DNA hairpin loops in solution. Correlation between primary structure, thermostability and reactivity with single-strand-specific nuclease from mung bean

Luigi E.Xodo*, Giorgio Manzini, Franco Quadrifoglio[1], Gijs van der Marel[2] and Jacques van Boom[2]
Department of Biochemistry, Biophysics and Macromolecular Chemistry, University of Trieste, I-34127, [1]Institute of Biology, Faculty of Medicine, University of Udine, I-33100, Italy and [2]Gorlaeus Laboratories, State University of Leiden, Leiden, The Netherlands

Received December 31, 1990; Revised and Accepted March 1, 1991

ABSTRACT

Hairpin structures formed by seven DNA inverted repeats have been studied by PAGE, UV(CD)-spectroscopy and nuclease cleavage. The hairpins consisted of $(CG)_3$ stems and loops of 2, 3 and 4 residues. Thermal stabilities (T_m) have been determined in low and high ionic strength buffers, where the hairpins were structured in the B- and Z-DNA form respectively. The thermodynamic parameters of hairpin formation have been obtained by a two-state analysis of the hairpin-coil transitions. It is found that, on increasing the number of bases in the loop from 2 to 3 and 4, the T_ms of the B-hairpins decrease, whereas the T_ms of the same hairpins in the Z-form increase. This confirms previous evidence (1,2) that in a hairpin molecule the size and structure of the loop are modulated by the conformation of the helical stem. Moreover, B-hairpins with loops comprising 2, 3 and 4 bases have been digested with the single-strand-specific nuclease from mung bean. In our experimental conditions (0°C) the nuclease preferentially cleaves the unbonded nucleotides of the loops. However, the rates of loop hydrolysis, which roughly follow a first-order kinetics, markedly depend on the size of the loop. At a ratio of 3 enzyme units/μg DNA, the half-lives of hairpins which are expected to form loops of 4, 3 and 2 residues are 90, 145 and 440 minutes respectively. Thermostability and enzymatic digestion data suggest that two-membered loops can be formed in B-hairpins but not in Z-hairpins.

INTRODUCTION

Palindromes, the sequences of two-fold symmetry often found in the regulatory regions of the genome (3), have been widely investigated for their ability to form cruciform structures. The DNA cruciforms comprise two opposed hairpins and are supposedly involved in gene expression (4). In order to understand the stability and formation of DNA cruciforms, it is important to learn about the structure and thermostability of individual hairpins, in particular as a function of the DNA sequence.

Significative contributions to the physico-chemical characterization of hairpin structures have been obtained with studies based on UV-absorption (1,5−8), NMR (9−17), X-ray crystallography (18) and enzymatic digestions (19,20).

Although these studies allowed a good description of hairpin structures, some aspects of the DNA folding in solution remain unclear. For instance, it is not well understood whether (i) the formation of two-membered loops in B-DNA hairpins is sequence-dependent; (ii) the type of base pair closing the loop affects its structure; (iii) the conformation of the stem helix influences the structure of the loop; (iv) the nucleotides in the loop region are differently exposed to the action of hydrolyzing agents like single-strand-specific nucleases.

Previous investigations have shown that short palindromic sequences exhibit in solution the following transformations:

$$\text{Duplex} = 2\ \text{Hairpins} = 2\ \text{Single strands}$$

While the duplex form is favoured by high DNA strand concentration (7), low temperature and increased ionic strength (1,7), the unimolecular hairpin form is strongly favoured by mismatched base pairs at the centre of the sequence (12) and by a high C+G content in the arms of the palindrome (21). The thermodynamic data of hairpin formation are useful for studying the structure of hairpins in solution (1,5−8). They should be regarded as complementary to those obtained by NMR-spectroscopy and X-ray crystallography, so that the combination of data of different origin should allow to develop a general principle on the folding of DNA in solution.

In this work we have synthesized seven inverted repeat sequences, of the type $(m^5CG)_3$-X_n-$(m^5CG)_3$ (n= 2, 3, 4) and $(CG)_3$-X_n-$(CG)_3$ (n=2,3), which can form hairpin structures

* To whom correspondence should be addressed

with loops comprising two, three and four nucleotides. The hairpin-coil transitions of these molecules have been studied in three different media, and the secondary structures of the hairpins have been correlated with the thermodynamic parameters of hairpin formation. In addition, the methylated hairpins have been digested by the single-strand-specific nuclease from mung bean and the loop size has been correlated with the digestion rates.

