Journal of Biomolecular NMR, **19:** 33–48, 2001.
KLUWER/ESCOM

Enhanced loop DNA folding induced by thymine-CH_3 group contact and perpendicular guanine-thymine interaction

Shan-Ho Chou[a,b,*], Ko-Hsin Chin[a] & Carton W. Chen[c]
[a]*Institute of Biochemistry and* [b]*Chemistry Department, Taipei, Taiwan National Chung-Hsing University, Taichung, 40227, Taiwan; and* [c]*Institute of Genetics, National Yang-Ming University, Taipei, Taiwan*

Received 14 August 2000; Accepted 13 October 2000

Key words: minor groove folding, perpendicular guanine-thymine interaction, T-methyl hydrophobic interaction, TTTG loop, UUCG loop

Abstract

A remarkable stabilizing effect induced by T-CH_3 group and perpendicular guanine–thymine interactions in the DNA loop conformation has been demonstrated for the d(TTTG) loop structure using UV melting, high resolution NMR, distance geometry, and molecular dynamics studies. Contrary to the previously published d(TTCG) sequence that exhibits no specific inter-nucleotide interaction, we have found that d(TTTG), which differs only by one nucleotide with the d(TTCG) sequence (C7 → T7), forms a rather stable and well-defined loop structure. Two characteristic structural features account for the stabilization of an otherwise flexible loop structure; the second loop T (T6) residue folds into the minor groove and engages in perpendicular interaction with the G8-NH_2, while the third loop T (T7) residue stacks well upon the closing T5•G8 wobble base pair and exhibits good contacts with many of the loop T5 and T6 sugar protons, which may form a hydrophobic core in the loop region. The importance of the bulky T7-CH_3 was also proved by the UV melting study; while d(TTCG) hairpin exhibits a lower melting point (74.5 °C) than d(TTTG) hairpin (80.5 °C), d($TT^{5-methyl}CG$) hairpin resumes the same higher melting point (80 °C). Similarly, the fact that the melting temperature (74 °C) of d(TTTI) is lower than that of d(TTTG) indicates the critical role played by the G8-NH_2 group. Our structural studies of the d(TTTG) loop indicate that DNA and RNA use a different strategy to establish stable tertiary folds. Comparison with several other pyrimidine-rich loop hairpins suggests that different minor-groove folding modes exist for the folding thymidine residue.

Introduction

GNRA, UNCG, and CUUG tetraloop hairpins are common RNA secondary structures (Woese et al., 1990), and have been suggested to play important roles in nucleating three-dimensional RNA folding (Tuerk et al., 1988; Uhlenbeck, 1990). The interaction of GNRA tetraloop with its receptor through platform formation (Cate et al., 1996; Butcher et al., 1997; Costa and Michel, 1997) is a good example. Such tetraloops have, therefore, been subjected to extensive experimental (Cheong et al., 1990; Heus and Pardi, 1991; Varani et al., 1991; Pley et al., 1994b; Allain and Varani, 1995; Jucker and Pardi, 1995; Sich et al., 1997) and theoretical studies (Zich, 1995; Akke et al., 1997; Miller and Kollman, 1997; Williams and Hall, 2000a, b). It is also generally believed that the ribose 2′-OH group is of critical importance in stabilizing RNA tertiary structure (Chastain and Tinoco, 1991; Strobel and Cech, 1993; Pley et al., 1994a), and assignment of a particular U2′-OH resonance improved the structural refinement of an UUCG tetraloop (Allain and Varani, 1995). Interestingly, the DNA analog of the UUCG tetraloop showed no specific base–base or base–phosphate interaction in the loop under similar conditions. The loop was therefore regarded as highly flexible and dynamic (James and Tinoco, 1993).

*To whom correspondence should be addressed. E-mail: shchou@dragon.nchu.edu.tw

We have been studying pyrimidine-rich triloops in the past few years (Chou et al., 1999b, 2000), and are working on the d(CTTG) loop hairpin that is found at the 3′-end secondary structure of a linear pSCL1 plasmid (Huang and Chen, 1997). We found that it can form a rather structured and compact motif (Chou et al., unpublished result). We were thus interested in knowing if the d(TTTG) sequence, which differs only by one nucleotide (C5 → T5) from the d(CTTG) sequence, and one nucleotide (C7 → T7) from the d(TTCG) sequence, can form a defined structure. Interestingly, the d(TTTG) sequence did turn out to be well behaved, forming a rather compact motif. Our structural studies indicate that the d(TTTG) loop is closed by a wobble T5•G8 base pair with G8 in the *anti* domain. The base and deoxyribose of the third loop T (T7) stack upon the T5 and G8 bases of the T5•G8 wobble base pair, respectively. Interestingly, T7-CH_3 was found to interact extensively with its surrounding sugar protons. Furthermore, T6 is folded into the minor groove to interact with G8-$^2NH_2$ in a perpendicular way. The importance of such interactions was further substantiated by the UV melting studies of several d(TTTG) loop hairpin analogs. Such results indicate that the T-CH_3 group can make a big difference in stabilizing an otherwise unstable DNA tetraloop structure. Our observation may have significant biological implications in that a methyl group is frequently added to cytosine in vivo and 5-methyl cytosine was found to greatly enhance the formation of left-handed Z-DNA (Fujii et al., 1982; Giessner-Prettre et al., 1984; Wang et al., 1985; Pearlman and Kollman, 1990).

Materials and methods

Sample preparation

All DNA samples were synthesized on a 6 μmol scale on an Applied Biosystems 380B DNA synthesizer with the final 5′-DMT groups attached. The samples were purified and prepared for NMR studies as described before (Chou and Tseng, 1999).

Thermodynamic analysis of the optical melting data

An absorbance (OD) versus temperature profile was obtained at 260 nm with a Spectronic Genesys2PC spectrometer equipped with a temperature controller. The temperature was increased from 25 °C to 95 °C at a rate of 0.5 °C/min in each run. Prior to experiments, samples were heated at 95 °C for several minutes and re-annealed quickly in ice water. Each run was repeated for three times under buffered conditions (pH 6.8, 10 mM sodium phosphate and 20 mM NaCl). The T_m values were determined from the maximum of first differential melting curves.

NMR experiments

All NMR experiments were obtained on a Varian Unity Inova 600 MHz spectrometer. One-dimensional imino proton spectra at 0 °C were acquired using the jump-return pulse sequence (Plateau and Gueron, 1982). The spectral width was 16 000 Hz with the carrier frequency set at the resonance of water. The maximum excitation was set at 10.5 ppm. For each experiment, 4K complex points were collected and 64 scans were averaged with a 2 s relaxation delay.

A 2D NOESY experiment in 90% H_2O/10% D_2O was performed at 0 °C in a low salt (20 mM) solvent with the following parameters: delay time 1 s, mixing time 0.12 s, spectral width 12136 Hz, complex points 2048, number of transients 96, and number of increments 500.

NOESY experiments in D_2O were also carried out at 0 °C in the hypercomplex mode with a spectral width of 4705 Hz. Spectra were collected using three mixing times of 60, 120, and 240 ms with a relaxation delay of 1 s between each transient and with 2048 complex points in the t_2 and 400 complex points in the t_1 dimension. For each t_1 incrementation, 64 scans were averaged.

A proton-detected ^{31}P-^{1}H heteronuclear correlation spectrum (Sklenar et al., 1986) was collected in the hypercomplex mode with spectral widths of 5000 Hz in the ^{1}H dimension and 2000 Hz in the ^{31}P dimension. Total numbers of 1024 complex points in the t_2 (^{1}H) dimension and 128 complex points in the t_1 (^{31}P) dimension were collected. Protons were presaturated for 1.0 s and 192 scans were accumulated for each t_1 incrementation.

The 2D ^{1}H-^{13}C HSQC (heteronuclear single-quantum coherence) spectrum was acquired with broad-band decoupling (Bax et al., 1983). The delay $1/(2^1J_{CH})$ was tuned to 3 ms for optimum excitation of sugar signals. Heteronuclear decoupling was achieved with the GARP-1 sequence (Shaka et al., 1985). In total, 160 t_1 increments of 2K complex data points were collected. Each FID in the t_1 dimension was further linear-predicted to 320 data points. The repetition delay was 1 s, and 16 scans were averaged for each FID. The carrier was positioned at 4.6 ppm for protons and 90.8 ppm for carbons. The spectral width

was 4669 Hz (7.8 ppm) in the proton dimension and 27183 Hz (180 ppm) in the carbon dimension. Total acquisition time was 1.5 h.

The high resolution ^{1}H-^{13}C HSQC spectrum was recorded with 800 t_1 increments of 2K complex data points and 32 scans, each with broad-band GARP-1 decoupling (Shaka et al., 1985). The carrier was positioned at 85.66 ppm for carbons. The spectral width was 4669 Hz (7.78 ppm) in the proton dimension and 2375 Hz (15.74 ppm) in the carbon dimension, leading to extensive folding. The total measuring time was 20 h.

The HMBC (heteronuclear multiple-bond coherence) experiment was recorded with 400 t_1 increments of 2K complex data points and 128 scans each. The delay $1/(2^1J_{CH})$ was tuned to 2.5 ms and the mixing time to 62 ms for observing smaller long-range heteronuclear couplings. Each FID in the t_1 dimension was further linear-predicted to 512 data points. The carrier was positioned at 4.6 ppm for protons and 90.8 ppm for carbons. The spectral width was 4669 Hz (7.8 ppm) in the proton dimension and 27183 Hz (180 ppm) in the carbon dimension. The total measuring time was 49 h.

The acquired data were transferred to an IRIS 4D workstation and processed by the software FELIX (MSI Inc.) (Chou and Tseng, 1999).

