of values for the dihedral angle constraints, the loop region of the G-mer and C-mer hairpins are not well determined. However, the superposition of structures (Fig. 6) clearly shows the overall fold of the hairpins.

In our previous study (1) it was shown that the C-mer hairpin has a 50 ± 1°C melting point, slightly higher than the G-mer hairpin melting point, 47 ± 1°C, under equivalent conditions of DNA concentration, salt concentration and pH value. This difference in melting point also has a corresponding difference in denaturation enthalpy. The enthalpy of denaturation ($\Delta H_{HC}$) for the C-mer hairpin was found to be 29 ± 2 kcal/mol, with 23 ± 2 kcal/mol for the G-mer hairpin. Since these decamers only differ in the composition of the loop, it was concluded that the stacking of the guanines in the G-mer loop must be less favorable than the stacking in the C-mer loop. To examine this point further, views of both loops (residues 4–6) are shown in Figure 7a–d. Figure 7a and b show the superposition of structures. As indicated by the RMS deviations in both structures (Fig. 6), Figure 7a and b shows that even though the superposition of loop structures for the C-mer and G-mer do not yield a well-defined structure, the stacking of the loop nucleotides is significantly different for the C-mer and G-mer hairpin loops. The C-mer hairpin nucleotides $C_4$, $C_5$ and $C_6$ appear to stack more or less in the same way as they would in B-DNA (Fig. 7e). In contrast, the G-mer nucleotides $C_4$, $G_5$ and $G_6$ appear to stack almost directly on top of each other, in contrast to the stacking found in a similar 5′-CGG piece of B-DNA (Fig. 7f). The local helical twist angle between successive bases defined by the projection of the successive $C_{1'}$–N vectors onto the plane perpendicular to the helix axis (45,46) were calculated. This type of helical twist has been reported to be in the range 30–40° for B-DNA (47). As expected, the stem regions of both hairpins show helical twist angles in the B-DNA range. The twists between bases in the C-mer loop do not deviate from this range except for a larger twist between nucleotides $C_5$ and $C_6$ (55°). The twist between nucleotide bases of the G-mer show much greater helical twists between nucleotides $C_4$ and $G_5$ (75°), $G_5$ and $G_6$ (56°) and $G_6$ and $G_7$ (72°). The up-field chemical shifts of the $H_8$ protons of $G_5$ and $G_6$ indicates an unusual stacking of the purines supporting these molecular dynamics results. Since the stacking of bases in B-DNA is the energetically favorable conformation, it may be concluded that since the stacking of the guanine residues in the G-mer hairpin loop do not assume the most stable orientation the G-mer hairpin is a slightly less stable hairpin conformation compared with the C-mer hairpin conformation. These results are also supported by the thermodynamic data derived in our previous study (1).