MATERIALS AND METHODS

Oligodeoxynucleotides

The hairpins investigated in this study, numbered from **1** to **7**, are shown in Figure 1. They have been synthesized in solid-phase, according to a modified phosphotriester method (22). Purification was performed by standard gel permeation chromatography using a G-50 Sephadex resin and tetraethylammonium bicarbonate as eluent. The purity of the samples was checked by ion exchange HPLC and PAGE electrophoresis.

Buffers

The high stability of hairpins **1**–**7** imposed the use of a low concentration buffer: 0.5 mM Tris-HCl (pH=7.4), 0.5 mM NaCl, 0.1 mM EDTA (buffer A). However, since hairpin **1** did not melt completely even in buffer A, samples **1**–**3** have been melted also in 2 mM Tris.HCl (pH=7.4), 0.5 mM NaCl, 0.1 mM EDTA, 20% formamide (buffer B). Furthermore, hairpins **1**–**4** and **7** have been melted in a high ionic strength medium, where they are structured in the left-handed Z-DNA conformation, containing 5 (or 4.6) M $NaClO_4$, 10 mM sodium phosphate (pH=7.4), 0.1 mM EDTA (buffer C).

Optical experiments

The UV-absorbance versus temperature profiles were obtained with a Cary 219 (Varian) spectrophotometer. The temperature was increased at a rate of 0.5°C/min by means of a Haake PG 10 temperature programmer, connected to a Haake water-circulating bath. The melting curves were monitored at 270 nm in both denaturation and renaturation. The oligomer concentrations were determined by UV absorption in the denaturated state, assuming as extinction coefficients 7500, 8500, 12500 and 15000 $M^{-1}cm^{-1}$ for C, T, G and A, respectively (23). The DNA concentrations used in UV melting varied from 1.5 to 40 μM in strands.

Circular dichroism spectra were recorded with a Jasco J-500 A spectropolarimeter, equipped with a thermostatted cuvette holder which allowed measurements at controlled temperature.

Gel electrophoresis

PAGE experiments were carried out on 20% polyacrylamide gels, run under non-denaturing conditions (Tris.HCl 50 mM, NaCl 10 mM, pH=7.4), in a thermostatted apparatus. Electrophoreses were carried out at 30 mA and 10–15 V/cm. Bromphenol blue dye was used as marker and the gels were stained with a solution of 'stains all' dye dissolved in 1:1 water-formamide.

Cleavage of hairpin loops by mung bean nuclease

Solutions of **1**, **2** and **3** in 50 mM acetate (pH=5), 10 mM NaCl, 10 μM $ZnCl_2$ (about 10 μg in 20 μl of buffer) were heated at 90°C for 15′ and immediately cooled in a bath at 0°C. This procedure ensured the total conversion of the samples in the unimolecular hairpin structure (see '**Electrophoresis**'). After equilibration at 0°C (15′), an amount of mung bean nuclease (from Sigma) dissolved in the reaction buffer (4.0 units/μl) was added to each solution. Aliquots (5 μl) of the reaction mixture, taken at increasing intervals, were quenched in 70 μl of NaOH at pH=12. The strong alkaline pH served to stop the enzymatic reaction and to denature the DNA. The denatured DNA fragments were separated by HPLC according to their size on a TSK DEAE-5W column (Beckman), using a gradient of $NaClO_4$, from 0 to 1 M in 300 minutes (see Figure 5). The DNA was detected and quantified through its absorbance at 260 nm, by using an integrator recorder (Waters 740 Data Module).

Analysis of transition data

The thermally induced hairpin-coil transitions were analyzed according a two-state model. Using the non-linear least-square algorithm of Marquardt (24), the experimental absorbance curves were fitted to:

$$A(T) = \epsilon_c C_T /(1+K^{-1}) + \epsilon_h C_T/(1+K) \qquad \text{(ref. 21)}$$

where ϵ_c and ϵ_h are the extintion coefficients of the oligomer in the coil and hairpin state respectively, C_T is the oligomer strand concentration and K is the equilibrium coil/hairpin constant which is related to the thermodynamic parameters by $K=\exp(-\Delta H/RT + \Delta S/R)$. The algorithm provided for the adjusted ΔH and ΔS parameters the 95% confidence limits, which, in general, had a magnitude of ±7%. The semitransition temperature T_m was determined with an accuracy of ±0.5°C.