Structure determination

Three-dimensional structures of the d(TTTG) loop hairpin were generated by distance geometry and molecular dynamics calculations using distance and torsional angle constraints derived from NMR experiments. Most distance constraints from NOESY spectra in D_2O were classified as either strong, medium, or weak, based on their relative intensities at 120 ms mixing time and were given distance bounds of 2.0–4.0 Å, 3.0–5.0 Å, or 4.0–6.0 Å, respectively. Canonical hydrogen-bond distances of 1.8–2.1 Å were assigned to the Watson–Crick base pairs. A large number of distance constraints involving exchangeable protons were also derived from H_2O/NOESY spectra and were given only two wide distance bounds of either 2.0–5.0 Å or 3.0–6.0 Å, due to exchange phenomena. The β and γ torsional angle constraints were determined primarily semi-quantitatively from the ^{31}P-^{1}H heteronuclear correlation data (Chou and Tseng, 1999; Chou et al., 1999a, b, c, 2000) using the in-plane 'W' rule (Sarma et al., 1973). If the long-range (n)P $\leftrightarrow$ (n)H4′ four-bond couplings were detected, then the β and γ torsional angles were constrained to the *trans* (180° ± 30°) and *gauche*$^+$ (60° ± 30°) domains, respectively. Otherwise only the β torsional angle was constrained to the *trans* domain from the observation of the $^{3}J_{P-C4'}$ heteronuclear coupling of about 8 Hz (see Table 2). The ε torsional angle can only be located in either the *trans* or the *gauche*$^-$ domain (Altona, 1982). The *gauche*$^+$ conformation is not sterically allowed, which is also borne out from the observation that the three-bond $^{3}J_{P-H3'}$ heteronuclear coupling is not larger than 6.5 Hz (Table 2). Based on the small values of the long-range $^{4}J_{H2'-P}$ coupling (< 2 Hz), all ε torsion angles could be constrained to the *trans* domain (180° ± 30°) (Marino et al., 1999). The ζ and α dihedral angles were all left unconstrained. The χ dihedral angles were constrained to −100° ± 30° (ideal B-DNA values) when no aromatic–anomeric cross peaks of comparable intensity with the CH5/CH6 cross peaks were detected. These NOE distance (220 in total) and torsional angle (44 in total) constraints were used to generate initial structures using the DGII program (MSI, Inc.). The initial 20 DG structures (rmsd values of 0.71 ± 0.12 Å) were further refined by restrained molecular dynamics using the program DISCOVER (MSI, Inc.). A 2 ps dynamics run was performed at 300 K with a step size of 1.0 fs, which was followed by a conjugate gradient minimization of 200 iterations looped 10 times. Well-converged final structures with pair-wise rmsd values of 0.65 ± 0.08 Å were obtained after molecular dynamics calculations.

Results

Thermodynamics studies

In order to understand why the d(TTTG) loop hairpin is stable, we have synthesized 5′-(CCGC<u>TTTG</u>GCGG)-3′ [d(TTTG)] (the loop residues are underlined), 5′-(CCGC<u>TT***C**G*</u>GCGG)-3′ [d(TT***C***G)], 5′-(CCGC<u>TT $^{5-Me}$***C**G*</u>GCGG)-3′ [d(TT^{5-Me}***C***G)] ($^{5-Me}$C stands for 5-methyl-cytosine), and 5′-(CCGC<u>TTT***I***</u>GCGG)-3′ [d(TTT***I***)] analogs and measured their UV melting profiles from 25 to 95 °C. The melting studies were performed under buffered conditions of 10 mM pH 6.8 sodium phosphate, 20 mM NaCl, and 0.1 mM EDTA. Figure 1 shows the melting curves (optical density versus temperature) of d(TTTG), d(TTCG), and d(TT^{5-Me}***C***G) (a), as well as d(TTT***G***) and d(TTT***I***) (b) oligomers. The vertical lines mark the melting temperatures determined from the first differential curves. From the figures it is clear that while

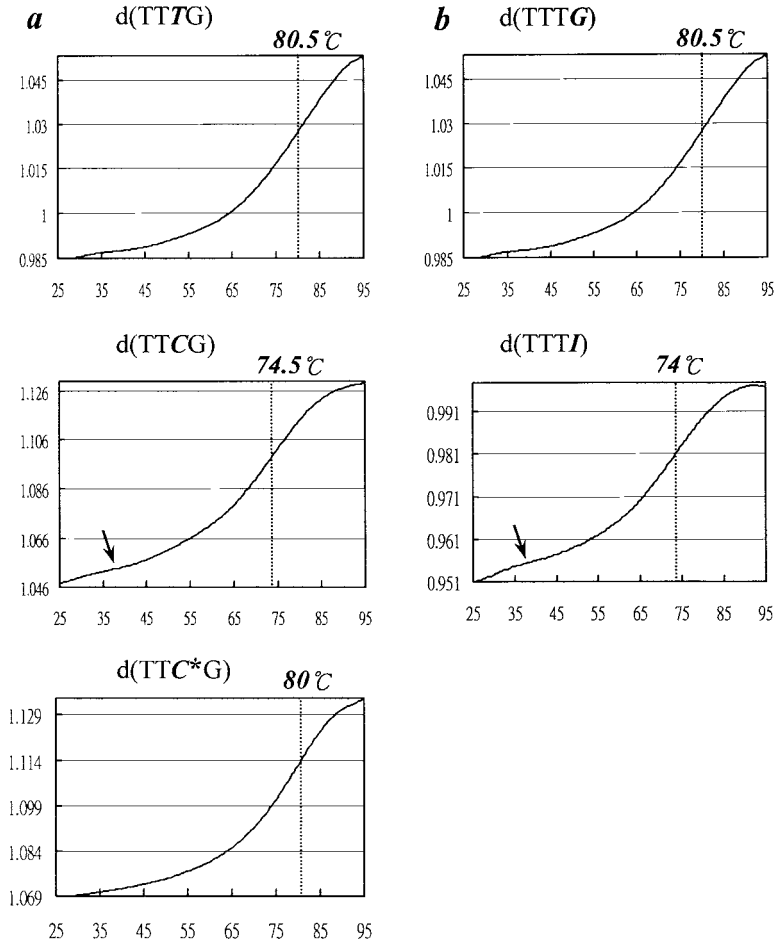

Figure 1. The absorbance versus temperature melting curves of d(TT***T***G), d(TT***C***G) and d(TT^{5-Me}***C***G) (a), as well as d(TTT***G***) and d(TTT***I***) (b) oligomers under buffered conditions. T_m values were determined from the maximum of the first differential melting curves and are marked with dotted vertical lines. Tm values of about 80 °C were observed for both d(TTTG) and d(TT^{5-Me}CG) oligomers, while Tm values of approximately 6 °C less (74.5 °C) were found for both d(TTCG) and d(TTTI) oligomers. The importance of thymine methyl group and perpendicular base–base interaction in stabilizing the d(TTTG) loop structure is evident from such UV melting studies. In both d(TTCG) and (TTTI) melting curves, a first transition temperature of about 37 °C was also detected (marked with an arrow), possibly due to the formation of a duplex form with a central track of four mismatches (Holbrook et al., 1991; Cruse et al., 1994).

d(TTCG) hairpin has a lower melting point (74.5 °C, middle of Figure 1a) than d(TTTG) (80.5 °C, top of Figure 1a), d(TT^{5-Me}CG) resumes the same higher melting point of d(TTTG) (80 °C, bottom of Figure 1a). We also measured the melting curve of the TTdUG loop hairpin and found that it exhibits a low melting temperature at approximately 60 °C (data not shown). Such UV melting studies indicate that the second loop T-CH_3 group plays an important role in stabilizing the d(TTTG) hairpin. The importance of the G8-NH_2 group is also revealed by the comparative melting studies of the d(TTT***G***) and d(TTT***I***) analog hairpins; a lower melting temperature was observed for the d(TTT***I***) hairpin (74 °C, bottom of Figure 1b) than for d(TTT***G***). The lack of a G-amino group thus accounts for a stability decrease of approximately 6 °C in the melting temperature.

NMR studies

Since the thermodynamics studies described above have indicated that both unpaired loop thymidines are important in stabilizing the d(TTTG) loop structure, we have further used NMR, distance geometry and molecular dynamics methods to determine its solution structure and tried to explain the critical roles played by the unpaired loop thymidines. The one-dimensional imino and amino/aromatic proton spectrum of the d(TTTG) sequence under low salt conditions (20 mM

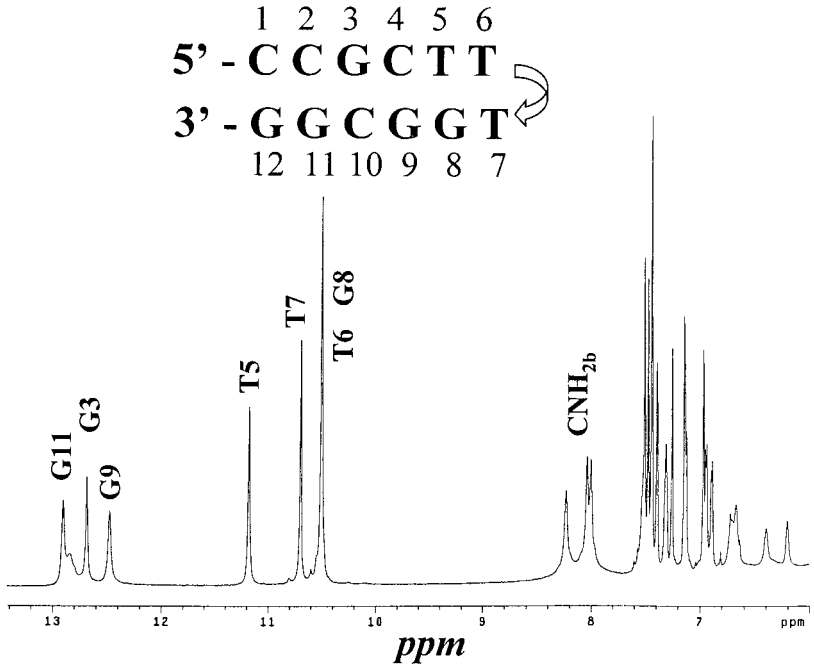

Figure 2. One-dimensional 600 MHz imino, amino and aromatic proton NMR spectrum of the d(TTTG) oligomer under low salt (20 mM NaCl, 0.1 mM EDTA) and low temperature (0 °C) conditions. Three sharp signals accounting for four imino protons between 10.4 and 11.4 ppm were clearly observed until over 40 °C.