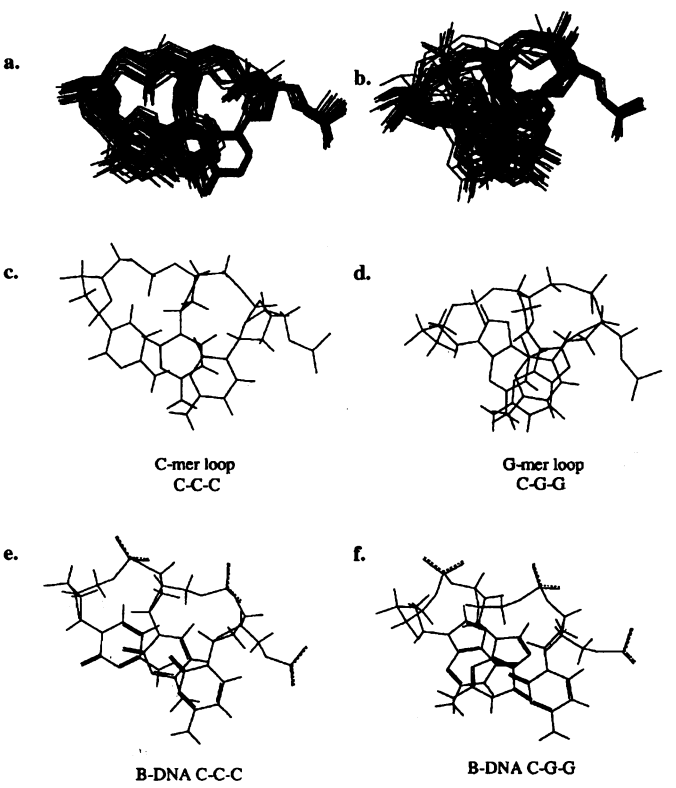


**Figure 7.** A view of nucleotides 4–6 showing the stacking of the hairpin loops. (**a**) The superposition of all 30 structures for the C-mer. (**b**) The superposition of all 30 structures for the G-mer. (**c**) The average structure of the C-mer. (**d**) The average structure of the G-mer. (**e**) A 5′-CCC trinucleotide showing the stacking of cytosines in B-DNA. (**f**) A 5′-CGG trinucleotide showing the stacking of two guanines in B-DNA. The hydrogen atoms are not shown for clarity.

## CONCLUSIONS

The hairpin forms of the two decamers d(CAACCCGTTG) and d(CAACGGGTTG) each consist of a stem region composed of four Watson–Crick base pairs and a loop formed by the two remaining central nucleotides. Through the use of NOE-derived distance constraints, proton–proton, proton–phosphorus and carbon–phosphorus coupling constants, dihedral angle constraints were derived and used in molecular dynamics calculations. Both C-mer and G-mer hairpins showed a number of unusual backbone coupling constants, as measured by a $^1H$-$^{13}C$ HSQC experiment. The $C_{4'}$–phosphorus coupling provided constraints for both β and ε torsion angles in the stems of the hairpins. The same coupling also showed that the β and ε torsion angles for nucleotides 5–7 were not in the *trans* conformation. Unfortunately, the actual values of β and ε torsion angles in the loops could not be obtained. Instead, a range of possible angles was used to restrain the backbone in the loop. In addition to the 3′-end loop-closing guanine $\gamma^{tr}$ torsion angle described by others (12,14,15,48), here we report that the 3′-end loop-closing guanine β torsion angle need not necessarily be in the *gauche*$^+$ conformation. In the loop of the C-mer the β and ε torsion angles of $G_7$ were found to adopt *gauche*$^-$ conformations, while the derived β and ε torsion angles of $C_6$ conform to *gauche*$^+$ orientations. The loop of the G-mer is less well defined. Like other mini-hairpin loops, the $G_7$ nucleotide adopts a $\gamma^{tr}$ conformation. The β torsion angle of $G_6$ adopted a *gauche*$^+$ conformation, while its ε dihedral angle may be in either the *gauche*$^+$ or the *gauche*$^-$ rotamer. No conclusions could be made for the β and ε torsion angles of $G_7$, $G_5$ or $C_5$ of the C-mer hairpin, because of the lack of a sufficient number of NOE distance constraints and the large range of possible backbone torsion angles. The stacking of the two guanines in the G-mer loop was found to have large

helical twist angles compared with guanine stacking in B-DNA, whereas the stacking of cytosines in the C-mer were found to be similar to cytosine stacking in B-DNA. In general, both hairpins adopt an **L1**-type fold. These results may be considered to be in contrast to previous predictions (15) that the pyrimidine-containing loop should form an **L2** fold with $C_6$ folding back into the minor groove. At this date, however, there have not been enough mini-hairpin structures solved with different loop and different stem sequences to be able to accurately predict the loop folding. The mini-hairpin loop d($C_1G_2C_3T_4A_5G_6C_7G_8$) forms an **L2** loop, with the thymine forming three hydrogen bonds in the minor groove with nucleotides $G_2$ ($T_4$–$O_4$ to $H_2$), $G_6$ ($T_4$–$O_2$ to $H_2$) and $C_7$ ($T_4$–$H_3$ to $O_2$) (13,15). No similar hydrogen bonds are likely to form in the C-mer hairpin, since it has cytosines in the loop and A–T pairs in the stem. Therefore, the **L1** conformation is favored.

Other studies (49–51) have implied that hairpins whose loops are formed by purines are less stable then those whose loops are formed by pyrimidines, because of the inability of purines to adopt a favorable stacking conformation in a restrained loop. However, these conclusions were not substantiated by structural evidence. In a previous study (1) we showed that the C-mer hairpin was slightly more thermodynamically stable than the G-mer hairpin, consistent with the literature. In this study we provide structural evidence of how base stacking in loops containing either purines or pyrimidines may influence the stability of the hairpin structure.