RESULTS

Design of DNA hairpins

In virtue of their two-fold simmetry the palindromic sequences **1**–**7** can assume in solution either a linear duplex or a folded hairpin conformation. The number of bases intervening between the two complementary arms of **1**–**7** is variable, so that the palindromes can, in principle, form DNA hairpins with a loop of two, three and four residues as shown in Figure 1. The structures of hairpins similar to **2**, **7** and **3** have been determined by NMR (11, 13). The first three hairpins, whose cytosines are methylated in C_5, were designed to correlate the size of the loop with their thermostability and their reactivity with the single-stranded specific nuclease from mung bean, which cleaves the unbonded nucleotides of the loop. Hairpins **4**, **5** and **6** were planned to determine: (i) whether the substitution in **6** of the TT adjacent to the stem with either a GC (hairpin **4**) or a CG (hairpin **5**) does result in a substantial increment of the hairpin T_ms, as consequence of the formation of a two-membered loop; (ii) whether flipping the base pair closing the central TT from G-C (hairpin **4**) to C-G (hairpin **5**), influences the T_ms of the two

	T	T T				T
T A	T T	T T	T TT	T T	T T	T T
G-C	G-C	G-C	**G-C**	**C-G**	**T T**	G-C
C-G	C-G	C-G	C-G	C-G	C-G	C-G
G-C	G-C	G-C	G-C	G-C	G-C	G-C
C-G	C-G	C-G	C-G	C-G	C-G	C-G
G-C	G-C	G-C	G-C	G-C	G-C	G-C
C-G	C-G	C-G	C-G	C-G	C-G	C-G
1	**2**	**3**	**4**	**5**	**6**	**7**

Figure 1: DNA hairpins; C= 5-CH_3-Cytosine

hairpins. Finally, most of the **1–7** sequences are strongly Z-helicogenic, therefore they can be studied in a high ionic strength buffer, where they are structured in the Z-DNA conformation. This allows a comparative study between hairpins in the B- and Z-DNA conformations.

Gel electrophoresis

PAGE shows that sequences **1–7** each migrate in a polyacrylamide gel as a slowly interconverting mixture of hairpin and duplex forms, as previously observed for similar palindromic sequences (1,21). The relative fractions of the two forms depend (i) on the size of the sequence intervening between the arms of the inverted repeat; (ii) on the thermal treatment of the samples prior to loading in the gel, (iii) on the ionic strength, as well as the DNA concentration. This is exemplified by the methylated sequences **1–3** in Figure 2. Gel A, run at 5°C, shows that all the three sequences migrate, at the concentration of 3–4 mM (in bases), as a mixture of duplex and hairpin forms. It is noteworthy that four thymidines are sufficient to destabilize the duplex of **3** (lane a), while the absence of mismatched pairs in **1** results in a very small amount of folded structure (lane c). By contrast, **2** migrates with comparable quantities of duplex and hairpin forms. However, at 50°C the mobility of **1–3** is similar to that of 6-mer d(m^{5}CGm^{5}CGm^{5}CG), which clearly indicates that at this temperature the samples completely convert into the hairpin conformation (gel B). Finally, the effect of heating **1**, **2**, **3** before loading them in a gel, which was pre-cooled at 5°C, is shown in gel C. In this case each sequence migrates only as a hairpin during the entire time of the experiment (7h). This is due to the very slow kinetics of the hairpin-to-duplex transformation of **1–3**.

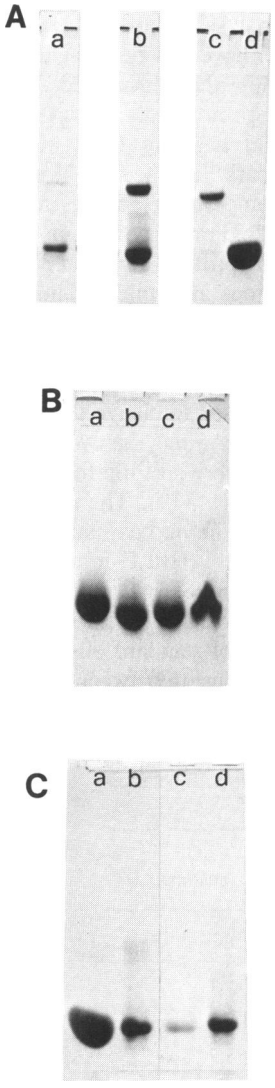
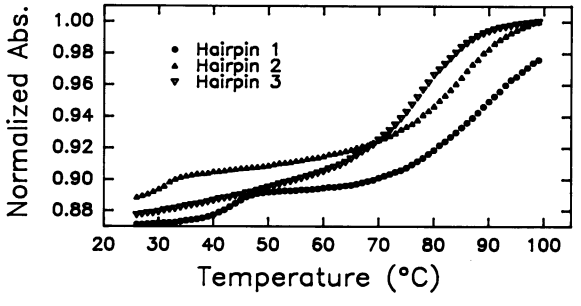