Table 1. Proton and phosphorus chemical shifts of the 5′-d(CCGC-TTTG-GCGG)-3′ hairpin

	H1/H3	NH_2	H1′	H2′/H2″	H3′	H4′	H5′$^{(Pro-S)}$	H5′$^{(Pro-R)}$	H6/H8	H5/H2/CH_3	P
1C		6.69/7.45	5.49	1.59/2.03	4.18	3.63	3.23	3.23	7.28	5.44	
2C		6.64/8.15	5.08	1.72/1.97	4.39	3.74	3.92	3.97	7.10	5.19	−3.90
3G	12.61		5.49	2.19/2.26	4.52	3.91	3.96	4.06	7.48		−3.69
4C		6.37/7.96	5.61	1.44/1.91	4.33	3.69	4.16	4.04	6.86	4.96	−3.95
5T	11.13		5.47	1.87/1.91	4.39	3.74	3.98	3.93	7.22	1.36	−4.09
6T	10.43		5.54	1.51/1.86	4.20	3.75	3.99	3.90	7.11	1.32	−4.16
7T	10.62		5.21	1.38/1.51	4.07	2.84	3.40	3.66	6.95	1.17	−4.74
8G	10.43	5.88	5.00	2.20/2.24	4.29	3.77	3.82	3.76	7.40		−4.33
9G	12.39		5.34	2.21/2.18	4.49	3.89	3.60	3.60	7.44	4.95	−3.83
10C		6.18/ 7.91	5.24	1.46/1.85	4.38	3.73	4.08	4.02	6.91		−3.94
11G	12.88		5.15	2.20/2.26	4.50	3.87	3.92	4.00	7.42		−3.76
12G	12.82		5.68	2.19/1.84	4.19	3.75	4.13	4.03	7.37		−3.70

[a]Both chemical shifts of exchangeable and non-exchangeable protons were determined at 0 °C.
[b]The Pro-S and pro-R H5′ protons were assigned from the combination of high-resolution ^{1}H-^{13}C HSQC and HMBC experiments as described previously (Chou et al., 2000).
[c]Phosphorus chemical shifts were determined from the ^{1}H-^{31}P correlation experiment.

NaCl) at 0 °C is shown in Figure 2. Three sharp imino proton peaks with an intensity ratio of 1/1/2 were clearly observed between 10.4 and 11.2 ppm. We also detected three sharp and one broad peaks, further downfield between 12.5 and 13.0 ppm. Such peaks are characteristic of the four Watson–Crick base-paired G imino protons, and were assigned through the two-dimensional H_2O/NOESY spectrum shown in Figure 3 (Tseng and Chou, 1999). Temperature dependent imino proton studies indicate that their linewidths became broader when the temperature was increased, but were observable until the temperature is higher than

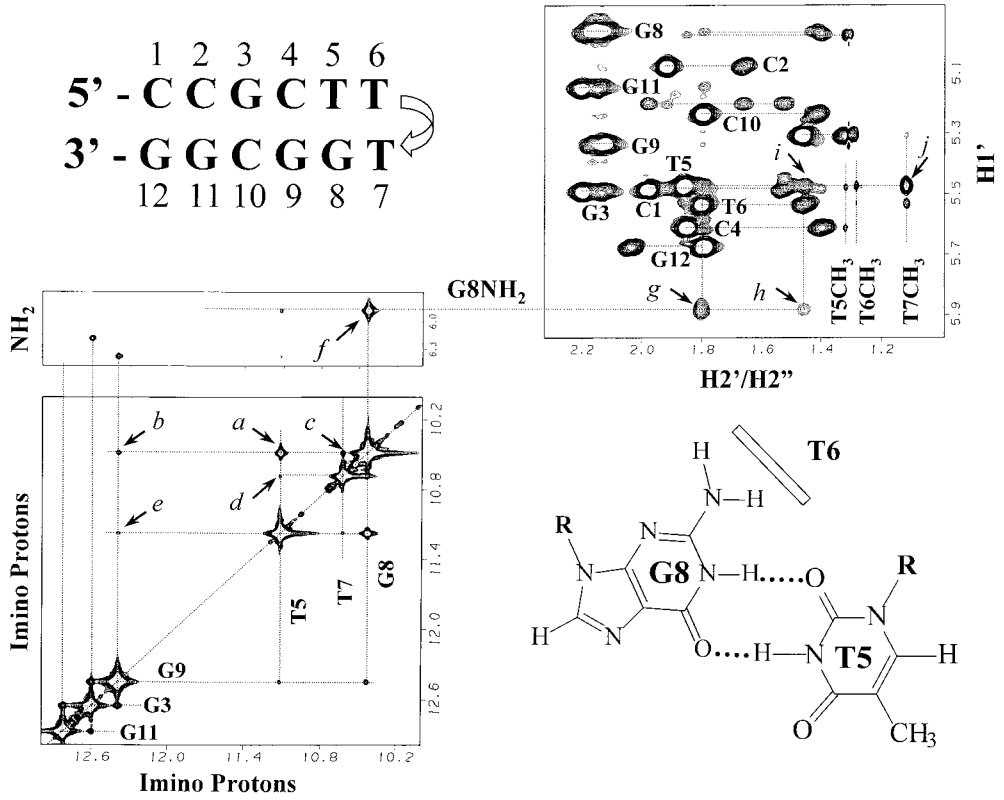

Figure 3. Portions of expanded H_2O/NOESY contour plots (mixing time 0.12 s) of the d(TTTG) oligomer at 0 °C and the wobble G•T base pair scheme in the present TTTG loop. These plots cover NOE cross peaks from imino to imino (lower-left), imino to amino (middle-left), and H1′ to H2′/H2″ (upper-right) regions. Some unusual NOEs that are critical in the structural determination of d(TTTG) loop are marked with small letters and assigned as: (a) T5-H3–G8-H1; (b) G9-H1–G8-H1; (c) T7-H3–G8-H1; (d) T5-H3–T7-H3; (e) G9-H1–T5-H3; (f) G8-H1–G8-NH_2; (g) G8-NH_2–T6-H2′; (h) G8-NH_2–T6-H2″; (i) T5-H1′–T6-H2″; (j) T7-CH_3–T5-H1′.

40 °C. This result indicates that the d(TTTG) hairpin is stable at 0 °C (on the slow exchange NMR timescale), and forms a stable structure at this temperature without extraordinary dynamics. If the narrow imino proton linewidth is interpreted as being due to fast dynamics, then the linewidth will become even smaller as temperature is increased, which was not the case. The narrow imino proton linewidth therefore indicates that the d(TTTG) hairpin is in the slow exchange domain.

Since the r(UUCG) sequence has the potential to form a duplex structure with a track of non-Watson–Crick base pairs in the solid state (Holbrook et al., 1991; Cruse et al., 1994), the current d(TTTG) sequence was therefore carefully checked to see if it forms a hairpin or a duplex. After diluting the sample several times and reannealing it before data collection, no new imino proton peak was found. The concentration independence of the NMR spectrum and the UV melting studies indicate that the d(TTTG) sequence most likely forms a hairpin under the millimolar conditions used for the current NMR studies.

Figure 3 shows parts of the two-dimensional H_2O/NOESY of the d(TTTG) hairpin. Except for the terminal G12-H1, all other imino protons can be sequentially assigned from G11-H1, G3-H1 to G9-H1 through the NOE cross peaks (lower left). The connectivity could be further extended to T5-H3 (peak e) and G8-H1 (peak b). Such sequential connection, along with a strong NOE cross peak between T5-H3 and G8-H1 (cross peak a), indicates that T5 and G8 form a wobble type base pair (Kennard, 1985; Hare et al., 1986) that stacks upon the stem C4-G9 base pair. The two weak NOEs of the T5-H3–T7-H3 (cross peak d) and G8-H1–T7-H3 (cross peak c) proton pairs indicate that T7 stacks upon the T5•G8 wobble base pair (T7-H3 was assigned by a weak NOE from T7-H3 to T7-CH_3). Such unusual stacking of T7 upon T5 is validated by the abundant NOEs between T7-CH_3 and T5 sugar protons. One of these NOEs (T7-CH3–T5-H1′)

Table 2. Carbon-phosphorus and proton-phosphorus scalar couplings (Hz) of the 5′-d(CCGC-TTTG-GCGG)-3′ hairpin

	$^3J_{C2'-P}$	$^2J_{C3'-P}$	$^3J_{H3'-P}$	$^3J_{C4'-P}$[e]	$^4J_{H4'-P}$[f]
1C	< 2	−5.9	6.2		0[b]
				9.3[c]	
				8.1	
2C	< 2	−4.7	4.1		4.2
				9.3	
				10.4	
3G	< 2	−5.8	4.1		4.1
				10.4	
				9.3	
4C	< 2	−5.9	6.2		4.2
				9.3	
				8.3	
5T	< 2	−4.7	4.1		4.1
				9.2	
				8.4	
6T	< 2	− 4.7	6.2		4.1
				8.1	
				8.1	
7T	∼ 4	− 4.6	6.2		4.1
				8.2	
				8.1	
8G	< 2	−4.5	4.1		∼ 0
				8.1	
				9.2	
9G	< 2	−4.6	4.1		4.2
				8.1	
				9.2	
10C	< 2	−4.7	5.2		4.1
				9.3	
				9.3	
11G	< 2	−6.2	6.2		4.1
				9.2	
12G	< 2	na[a]	na[a]	8.1[d]	4.2

[a]na: non-available, no phosphate at 3′-end.
[b]Four-bond (n)H4′ − (n+1)P coupling vanishes when H4′-C4′-C3′-O3′-P atoms are not in the 'W' shape (Sarma et al., 1973; Chou et al., 1997).
[c]Three-bond (n)C4′ − (n+1)P coupling constant.
[d]Three-bond (n)C4′ − (n)P coupling constant.
[e]Each central residue has two $^3J_{C4'-P}$ values that are approximately equal, but could not be specifically identified.
[f]Sum of the (n)H4′ − (n)P and (n)H4′ − (n+1)P couplings, but the latter value is usually zero.