## ACKNOWLEDGEMENTS

We thank Dr Victor Hsu for useful discussions and editing this manuscript. This work was supported by grants from the American Cancer Society (CH-32) and NSF grant GM-35177. We are grateful to the NIH (1S01 RR03342) and the NSF for providing resources for the acquisition of the 500 MHz spectrometer.

## REFERENCES

1 Avizonis,D.Z. and Kearns,D.R. (1995) *Biopolymers*, **35**, 187–200.
2 Hilbers,C.W., Haasnoot,C.A.G., de Bruin,S.H., Joordens,J.J.M., van der Marel,G.A. and van Boom,J.H. (1985) *Biochemie*, **67**, 685–695.
3 Haasnoot,C.A.G., Hilbers,C.W., van der Marel,G.A., van Boom,J.H., Singh,U.C., Pattabiraman, N., and Kollman, P.A. (1986) *J. Biomol. Struct. Dynam.*, **3**, 843–857.
4 Germann,M.W., Kalisch,B.W., Lundberg,P., Vogel,H.J. and van de Sande,J.H. (1990) *Nucleic Acids Res.*, **18**, 1489–1498.
5 Roy,S., Weinstein,S., Borah,B., Nickol,J., Appella,E., Sussman,J.L., Miller,M., Sindo,H. and Cohen,J.S. (1986) *Biochemistry*, **25**, 7417–7422.
6 Wemmer,D.E., Chou,S.H., Hare,D.R. and Reid,B.R. (1985) *Nucleic Acids Res.*, **13**, 3755–3772.
7 Nadeau,J.G. and Gilham,P.T. (1985) *Nucleic Acids Res.*, **13**, 8259–8274.
8 Summers,M.F., Byrd,R.A., Gallo.,K.A., Samson,C.J., Zon,G. and Eagen,W. (1985) *Nucleic Acids Res.*, **13**, 6375–6386.
9 Williamson,J.R. and Boxer,S.J. (1989) *Biochemistry*, **28**, 2819–2831.
10 Williamson,J.R. and Boxer,S.J. (1989) *Biochemistry*, **28**, 2831–2836.
11 Ikuta,S., Chattopadhyaya,R., Hirataka,I., Dickerson,R.E. and Kearns,D.R. (1986) *Biochemistry*, **25**, 4840–4849.
12 Orbons,L.P.M., van der Marel,G.A., van Boom,J.H. and Altona,C. (1986) *Nucleic Acids Res.*, **14**, 4187–4195.
13 Pieters,J.M.L., de Vroom,E., van der Marel,G.A., van Boom,J.H., Koning,T.M.G., Kaptien,R. and Alona,C. (1990) *Biochemistry*, **29**, 788–799.
14 Orbons,L.P.M., van Beuzekon,A.A. and Altona,C. (1987) *J. Biomol. Struct. Dynam.*, **4**, 965–987.
15 Ippel,J.H., Lanzotti,V., Galeone,A., Mayol,L., van den Boogaart,J.E., Pikkemaat,J.A. and Altona,C. (1991) *J. Biomol. Struct. Dynam.*, **9**, 821–836.
16 Sakata,T., Hiroaki,H., Oda,Y., Tanaka,T., Ikehara,M. and Uesugi,S. (1990) *Nucleic Acids Res.*, **18**, 3831–3839.
17 Cheong,C., Varani,G. and Tinoco,I. (1991) *Biochemistry*, **30**, 3280–3289.
18 Orbons,L.P.M., van der Marel,G.A., van Boom,J.H. and Altona,C. (1987) *Eur. J. Biochem.*, **170**, 225–239.
19 Orbons,L.P.M., van der Marel,G.A., van Boom,J.H. and Altona,C. (1987) *J. Biomol. Struct. Dynam.*, **4**, 939–936.
20 Howard,F.B., Chen,C.,Ross,P.D and Miles,H.T. (1991) *Biochemistry*, **30**, 779–782.
21 Raghunathan,G., Jernigan,R.L., Miles,H.T. and Sasisekharan,V. (1991) *Biochemistry*, **30**, 782–788.
22 Blommers,M.J.J., Walters,J.A.L.I., Haasnoot,C.A.G., Aelen,J.M.A., van der Marel,G.A., van Boom,J.H. and Hilbers,C.W. (1989) *Biochemistry*, **28**, 7491–7498.