Figure 2. (A) 20% PAGE of **1** d(CGCGCGTTTTCGCGCG) (lane a), **2** d(CGCGCGTTTCGCGCG) (lane b), **3** d(CGCGCGTACGCGCG) (lane c) and **8** d(CGCGCG) (lane d), where C=m^{5}C, in 50 mM Tris-HCl (pH=7.4), 10 mM NaCl, 5°C; (B) PAGE as above but at 50°C; (C) PAGE in which the samples have been heated at 90°C for 5′ before loading them in the gel, pre-cooled overnight at 5°C.

Melting experiments

Absorbance versus temperature profiles for each hairpin were obtained in buffer A, at different DNA concentrations. Figure 3, top panel, shows the denaturation profiles for **1**, **2** and **3**. In accordance with the electrophoretic results, **1** and **2** exhibited biphasic profiles, characterized by a duplex-hairpin transition at low temperature and a hairpin-coil transition at high temperature.



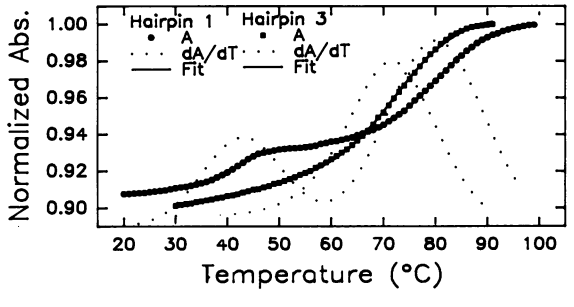


Figure 3. (Top) Absorbance (270 nm), normalized for the value at 99°C, versus temperature profiles for **1** d(CGCGCGTACGCGCG), **2** d(CGCGCGTTTCGCGCG) and **3** d(CGCGCGTTTTCGCGCG), where C=m^{5}C, in buffer A. The concentration of samples **1**, **2** and **3** were: 0.13, 0.17 and 0.24 mM/base, respectively. The curve of hairpin **1** has been shifted down by 0.02 units for design clarity; (Bottom) Absorbance versus temperature profile of **1** d(CGCGCGTACGCGCG) and **3** d(CGCGCGTTTTCGCGCG) in the presence of 20% formamide (buffer B).

By contrast, **3** as well as **4**–**7** (not shown) only showed the hairpin-coil transition, because in these sequences, due to the low ionic strength of buffer A, the duplex-hairpin transition is detected at DNA concentrations higher than those used for the melting experiments (1.5–40 μM in strands). At a heating rate of 0.5°C/min the hairpin-coil transition was fully reversible, and the curves were analyzed with a two-state model. The thermodynamic data of the hairpin-coil transition of **1**–**7**, obtained on such assumption, are reported in Table I. Although the ionic strength of buffer A was very low, the methylated hairpins **1**–**3** were found to be very stable in this medium. For instance, at 99°C **1** is not fully denatured, pointing to a T_m not lower than that of polydGdm^{5}C:polydGdm^{5}C (T_m = 86°C) at the same ionic strength (25). Hence, hairpins **1**–**3** were melted in the presence of formamide. The effect of formamide on the denaturation of DNA has been previously investigated (26). We melted **1**–**3** in buffer B, containing 20% (v/v) formamide. In such a buffer the circular dichroism spectrum of each hairpin did not significantly change, suggesting that formamide should not appreciably alter the B-DNA conformation. The relative stabilities of **1**–**3** in buffer B follow the same order as that observed in buffer A: **1** > **2** > **3**. Figure 3, bottom panel, shows representative denaturation profiles for **1** and **3**. The results of the thermodynamic analyses in buffer B are displayed in Table I.