Table 3. Constraints used to determine the structure of the 5′-d(CCGC-TTTG-GCGG)-3′ hairpin

Restraints	Values
Exchangeable NOEs	54
H-bonds (1.8–2.1 Å)	14
2.0–4.0 Å	4
3.0–6.0 Å	25
5–10.0 Å	11
Non-exchangeable NOEs	166
2.0–4.0 Å	24
3.0–5.0 Å	82
4.0–6.0 Å	55
5–10.0 Å	5
Total NOEs	220
Torsional angles	44
Backbone (β, γ, ε)	32
Glycosidic	12
NOEs per residue	18.3
NOEs and torsion angles per residue	22.0
Violations of experimental restraints[a]	
Distance restraints (> 0.15 Å)	0
Torsional angles restraints (> 10°)	0
rmsd	0.65 ± 0.10 Å

[a]Two out of the 20 final structures have one violation larger than 0.15 Å.

is indicated by cross peak j in the upper right spectrum. Other unusual NOEs from T6-H2″ to T5-H1′ (cross peak i) and T6-H2′/H2″ to G8-NH_2 (cross peaks g and h) indicate that the T6 sugar is also close to the T5•G8 base pair. This is revealed by the final NMR structure that shows the T6 nucleotide folding into the minor groove and the T6 base is in a position that is almost perpendicular to the G8 base. This can explain the sharp nature of one of the G8-NH_2 signals (the other amino proton is probably too close to the H_2O signals to be observed) and the well-defined NOEs between G8-NH_2-G8-H1 (cross peak f in the middle left spectrum) and G8-NH_2-T6-H2′/H2″ proton pairs (cross peaks g and h).

Further information necessary for determining the d(TTTG) loop structure was acquired from the D_2O/NOESY that was assigned using the well-established NOE cross-peak connectivity method (Hare et al., 1983). All aromatic, H1′, H2′/H2″, H3′, and H4′ protons were sequentially assigned through the tetraloop region without any problem (data not shown). Unlike the G(*syn*) conformation adopted for the reverse wobble U•G base pair (Varani et al., 1991), the G8 nucleotide in the d(TTTG) sequence is in

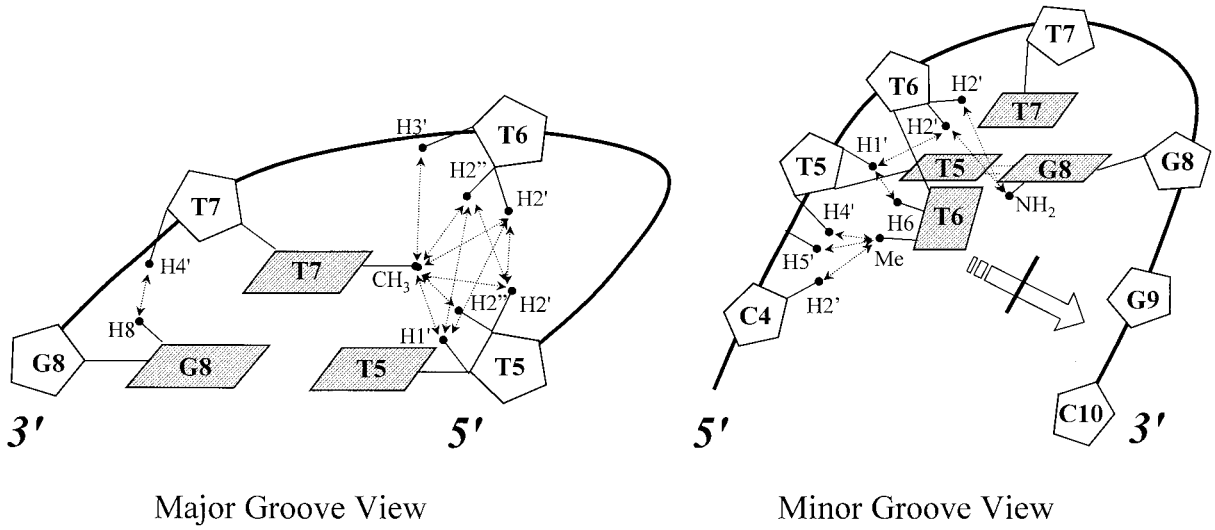

Figure 4. The characteristic NOEs of the d(TTTG) loop from the major groove (left) and minor groove view (right). The abundant T6-CH_3 and T7-CH_3 related NOEs are clear from this figure. Intriguingly, no NOE was observed for the folded-in T6 to the stem G9 and C10 residues (indicated by a blocked hollow arrow). This NOE was clearly observed for the d(TCC) (Chou et al., 1999b) and d(TTT) (Chou et al., 2000) loop hairpins with a similarly folded-in loop thymidine. This observation indicates that there exists an alternative conformation for the folded-in thymidine.

the regular *anti* domain, as proved by the weak-to-medium G8-H8 to G8-H1′ NOE intensity (data not shown). Many characteristic NOEs between the T5, T6, and T7 residues that were useful for determining the tertiary fold of this DNA hairpin were identified. Some of these important NOEs are listed in the cartoon shown in Figure 4. The chemical shifts for all assigned protons are listed in Table 1.

Information for constraining the backbone torsion angles can be derived from the natural abundance heteronuclear ^{1}H-^{31}P and ^{1}H-^{13}C spectra as described before (Schmieder et al., 1992; Avizonis and Kearn, 1995; Ippel et al., 1995; Chou et al., 2000). All the measured heteronuclear coupling values are listed in Table 2.

Structural features

Due to the sharp nature of the exchangeable imino-, amino, and non-exchangeable proton signals, we were able to detect many NOEs that are critical in determining this d(TTTG) loop structure. These NOEs around the loop region are summarized in Figure 4 from two different views, with the derived distance constraints listed in Table 3. Figure 5 shows the superimposition of the final 20 (out of 30) structures before (Figure 5a) and after (Figure 5b) molecular dynamics calculation. It is clear from the figure that the DG structures have already converged quite well (rmsd value: 0.71 ± 0.12 Å). After restrained molecular dynamics calculation, the rmsd value further decreased to 0.65 ± 0.10 Å. All glycosidic angles are in the *anti* range, and all sugar puckers are in the C2′-endo domain. The first and last residues form a wobble G8•T5 base pair, and the stacking continues from the stem to the loop G8•T5 base pair. The first loop T (T6) residue folds into the minor groove and is almost perpendicular to the G8 base, while the second loop T (T7) residue stacks upon the closing T5•G8 wobble base pair.

Figure 6 shows the important structural features that may be responsible for the stability of the d(TTTG) loop. Figure 6a shows the excellent stacking of the T7 base upon the T5 base and the T7-deoxyribose upon the G8 base. Such deoxyribose-base stacking has been observed many times in the purine-rich single-residue loop hairpins (Chou et al., 1994, 1996a, b, 1999c; Zhu et al., 1995), and can explain the upfield shifting of the T7-H4′ and the T7-H5′/H5″ protons in the present case. The extensive hydrophobic interaction induced by T7-CH_3 in the loop region is shown in Figure 6b. While the T5-CH_3 group is turned away from the loop backbone, the T7-CH_3 group instead points toward it and forms good contacts with many of the T5 and T6 sugar protons. Such contacts are supported by the many medium strength NOEs, as illustrated in Figure 4 and by the melting temperature profile shown in Figure 1a.