23 Marion,D. and Wüthrich,K. (1983) *Biochem. Biophys. Res. Commun.*, **113**, 967–974.
24 Clore,G.M., Kimber,B.J. and Gronenborn,A.M. (1983) *J. Magn. Res.*, **54**, 170–172.
25 Hasnoot,C.A.G. and Hilbers,C.W. (1983) *Biopolymers*, **22**, 1259–1266.
26 Marion,D. and Bax,A. (1988) *J. Magn. Res.*, **80**, 528–533.
27 Schmieder,P., Ippel,J.H., van den Elst,H., van der Marel,G.A., van Boom,J.H., Altona,C. and Kessler,H. (1992) *Nucleic Acids Res.*, **20**, 4747–4751.
28 Shaka,A.S. and Freeman,R. (1983) *J. Magn. Reson.*, **53**, 169–172.
29 Bax,A. and Davis,D.G. (1985) *J. Magn. Reson.*, **65**, 355–360.
30 Sklenar,V. and Bax,A. (1986) *FEBS Lett.*, **208**, 4176–4178.
31 Sklenar,V. and Bax,A. (1987) *J. Am. Chem. Soc.*,**109**, 7525–7526.
32 Bodenhausen,M.J.J., Haasnoot,C.A.G., Hilbers,C.W., van Boom,J.H. and van der Marel,G.A. (1987) *Structure and Dynamics of Biopolymers*. NATO ASI ser. E., Vol. 113, pp. 78–91.
33 Bodenhausen,M.J.J. and Ruben,D.J. (1980) *Chem. Phys. Lett.*, **69**, 185–189.
34 Norwood,T.J., Boyd,J., Heritage,J.E., Soffe,N. and Campell,I.D. (1990) *J. Mag. Reson.*, **87**, 488–501.
35 Haasnoot,C.A.G., Westerink,H.P., van der Marel,G.A. and van Boom,J.H. (1983) *J. Biomol. Struct. Dynam.*, **1**, 131–149.
36 Kearns,D.R., Assa-Munt,N., Behling,R.W., Early,T.A., Feigon,J., Granot,J., Hillen,W. and Wells,R.D. (1981) In Sarma,R.H. (ed.), *Biomolecular Stereodynamics*. Adenine Press, New York, NY, Vol. I, pp. 345–365.
37 Wuthrich,K. (1986) In *NMR of Proteins and Nucleic Acids*. John Wiley and Sons, New York, NY.
38 Hosur,R.V., Govil,G. and Miles,H.T. (1988) *Magn. Resonance Chem.*, **26**, 927–944.
39 Majumdar,A. and Hosur,R.V. (1992) *Prog. Nucl. Magn. Spect.*, **24**, 109–158.
40 Altona,C. (1982) *Recl. Trav. Chim. Pays-Bas*, **101**, 413–433.
41 Kim,S.G, Lin,L. and Reid,B.R. (1992) *Biochemistry*, **31**, 3564–3574.
42 Lankhorst,P.P., Haasnoot,C.A., Erkelens,C. and Altona,C. (1984) *J. Biomol. Struct. Dynam.*, **1**, 1387–1405.
43 Nikonowicz,E.P. and Gorenstein,D.G. (1990) *Biochemistry*, **29**, 8845–8858.
44 Avizonis,D.Z. (1993) Doctoral Dissertation, University of California at San Diego, La Jolla, CA.
45 Dickerson,R.E. (1989) *EMBO J.*, **8**, 1–4.
46 Lavery,R. and Sklenar,H. (1988) *J. Biomol. Struct. Dynam.*, **6**, 63–91.
47 Dickerson,R.E. and Drew,H.R. (1981) *J. Mol. Biol.*, **149**, 761–786.
48 Wickstrom,E. and Tinoco,I. (1974) *Biopolymers*, **13**, 2367–2383.
49 Senior,M.M., Jones,R.A. and Breslauer,K.J. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 6242–6246.
50 Xodo,L.E., Manzini,G., Quadrifolio,F., Yathindra,N., van der Marel,G.A. and van Boom,J.H. (1988) *Biochemistry*, **27**, 6321–6326.
51 Xodo,L.E., Manzini,G., Quadrifoglio,F., van der Marel,G. and van Boom,J.H. (1989) *Biochimie*, **71**, 793–803.