Hairpins **1**–**4** and **7** have also been melted in buffer C. In this buffer **1**–**4** and **7** adopt the left-handed Z-DNA conformation as revealed by CD spectra (not shown). Each CD spectrum exhibits a strong and negative band at 198 nm, which is considered the hallmark for the Z-DNA conformation (27). Absorbance versus temperature profiles for each hairpin were determined over a tenfold range nucleotide concentration. Figure 4 shows typical melting profiles for **1** and **3**. The thermodynamic data for these transitions are collected in Table II.

Table I. Thermodynamic parameters for B-DNA hairpin formation

Hairpin[a]	T_m[b] (°C)	−ΔH (kcal/mol)	−ΔS (e.u.)	−ΔG(37°C) (kcal/mol)
a) in 0.5 mM Tris.HCl, 0.5 mM NaCl, 0.1 mM EDTA, pH=7.4.				
1 (C̲G)$_3$TA(C̲G)$_3$	86–88	(not fully melted at 98°C)	–	
2 (C̲G)$_3$T$_3$(C̲G)$_3$	82	51 (±3)	143 (±8)	6.6
3 (C̲G)$_3$T$_4$(C̲G)$_3$	76	54 (±3)	155 (±8)	5.9
4 (CG)$_3$T$_2$(CG)$_3$	78	47 (±3)	134 (±8)	5.4
5 (CG)$_2$CCTTGG(CG)$_2$	82	48 (±3)	135 (±8)	6.1
6 CGCGCT$_4$GCGCG	72	42 (±2)	122 (±6)	4.1
7 (CG)$_3$T$_3$(CG)$_3$	74	51 (±3)	147 (±9)	5.4
8 C̲GC̲GC̲G		54 (±2)	–	
9 CGCGCG		51 (±2)	–	
b) in 2mM Tris.HCl, 0.5 mM NaCl, 0.1 mM EDTA, 20% formamide (pH=7.4):				
1 (C̲G)$_3$TA(C̲G)$_3$	80	43 (±2)	121 (±4)	5.5
2 (C̲G)$_3$T$_3$(C̲G)$_3$	76	46 (±3)	131 (±8)	5.4
3 (C̲G)$_3$T$_4$(C̲G)$_3$	72	47 (±3)	135 (±9)	5.1

ΔH, ΔS and T_m are the average of several values obtained from independent melting curves, and are expressed as round numbers. The number in parenthesis is the 95% confidence limits; (a) C̲= 5-CH$_3$-cytosine (b) Error on T_m is ±0.5°C.

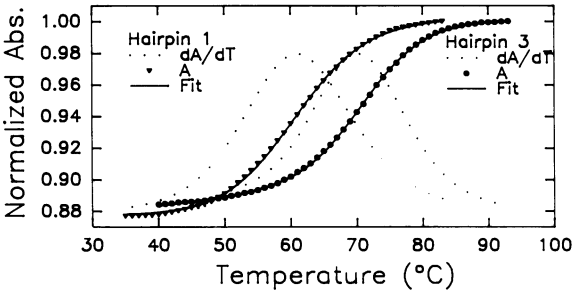


Figure 4. Normalized absorbance (270 nm) versus temperature profile for **1** d(C̲GC̲GC̲GTAC̲GC̲GC̲G) and **3** d(C̲GC̲GC̲GTTTTC̲GC̲GC̲G), where C̲=m^{5}C, in the Z-DNA conformation (buffer C). Sample concentrations were 0.032 and 0.2 mM/base for **1** and **3** respectively.

Cleavage of hairpin loops by mung bean nuclease

Single-strand-specific nucleases from *Aspergillus oryzae* (S-1) and mung bean have been used to cleave the unbonded nucleotides of hairpin loops (19,20). Among the two nucleases, the one from mung bean was found more specific for cutting the loop region of small hairpin molecules (19). The methylated hairpins **1**, **2** and **3** were treated with mung bean nuclease and the digestion products were analyzed by HPLC in denaturing conditions (at pH=12 as described above), where the DNA fragments are separated according to their size (28). Figure 5 reports the HPLC separation of a number of standard oligomers of different size, exhibiting a linear relationship between retention time and DNA size (insert). The enzymatic reactions were performed in the same

Table II. Thermodynamic parameters for Z-DNA hairpin formation. Data have been obtained in 10 mM phosphate, 5 M (or 4.6 M) NaClO$_4$, pH=7.4