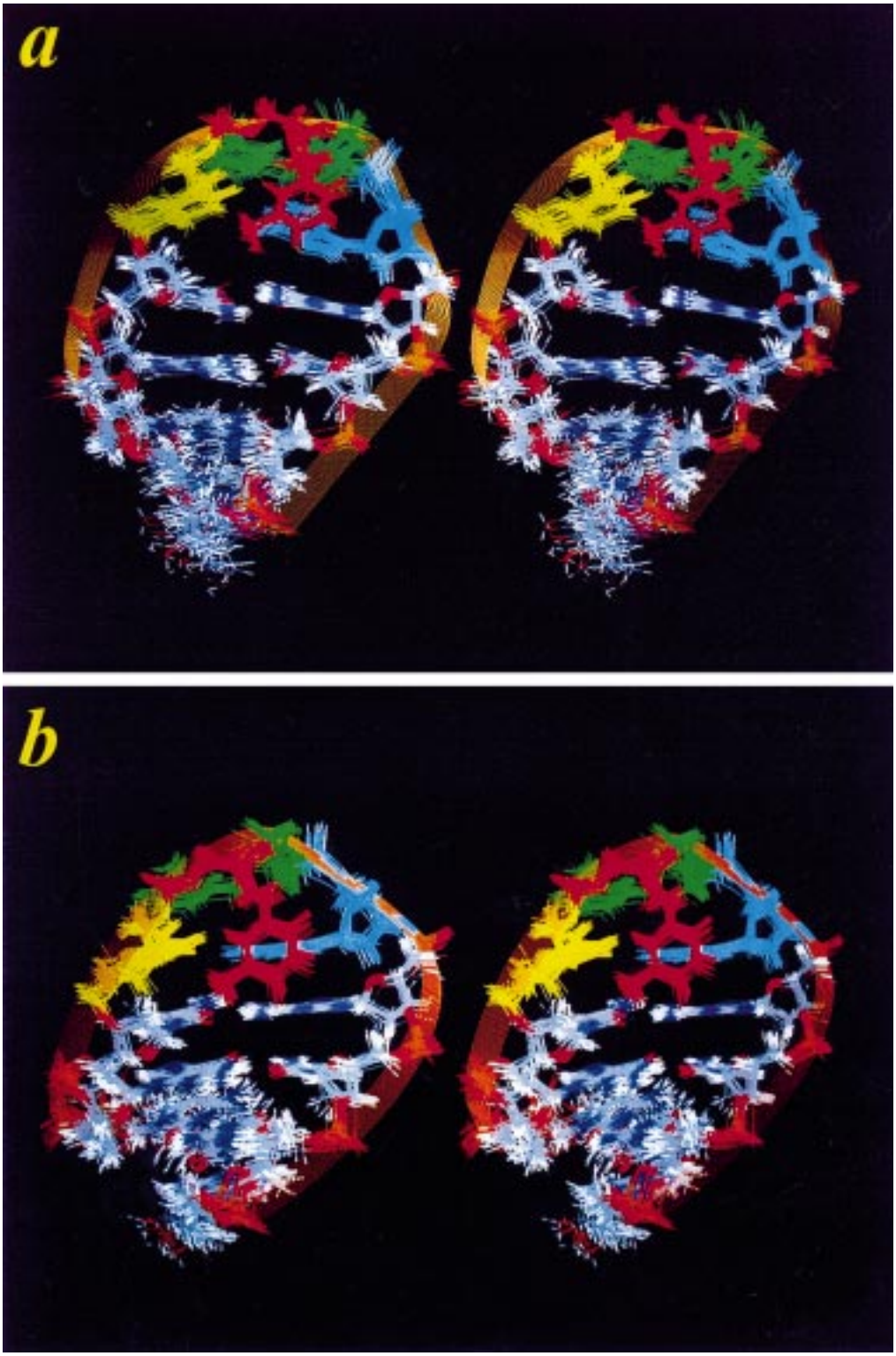

Figure 5. Superimposed wide-eye stereoviews of the final 20 structures of the d(TTTG) loop hairpin before (a) and after (b) molecular dynamics calculation. Before the molecular dynamics calculation, the 20 DG structures overlap with one of the selected structures with rmsd values of 0.71 ± 0.12 Å (a). After further molecular dynamics calculation, the 20 final structures overlap with one of the selected structures with rmsd values of 0.65 ± 0.10 Å (b). The T5 residue is colored in yellow, T6 in red, T7 in green, and G8 in blue. Orange ribbons were drawn to connect the hairpin backbones. Structures usually improve in helicity and hydrophobic stacking after a 2 ps molecular dynamics calculation.

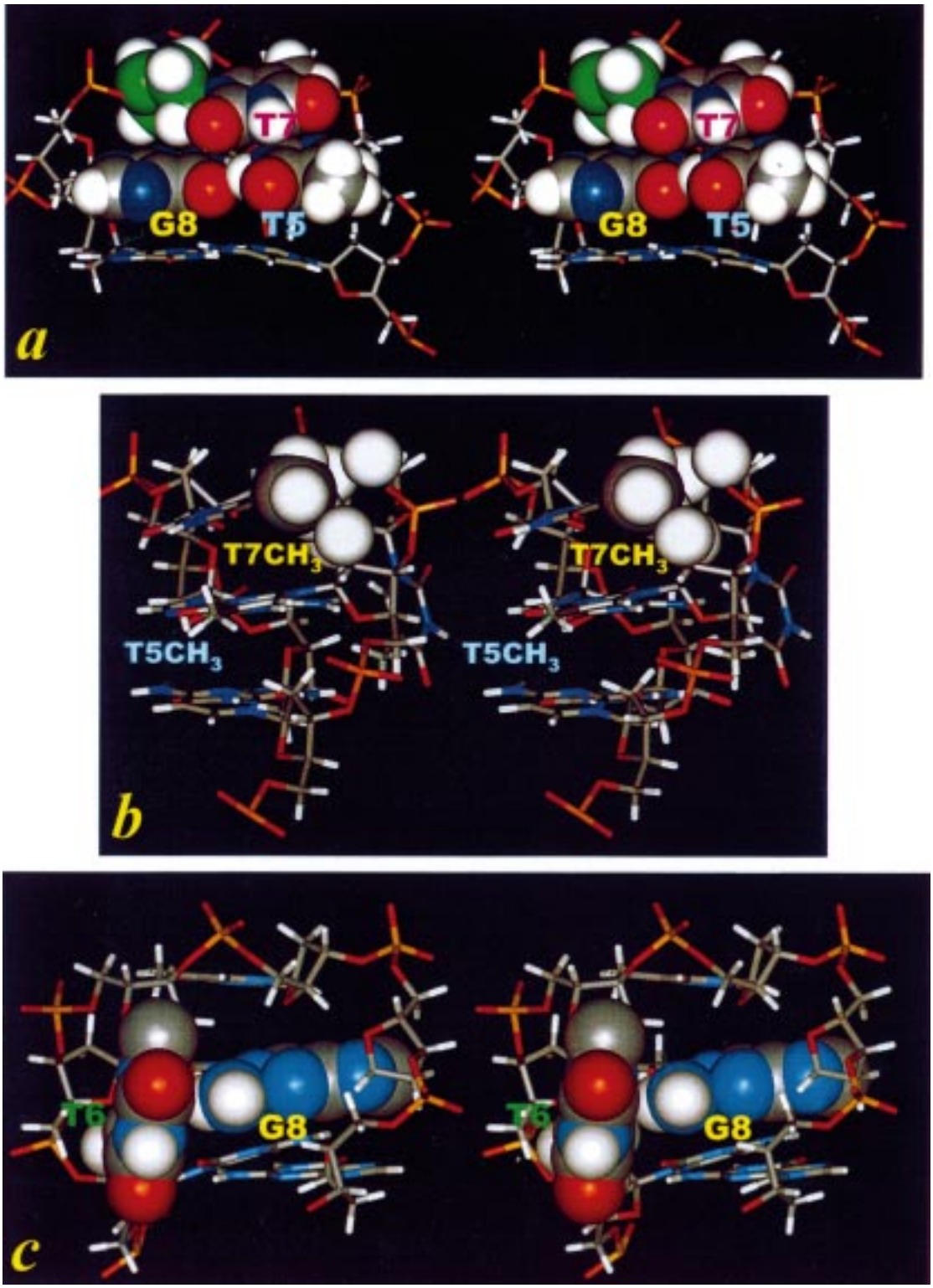

Figure 6. Wide-eye stereoviews of the dC(TTTG)G loop structure showing the structural features that are important in stabilizing this loop conformation. The carbons are colored in gray, oxygens in red, phosphorus in orange, nitrogen in deep blue, and protons in white, except that carbons in the T7-deoxyribose are colored in green. The T7 residue stacks well with the T5•G8 wobble base pair, with its base stacking upon the base of T5, and its sugar upon the base of G8. All these structural elements are drawn in space-filling in (a). In (b), only the T7-CH_3 group and sugar protons in close contact with the T7-CH_3 group are drawn in space-filling. Extensive hydrophobic interaction induced by the T7-CH_3 group is evident from this figure. In (c), only the bases of T6 and G8 are drawn in space-filling to reveal the perpendicular interaction between the unpaired T6 base and G8-NH_2.

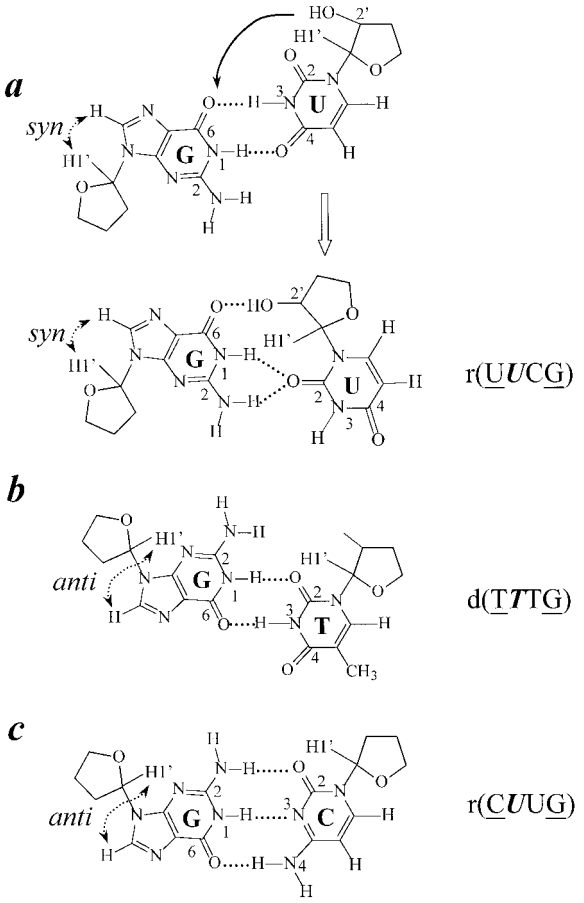

Figure 7. Pairing schemes of the closing U•G, T•G, and C•G base pairs in the r(UUCG), d(TTTG), and r(CUUG) tetraloops: (a) a reverse wobble U•G base pair (top of a) published by Cheong et al. (1990) and Varani et al. (1991) is characterized by a G(*syn*) conformation and two H-bonds between G-H1–U^4O and G^6O–U-H3. The newly defined U•G base pair obtained by Allain and Varani (1995) has a different H-bonding pattern that is characterized by two bifurcated $GH1/G^2NH_2$–U^2O H-bonds and a base-sugar G^6O–U2′-OH H-bond (bottom of figure a); (b) the wobble T•G base pair in the present d(TTTG) loop structure is characterized by a G(*anti*) conformation and two H-bonds between GH1–U^2O and G^6O–UH3; (c) the canonical Watson–Crick G•C H-bonds in the r(CUUG) closing base pair.

Figure 6c gives a stereoview of another important structural feature of this d(TTTG) loop; a perpendicular interaction between the folded T6 base and the G8-NH_2 (drawn in space-filling). The G8-NH_2 is found pointing toward the center of the T6 base. This feature is correlated with the lower melting temperature of the d(TTTI) hairpin compared with that of the d(TTTG) hairpin (see Figure 1b). Such perpendicular N-H/π interaction can also explain the characteristic NOEs exhibited by the G8-amino proton (G-amino protons are usually highly labile and exhibit no NOE), and the upfield shifting of the G8-amino proton, which now resonates at 5.88 ppm, at least 1–2 ppm upfield from the regular values (between 7 and 8 ppm). While the importance of such a perpendicular interaction remains to be established, many instances of perpendicular C-H/π interactions (Chou et al., 1996a, b, 1997; Chou and Tseng, 1999), O4′/π interactions (Wang et al., 1981; Chou et al., 1994; Egli and Gessner, 1995; Nishinaka et al., 1997), and N-H/π interactions (Perutz et al., 1986; Levitt and Perutz, 1988) have already been described and were found to play a non-trivial role in stabilizing nucleic acids, nucleic acid-ligand complexes (Kopka et al., 1985; Coll et al., 1987; Teng et al., 1988), protein recognition (Nishio et al., 1995), and peptide β-turn formation (Jimenez et al., 2000).

The structural features described above may explain why the presence of thymine residues in the loop region can make such a big difference in the UV melting of the d(TTTG) analogs. Although similar CH_3 group induced stabilization has been observed for cytosine 5-CH_3 in the Z-DNA structure (Fujii et al., 1982), this is the first time to our knowledge that a single nucleotide change can cause such a dramatic change in the DNA loop structure. This indicates that, while RNA uses 2′-OH groups to form extra H-bonds or coordinating bonds to stabilize its tertiary fold, DNA can instead use thymine-CH_3 groups to form a hydrophobic core to stabilize its folding. Such hydrophobic interactions can be quite substantial, as clearly demonstrated in this paper.

Since the present d(TTTG) loop contains two unpaired thymidines, its dynamics behavior is certainly a concern that needs to be addressed, especially when the analogous d(TTCG) loop was found to be rather dynamic and unstructured (James and Tinoco, 1993). This situation also occurred in another r(CUUG) loop structure (Jucker and Pardi, 1995); considerable dynamics was detected in the loop region, as revealed by the absence of many expected NOE cross peaks and by the increased line widths for several of the loop resonances. However, similar to the cases of the d(TCC) (Chou et al., 1999b) and d(CTTTG) (Chou et al., 2000) loops, we have not found any significant dynamics on the NMR time scale for the present d(TTTG) loop; all resonance lines are sharp at low temperature but become broadened at higher temperature, and all expected and quite a few unexpected NOE cross peaks were clearly identified (see Figures 3 and 4). The d(TTTG) loop hairpin can therefore be considered

Table 4. Backbone torsion angles (deg) of the 5′-d(CCGC-TTTG-GCGG)-3′ hairpin

Residue	α	β	γ	δ	ε	ζ	χ
1C				112.3±12.6	−173.2±4.7	−86.4±2.6	−135.0±5.1
2C	−70.6±6.1	176.8±3.9	58.1±1.0	121.2±3.8	−172.8±1.9	−90.8±2.3	−129.4±6.3
3G	−81.3±5.0	177.0±4.8	55.1±3.6	121.8±3.4	−176.5±1.9	−93.4±8.0	−124.6±7.0
4C	−78.3±3.9	174.4±2.3	55.3±2.8	***83.2±10.2***	−164.9±3.3	−68.4±2.5	−145.1±8.1
5T	−72.5±5.3	179.4±2.4	69.7±1.6	143.4±5.2	165.9±2.5	−112.5±7.6	***−94.6±5.5***
6T	***−168.3±3.5***	149.8±2.9	89.9±2.9	128.0±4.5	−145.8±3.2	−53.3±3.8	−127.9±4.7
7T	−87.5±2.3	−162.7±5.7	62.3±1.0	138.3±3.7	***−110.2±2.5***	***62.7±2.2***	−133.1±2.2
8G	***109.2±7.2***	−176.3±1.8	***−174.7±1.2***	159.8±1.1	177.2±3.5	−102.6±7.6	***−74.0±2.6***
9G	−73.5±2.7	168.4±7.3	62.9±4.6	125.8±9.2	177.7±2.5	−100.7±7.5	−104.2±8.8
10C	−67.5±2.6	168.5±0.8	65.8±1.7	121.3±9.8	−177.7±2.9	−95.2±5.1	−118.8±8.0
11G	−74.2±2.9	176.3±1.5	58.1±2.9	124.5±6.3	−179±1.0	−96.5±6.2	−125.2±5.1
12G	−72.6±6.3	−179.6±3.4	58.3±2.8	116.9±18.3			−129±3.9

Parameters were calculated based on the values of 20 final structures. Values are expressed as mean values ± standard deviations. Special torsional angles are listed in ***italic-bold***.
The narrow ranges for some of the dihedral angles are possibly due to the applied AMBER force field, and do not necessarily mean that NMR methodology can determine biopolymer structure with such precision.

as relatively 'rigid', exhibiting no significantly differential dynamics between the loop and stem residues on the NMR timescale. The combined hydrophobic and perpendicular N-H/π interaction induced by the two unpaired thymidines may have made a substantial contribution to stabilize the d(TTTG) loop.

Discussion

Comparison of the structural features among r(UUCG), d(TTTG), and r(CUUG) tetraloops

Until now, structure determination of several DNA or RNA pyrimidine-rich triloops (Mooren et al., 1994; Chou et al., 2000) and tetraloops (Blommers et al., 1991; Hilbers et al., 1994; Allain and Varani, 1995; Jucker and Pardi, 1995; van Dongen et al., 1996) has been achieved by high resolution NMR techniques. It is therefore interesting to compare several of these structures with a common guanosine in the last loop position to determine the structural characteristics that contribute to the stability of such tetraloops. The base-pairing schemes of the three selected hairpins of r(UUCG), d(TTTG), and r(CUUG) tetraloops are drawn in Figure 7 with their characteristic features listed in Table 5 for comparison.

The r(UUCG) tetraloop (Allain and Varani, 1995) was refined from the previous reported structure (Cheong et al., 1990; Varani et al., 1991) and features a closing base-pair of unusual H-bonding. As shown at the bottom of Figure 7a, $G^6_{L4}O$ was found to pair with $U_{L1}2'$-OH (Ln designates the nth loop residue in the tetraloop) after successful assignment of the $U_{L1}2'$-OH and identification of several important NOEs from the G_{L4}-H1 proton to the U_{L1}-H1′, U_{L1}-H2′, and $U_{L1}2'$-OH protons (Allain and Varani, 1995). Such a special base-sugar H-bond causes a dramatic counterclockwise rotation of the U_{L1} base plane (larger than 60°, see top of Figure 7a) and formation of two novel bifurcated $G_{L4}H1/G_{L4}NH_2$–$U^2_{L1}O$ H-bonds. Such rotation also decreases the $G_{L4}C1'$-$U_{L1}C1'$ distance and relieves the buckle tension present in the originally proposed G•U base pair. The newly defined loop structure was indeed found to be more stable from an unconstrained molecular dynamics study (Miller and Kollman, 1997). In the d(TTTG) loop (middle of Figure 7) the G_{L4} base instead pairs with the T_{L1} base through the wobble $G_{L4}H1$–$T^2_{L1}O$ and $G^6_{L4}O$–$T_{L1}H3$ bonds normally found in DNA (Kennard, 1985; Hare et al., 1986). In the r(CUUG) loop, a canonical Watson–Crick $C_{L1} \bullet G_{L4}$ base pair is, however, found to close the tetraloop structure (bottom of Figure 7) with conversion of all four loop sugars from C3′-endo to C2′-endo (see below). The C1′-C1′ distances of the closing base pairs for the r(UUCG), d(TTTG), and r(CUUG) loops are nevertheless approximately equal and have values of 10.21 Å, 10.73 Å, and 10.21 Å, respectively.

Another major difference between these structures is the different sugar puckers incorporated to form such tetraloops. Since tetraloop hairpins have a loop of only two nucleotides, it is generally believed that