Hairpin[a]	T_m[b] (°C)	−ΔH (kcal/mol)	−ΔS (e.u.)	−ΔG(37°C) (kcal/mol)
1 (C̲G)$_3$TA(C̲G)$_3$	61	42 (±3)	126 (±9)	2.9
2 (C̲G)$_3$T$_3$(C̲G)$_3$	70	50 (±3)	146 (±10)	4.7
3 (C̲G)$_3$T$_4$(C̲G)$_3$	71	51 (±3)	148 (±10)	5.1
4 (CG)$_3$T$_2$(CG)$_3$[c]	56	40 (±2)	122 (±6)	2.2
7 (CG)$_3$T$_3$(CG)$_3$[c]	59	43 (±2)	129 (±6)	3.0
T$_4$ (CG)$_3$T$_4$(CG)$_3$[d]	64	45	133	3.7
TATA (CG)$_3$TATA(CG)$_3$[e]	67	–	–	–
T$_5$ (CG)$_3$T$_5$(CG)$_3$[f] [c]	63	50 (±3)	148 (±7)	4.1
(CGCGCG)[g]	–	48	–	–
(C̲GC̲GC̲G)	–	49 (±2)	–	–

ΔH, ΔS and T_m are the average of several values, and are expressed as round numbers. The number in parenthesis is the 95% confidence limits;(a) C̲= 5-CH3-Cytosine; (b) Error on Tm is ±0.5°C; (c) data obtained in 4.6 M NaClO$_4$4; (d) see ref. 14, data obtained in 4 M NaClO$_4$4; (e) see ref. 35, data obtained in 4 M NaClO$_4$; (f) see ref. 2; (g) see ref. 36.

conditions used by Baumann et al. to digest the hairpin loops formed by d(ATCCTA-An-TAGGAT) (n=3−6) (19). The reactions were carried out at 0°C to (i) minimize fraying ends in the hairpin molecules; ii) lower enzyme catalysis. Figure 6 shows typical HPLC profiles for **1**−**3**, after an incubation with mung bean nuclease for 272 minutes at 0°C. The primary digestion products of **1**−**3** are DNA fragments eluting with retention times comprised between that of the standard 6-mer d(m^{5}CGm^{5}CGm^{5}CG) (34.0 min) and 11-mer d(m^{5}CTm^{5}C-Tm^{5}CTTm^{5}CTTm^{5}C) (41.3 min) oligomers. This indicates that, in the experimental conditions used, the mung bean nuclease cleaves predominantly the loop region of **1**−**3**. For prolonged incubation times, and in particular for hairpins **2** and **3**, another peak appeared in the chromatograms, with a retention time (19 min) comparable to that observed for the single mononucleotides. This suggests that the nuclease removes single nucleotides from the loop region and from the fraying ends of the stem helix, probably present in minor extent even at 0°C. As a control experiment the 6-mer duplex formed by d(m^{5}CGm^{5}CGm^{5}CG), corresponding to the stem of **1**−**3**, was treated with mung bean nuclease. At 0°C the 6-mer was quite resistent to the nuclease, even after extended incubation. The only observable digestion product was a species eluting at about 19 min, attributable to nucleotides removed from the fraying ends. By comparing the retention times of the main digestion products of **3** (two overlapped peaks at 34.2 and 35.2 minutes and one at 39 minutes) with the retention times on the calibration line (insert of Fig. 5), one finds that this 16-mer hairpin is cleaved in a mixture of 6/7-mer and 9-mer fragments. Hairpin **1** is cut in a mixture of 6-mer and 8-mer fragments.

The reproducibility of these cleavage patterns was assessed by repeated enzymatic digestions at different enzyme/DNA ratios. Insert of Figure 6 reports the fraction of hairpin loops not yet cleaved by the nuclease as a function of increasing incubation times. It is readily seen that the rate of loop hydrolysis strongly depends on the size of the intervening sequence, namely TA in **1**, TTT in **2** and TTTT in **3**. For instance, after an incubation time of 272 minutes the fraction of uncleaved hairpins **1**, **2** and **3** are 0.7, 0.3 and 0.2 respectively. This variation in the rate of digestion reflects different loop structures in the three hairpins. The loss of hairpin loop roughly follows a first-order kinetics; from the slopes of these linear plots the half-lives for hairpins **3**, **2** and **1** resulted to be 90, 145 and 440 minutes respectively, showing that the rates of loop hydrolysis observed in hairpins **1**−**3** increase the order TA<TTT<TTTT.

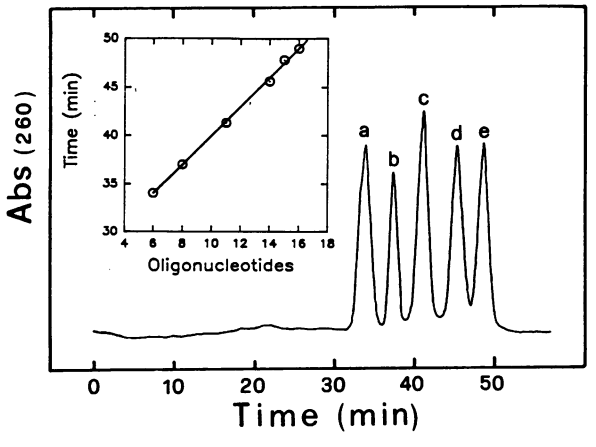


Figure 5. HPLC separation of oligodeoxynucleotides on a TSK DEAE-5PW column using a gradient of $NaClO_4$ (0 to 1 M in 300 min, 1ml/min) in NaOH at pH=12. The DNA oligomers: (a) d(CGCGCG); (b) d(CGCGAGCG); (c) d(CTCTCTTCTTC); (d) d(CGCGCGTACGCGCG); (e) d(CGCGCGTTTT-CGCGCG), where C=m^{5}C, are denatured at pH=12 and separated according to their size. Insert reports the the retention time vs DNA size plot. The retention time of d(CGCGCGTTTCGCGCG) (47.4 min) is also reported.

DISCUSSION

Thermodynamics of B-DNA hairpins formation

The thermostability of a hairpin structure depends on the extent to which the negative enthalpic term from the helix stem, and from the stacking of bases in the loop, outweigh the positive

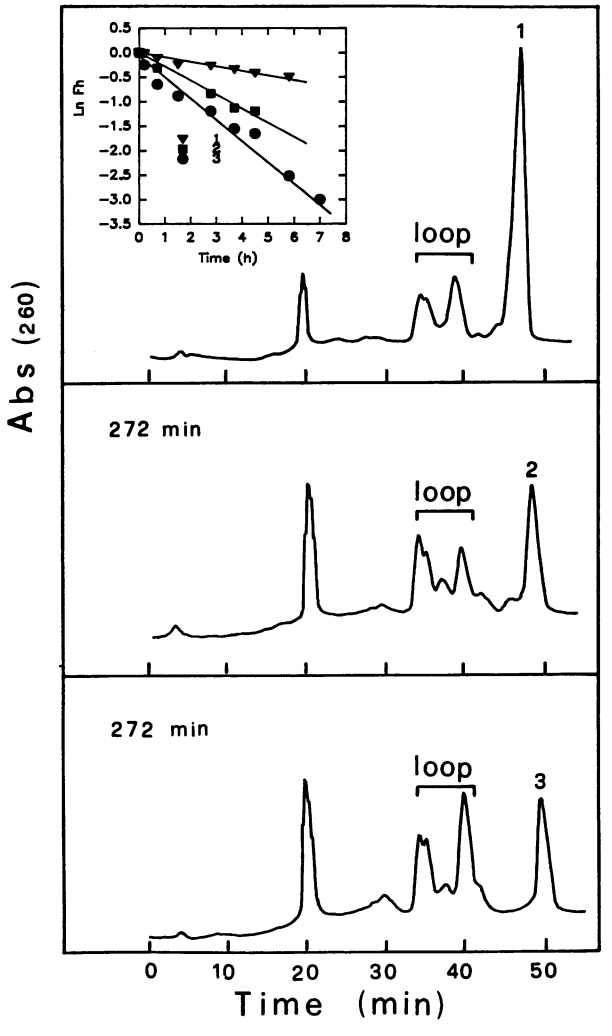


Figure 6. HPLC analyses of the digestion products obtained by incubations of **1** d(CGCGCGTACGCGCG), **2** d(CGCGCGTTTCGCGCG) and **3** d(CGCGC-GTTTTCGCGCG) with mung bean nuclease for 272 minutes, in 50 mM acetate (pH=5), 10 mM NaCl, 10 μM $ZnCl_2$, 0°C. Separations have been performed on a TSK DEAE-5W column used in denaturing condition (NaOH at pH=12) with a gradient of 0−1 M $NaClO_4$, over a period of 300 minutes. The retention times of the digestion products of length varying from 6 to 10 bases, which are originated by the nuclease cleaving the loop region of hairpins **1**, **2** and **3**, fall in the interval denoted by 'loop'. Insert: Rate of disappearance of hairpins **1**, **2**, **3** as a function of incresing digestion times. Fh is the fraction of uncleaved hairpins. The hairpin half-live values were determined from the slopes of ln Fh versus t plots.