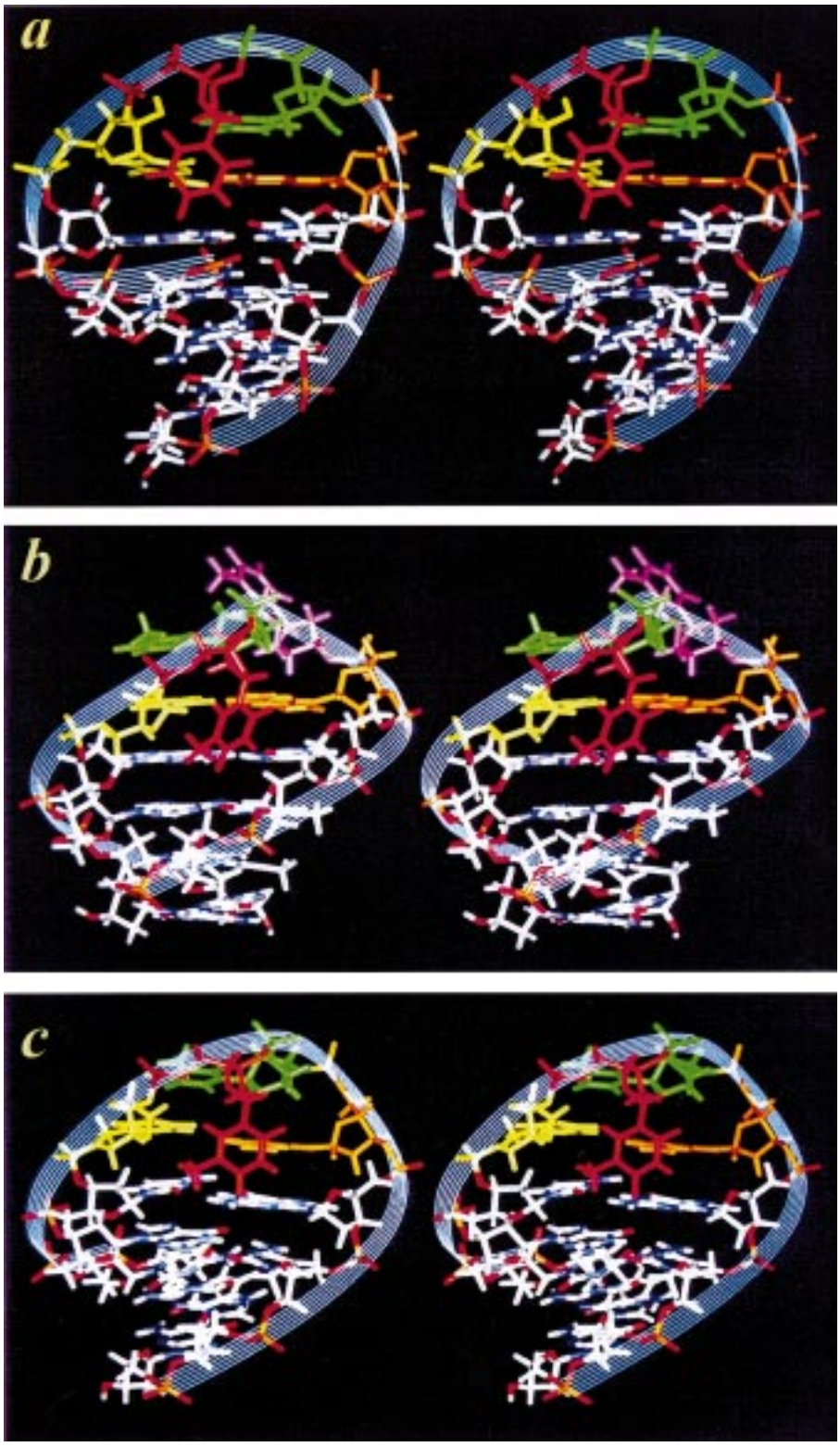

Figure 8. Different minor-groove folding modes for the r(CUUG) (a), d(CTTTG) (b), and d(TTTG) (c) hairpins. The hairpin backbones are connected by a blue ribbon, with the first Py_{L1} residue colored in yellow, the folding Py_{L2} in red, the stacking Py_{L3} in green, and the last G_{L4} in orange. The extra looping-out $T_{L2'}$ in d(CTTTG) is colored in purple. The different minor groove width is evident from this figure.

Table 5. Comparison of the structural features of the r(U***U***CG), d(T***T***TG), and r(C***U***UG) tetraloops[a]

	$\underline{U_{L1}}\boldsymbol{U}_{L2}C_{L3}\underline{G_{L4}}$[b]	$\underline{T_{L1}}\boldsymbol{T}_{L2}T_{L3}\underline{G_{L4}}$	$\underline{C_{L1}}\boldsymbol{U}_{L2}U_{L3}\underline{G_{L4}}$
H-bonding of the closing base pair	Bifurcated $G_{L4}H1/G^2_{L4}NH_2$–$U^2_{L1}O$ and base – deoxyribose $G^6_{L4}O$–$U_{L1}2'$-OH	Wobble $G_{L4}H1$ – $T^2_{L1}O$ and $G^6_{L4}O$ – $T_{L1}H3$	Canonical Watson–Crick C_{L1}•G_{L4} H-bonding
Sugar pucker	U_{L1} and G_{L4} C3′-endo U_{L2} and C_{L3} C2′-endo	All C2′-endo	All C2′-endo
Glycosidic angle	G_{L4} *syn*, others *anti*	All *anti*	All *anti*
Stacking	C_{L3}/U_{L1} Ribose (C_{L3})/G_{L4}	T_{L3}/T_{L1} deoxyibose (T_{L3})/G_{L4}	U_{L3}/G_{L4}
Minor groove folding mode	Partial folding of U_{L2}	Folding of T_{L2}	Folding of U_{L2}
Other special features	H-bonding between $C_{L3}NH_2$–U_{L2} backbone phosphate	(1) Hydrophobic interaction of T_{L3} (2) Perpendicular base / base interaction between G_{L4} and T_{L2}	(1) Three potential H-bonds between U_{L2} and stem G•C base pairs (2) Considerable dynamics for C_{L1}, U_{L3} and G_{L4}

[a] r(UUCG) data are from Allain and Varani (1995), and r(CUUG) data from Jucker and Pardi (1995). Closing base-pair residues are underlined, and folding residues *italicized*.
[b] L1, L2, L3, and L4 designate residues 1, 2, 3, and 4 of the tetraloop sequences, respectively.

these two nucleotides have to be in the C2′-endo domain to stretch the loop backbone and bridge the two opposite strands of the stem (Cheong et al., 1990). This requirement is probably strictly followed, since the two loop nucleotides U_{L2}, C_{L3} in the r(UUCG) tetraloop (Allain and Varani, 1995) and U_{L2}, U_{L3} in the r(CUUG) tetraloop (Jucker and Pardi, 1995) were indeed found to reside in the C2′-endo domain. Furthermore, the closing base-paired ribonucleotides C_{L1} and G_{L4} of the r(CUUG) tetraloop were also switched to the C2′-endo domain. All four ribonucleotides in the r(CUUG) tetraloop thus dwell in the C2′-endo domain rarely observed for RNA.

Several other special structural features also deserve mentioning for each tetraloop. In the r(UUCG) case, the loop C_{L3} amino group was proposed to engage in H-bonding with the backbone U_{L2} phosphate to stabilize this unusual loop fold. The polar $U_{L1}2'$-OH in the center of the loop also participates in H-bonding with the $G^6_{L4}O$ to prevent unfavorable entropy loss. However, the unpaired U_{L2} only partially folds into the minor groove and does not interact much with other hairpin residues, while in the d(TTTG) tetraloop, the folding T_{L2} exhibits a perpendicular interaction with the G_{L4} residues, as shown in Figure 6c. Finally in the r(CUUG) case, the U_{L2} also folds into the minor groove, but in a different orientation compared to that in the d(TTTG) case. The $G^2_{L4}NH_2$ does not interact with the U_{L2} base, but instead forms an H-bond with $U^2_{L2}O$. Two other H-bonds between the U_{L2} and the stem G•C pair ($U^3_{L2}N$-G^2NH_2 and $U_{L2}H3$-C^2O) are also proposed (Jucker and Pardi, 1995).

Different minor-groove folding modes for the r(CUUG), d(CTTTG), and d(TTTG) hairpins

Several pyrimidine-rich tetraloop hairpin structures with a closing Py_{L1} • Pu_{L4} base pair have already been described before with most structures highlighted by the folding of an unpaired Py_{L2} into the minor groove (Blommers et al., 1989, 1991; Hilbers et al., 1994; Ippel et al., 1995; Jucker and Pardi, 1995; van Dongen et al., 1996). Our detailed structural studies of the d(CTCCG) (Chou et al., 1999b), d(CTTTG) (Chou et al., 2000) and d(TTTG) hairpins indicate that there are subtypes within such minor-groove folding loop structures. The three selected minor-groove folding modes of the r(CUUG), d(CTTTG), and d(TTTG) hairpins are shown in Figure 8 for comparison. From the figure, it is clear that there is substantial variation with regard to the minor groove width among

these hairpin structures. In the d(CTCCG) (Chou et al., 1999b) and d(CTTTG) hairpins (Chou et al., 2000) (Figure 8b), the unpaired T_{L2} folds rather deep into the minor groove and makes good contacts with the stem nucleotides two or three base pairs away, which is supported by the medium-strength NOEs between the T6-H3–C10-H1′ and T6-H3–T11-H1′ proton pairs (Chou et al., 1999b, 2000). Three possible T6 mediated H-bonds are present when the stem sequence contains the 5′-GC-loop-GC-3′ motif. The two opposite strands are therefore drawn near, with a minor groove width of only 11.8 Å between the C5 phosphate and the T12 phosphate. In the present d(TTTG) loop sequence (Figure 8c), there are only two nucleotides to span the loop, so the unpaired T_{L2} is not able to penetrate deep enough into the minor groove. Therefore the folding T_{L2} is not in a suitable position to participate in H-bonding with the stem G•C base pairs. Instead, it is drawn near to the top loop region, with its base engaging in a perpendicular interaction with the G_{L4} base (also shown in Figure 6c). As a result, a wider minor groove width of approximately 15.7 Å was observed. Such a result is consistent with the H_2O/NOESY data; no NOE was detected between the folding T_{L2} with any stem nucleotide (data not shown), although its resonance peak is rather sharp (Figure 2). It is thus interesting to note that the two different folding modes have a difference of about 4 Å in terms of minor groove width. The situation of r(CUUG) (Figure 8a) is different from those of either d(CTTTG) or d(TTTG) in that, even though it has only two residues in the loop, the folding U_{L2} is still able to make H-bonds with the stem G•C base pair. This is possibly due to the different A-form stem adopted for the r(CUUG) hairpin, which is also characterized by an even wider minor groove width of 17.8 Å. There thus exists considerable difference in the three minor-groove folding modes of the pyrimidine-rich loops.

Conclusions

After detailed structural studies, the reasons why the d(TTTG) oligomer can form a stable hairpin are now clear. We propose that the presence of the T7-methyl group has induced the formation of a hydrophobic core in the loop region. Such a hydrophobic interaction, along with the novel perpendicular guanine/thymine interaction, may be responsible for the stable formation of the d(TTTG) loop hairpin. Several different minor-groove folding modes are available for the pyrimidine-rich loop structures, depending on the loop and stem sequences.

Acknowledgements

We thank the National Science Council and the Chung-Zhen Agricultural Foundation Society of Taiwan, ROC for the instrumentation grants and Y.-Y. Tseng for collecting NMR data. This work was supported by NSC grants 89-2113-M-005-010 to S.H.C.

References

Akke, M., Fiala, R., Jiang, F., Patel, D. and Palmer, A.G.I. (1997) *RNA*, **3**, 702–709.

Allain, F.H.-T. and Varani, G. (1995) *J. Mol. Biol.*, **250**, 333–353.

Altona, C. (1982) *Recl. Trav. Chim. Pays-Bas*, **101**, 413–433.

Avizonis, D.Z. and Kearn, D.R. (1995) *Nucleic Acids Res.*, **23**, 1260–1268.

Bax, A., Griffey, R.H. and Hawkins, B.L. (1983) *J. Magn. Reson.*, **55**, 301–315.

Blommers, M.J.J., van de Ven, F.J.M., van der Marel, G.A., van Boom, J.H. and Hilbers, C.W. (1991) *Eur. J. Biochem.*, **201**, 33–51.

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.

Butcher, S.E., Dieckmann, T. and Feigon, J. (1997) *EMBO J.*, **18**, 7490–7499.

Cate, J.H., Gooding, A.R., Podell, E., Zhou, K., Golden, B.L., Szewczak, A.A., Kundrot, C.E., Cech, T.R. and Doudna, J.A. (1996) *Science*, **273**, 1696–1699.

Chastain, M. and Tinoco Jr., I. (1991) *Prog. Nucleic Acids Res. Mol. Biol.*, **41**, 131–177.

Cheong, C., Varani, G. and Tinoco Jr., I. (1990) *Nature*, **346**, 680–682.

Chou, S.-H. and Tseng, Y.-Y. (1999) *J. Mol. Biol.*, **285**, 41–48.

Chou, S.-H., Tseng, Y.-Y., Chen, Y.-R. and Cheng, J.-W. (1999a) *J. Biomol. NMR*, **14**, 157–167.

Chou, S.-H., Tseng, Y.-Y. and Chu, B.-Y. (1999b) *J. Mol. Biol.*, **292**, 309–320.

Chou, S.-H., Tseng, Y.-Y. and Chu, B.-Y. (2000) *J. Biomol. NMR*, **17**, 1–16.

Chou, S.-H., Tseng, Y.-Y. and Wang, S.-W. (1999c) *J. Mol. Biol.*, **287**, 301–313.

Chou, S.-H., Zhu, L., Gao, Z., Cheng, J.-W. and Reid, B.R. (1996a) *J. Mol. Biol.*, **264**, 981–1001.

Chou, S.-H., Zhu, L. and Reid, B.R. (1994) *J. Mol. Biol.*, **244**, 259–268.

Chou, S.-H., Zhu, L. and Reid, B.R. (1997) *J. Mol. Biol.*, **267**, 1055–1067.

Chou, S.-H., Zhu, L. and Reid, B.R. (1996b) *J. Mol. Biol.*, **259**, 445–457.

Coll, M., Frederick, C.A., Wang, A.H.-J. and Rich, A. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8385–8389.

Costa, M. and Michel, F. (1997) *EMBO J.*, **16**, 3289–3302.

Cruse, W.B.T., Saludjian, P., Biala, E., Strazewski, P., Prange, T. and Kennard, O. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 4150–4164.

Egli, M. and Gessner, R.V. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 180–184.
Fujii, S., Wang, A.H., van der Marel, G., van Boom, J.H. and Rich, A. (1982) *Nucleic Acids Res.*, **10**, 7879–7892.
Giessner-Prettre, C., Pullman, B., Tran-Dinh, S., Neumann, J.-M., Huynh-Dinh, T. and Igolen, J. (1984) *Nucleic Acids Res.*, **12**, 3271–3281.
Hare, D., Shapiro, L. and Patel, D.J. (1986) *Biochemistry*, **25**, 7445–7456.
Hare, D.R., Wemmer, D.E., Chou, S.-H., Drobny, G. and Reid, B.R. (1983) *J. Mol. Biol.*, **171**, 319–336.
Heus, H.A. and Pardi, A. (1991) *Science*, **253**, 191–194.
Hilbers, C.W., Heus, H.A., van Dongen, M.J.P. and Wijmenga, S.S. (1994) In *Nucleic Acids and Molecular Biology* (Eckstein, F. and Lilley, D.M.J., Eds.), Vol. 8, Springer-Verlag, Berlin, Heidelberg, pp. 56–103.
Holbrook, S.R., Cheong, C., Tinoco Jr., I. and Kim, S.-H. (1991) *Nature*, **353**, 579–581.
Huang, C.-H. and Chen, C.W. (1997) *Telomere DNA of Streptomyces Chromosome and Linear Plasmid*, National Yang-Ming University, Taiwan.
Ippel, J.H., Lanzotti, V., Galeone, A., Mayol, L., van den Boogaart, J.E., Pikkemaat, J.A. and Altona, C. (1995) *J. Biomol. NMR*, **6**, 403–422.
James, J.K. and Tinoco Jr., I. (1993) *Nucleic Acids Res.*, **21**, 3287–3293.
Jimenez, A.I., Cativiela, C., Gomez-Catalan, J., Perez, J.J., Aubry, A., Paris, M. and Marraud, M. (2000) *J. Am. Chem. Soc.*, **122**, 5811–5821.
Jucker, F.M. and Pardi, A. (1995) *Biochemistry*, **34**, 14416–14427.
Kennard, O. (1985) *J. Biomol. Struct. Dyn.*, **3**, 205–226.
Kopka, M.L., Yoon, C., Goodsell, D., Pjura, P. and Dickerson, R.E. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1376–1380.
Levitt, M. and Perutz, M.F. (1988) *J. Mol. Biol.*, **201**, 751–754.
Marino, J.P., Schwalbe, H. and Griesinger, C. (1999) *Acc. Chem. Res.*, **32**, 614–623.
Miller, J.L. and Kollman, P.A. (1997) *J. Mol. Biol.*, **270**, 436–450.
Mooren, M.M.W., Pulleyblank, D.E., Wijmenga, S.S., van de Ven, F.J.M. and Hilbers, C.W. (1994) *Biochemistry*, **33**, 7315–7325.
Nishinaka, T., Ito, Y., Yokoyama, S. and Shibata, T. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 6623–6628.
Nishio, M., Umezawa, Y., Hirota, M. and Takeuchi, Y. (1995) *Tetrahedron*, **51**, 8665–8701.
Pearlman, D.A. and Kollman, P.A. (1990) *Biopolymers*, **29**, 1193–1209.
Perutz, M.F., Fermi, G., Abraham, D.J., Poyart, C. and Bursaux, E. (1986) *J. Am. Chem. Soc.*, **108**, 1064–1078.
Plateau, P. and Gueron, M. (1982) *J. Am. Chem. Soc.*, **104**, 7310–7311.
Pley, H.W., Flaherty, K.M. and McKay, D.B. (1994a) *Nature*, **372**, 111–113.
Pley, H.W., Flaherty, K.M. and McKay, D.B. (1994b) *Nature*, **372**, 68–74.
Sarma, R.H., Mynott, R.J., Wood, D.J. and Hruska, F.E. (1973) *J. Am. Chem. Soc.*, **95**, 6457–6459.
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.
Shaka, A.J., Barker, P.B. and Freeman, R.J. (1985) *J. Magn. Reson.*, **64**, 547–552.
Sich, C., Ohlenschlager, O., Ramachandran, R., Gorlach, M. and Brown, L.R. (1997) *Biochemistry*, **36**, 13989–14002.
Sklenar, V., Miyashiro, H., Zon, G., Miles, H.T. and Bax, A. (1986) *FEBS Lett.*, **208**, 94–98.
Strobel, S.A. and Cech, T.R. (1993) *Biochemistry*, **32**, 13593–13604.
Teng, M.-k., Usman, N., Frederick, C.A. and Wang, A.H.-J. (1988) *Nucleic Acids Res.*, **16**, 2671–2690.
Tseng, Y.-Y. and Chou, S.-H. (1999) *J. Chin. Chem. Soc.*, **46**, 699–706.
Tuerk, C., Gauss, P., Thermes, C., Groebe, D.R., Gayle, M., Guild, N., Stormo, G., d'Aubenton-Carafa, Y., Uhlenbeck, O.C., Tinoco Jr., I., Brody, E.N. and Gold, L. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 1364–1368.
Uhlenbeck, O.C. (1990) *Nature*, **346**, 613–614.
van Dongen, M.J.P., Wijmenga, S.S., van der Marel, G.A., van Boom, J.H. and Hilbers, C.W. (1996) *J. Mol. Biol.*, **263**, 715–729.
Varani, G., Cheong, C. and Tinoco Jr., I. (1991) *Biochemistry*, **30**, 3280–3289.
Wang, A.H.-J., Gessner, R.V., van der Marel, G.A., van Boom, J.H. and Rich, A. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 3611–3615.
Wang, A.H.-J., Quigley, G.J., Kolpak, F.J., van der Marel, G., van Boom, J.H. and Rich, A. (1981) *Science*, **211**, 171–176.
Williams, D.J. and Hall, K.B. (2000a) *J. Mol. Biol.*, **297**, 1045–1061.
Williams, D.J. and Hall, K.B. (2000b) *J. Mol. Biol.*, **297**, 251–265.
Woese, C.R., Winker, S. and Gutell, R.R. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 8467–8471.
Zhu, L., Chou, S.-H. and Reid, B.R. (1995) *Nat. Struct. Biol.*, **2**, 1012–1017.
Zichi, D.A. (1995) *J. Am. Chem. Soc.*, **117**, 2957–2969.