entropic $-T\Delta S$ term for hairpin formation. Thus the T_m of a hairpin, which depends on the balance between ΔH and ΔS, is a parameter associated with the structure of the hairpin, in terms of stem and loop size. Table I shows that the thermostabilities of hairpins **1**, **2** and **3** decrease as the sizes of the intervening sequences increase from two, to three and to four nucleotides. This trend is obtained also in the presence of the destabilizing agent formamide (Table I). This finding is in keeping with the presence in **1** of a two-membered loop structure. In fact, if we assume that **1** forms a four-membered loop, its T_m should be considerably lower than that of **3**, since **1** would have a stem with one C-G pair less than that of **3**. On the contrary, it is experimentally observed that **1** is 8°C more stable than **3**. A similar conclusion can be inferred by analysing the thermodynamic data for **4**–**7** (Table I). Hairpin **6** can be used for an interesting comparison with **4** and **5**: if **4** and **5** were to form a four-membered hairpin loop structure by breaking the last C-G or G-C pair they should exhibit a stability comparable to that of **6**. However, the finding that **4** and **5** are 6°C and 10°C more stable than **6**, (1.3 and 2.0 kcal/mol in free energy), is a good piece of evidence that **4** and **5** should be characterized by a two-membered loop. Arguments could be issued suggesting that all **4**, **5**, **6** form a four-membered loop and that the different stabilities observed are due to a different composition of the loops: TTTT vs CTTG vs GTTC. This can be ruled out on the basis of reported data showing that when a TTTT loop is substituted with a CCCC or GGGG, the Tm is influenced by about 1–2°C: in our case the effect would be even lower (1, 8, 29–31). The higher stability of **5** with respect to **4** (4°C) results from a different orientation of the base pair closing the TT loop: G:C (hairpin **4**) versus C:G (hairpin **5**) (Figure 1). The physical origin of this phenomenon may be due to: (i) different nearest-neighbour interactions of the last C-G or G-C pair over the preceding C-G pair; (ii) the fact that a G-C pair closing a two membered TT loop has been found to be markedly buckled: a feature that seems to create more hindrance for hydrogen bond formation than for vertical stacking interactions (16). This effect has also been observed in RNA hairpins (32–34). The occurrence of two-membered hairpin loops in B-DNA has been previously reported (1,12,16). Energy-minimized structures of hairpin loops showed that the phosphate-phosphate distance across the minor groove of a B-helix, 17.5 A, can be bridged by only two nucleotides. The proposed structure of a two-membered loop shows that the base-base stacking interactions are propagated from the 3-side of the stem, involving the first loop residue, and to a minor extent, the second one (12). Very recently it has been reported that also the RNA sequence r(GGACUUCGGUCC) adopts in solution a hairpin structure with a loop of only two nucleotides (U_6 and C_7) with G_8 of the U_5-G_8 base pair in the syn conformation (17).

Loop cleavage by mung bean nuclease

Single-strand-specific nucleases bind single-strand DNA and cut the O-3′-P phosphodiester bond in the presence of divalent cation cofactors (20). The nuclease from mung bean has been successfully used as a structural probe to investigate the structure and the length of the loop in hairpin molecules formed by d(AT-CCTA-An-TAGGAT) (n=3,4,5,6)(19). The observation that the rates of loop hydrolysis of **1**–**3** markedly decrease as the length of the intervening sequence decreases from four to three and two nucleotides, suggests that the enzyme-substrate interactions of **1** are weaker compared to those of **2** and **3**. This is in keeping with the view that **1**–**3** should form loops of different structure and size: while the nucleotides of a four-member loop are sufficiently flexible or exposed in the solvent to interact positively with the nuclease, the nucleotides in a two-member loop are more constrained and therefore less keen to recognize the catalytic site of the enzyme.

Thermodynamics of Z-DNA hairpin formation

Table II shows that the relative stabilities of **1**–**3** in the Z-DNA form as a function of loop size follow a trend which is opposite to that observed for the same hairpins in the B-DNA form: **1** is less stable than **2** and **3**. The same stability trend holds true for **4**, **7**, [T4], [TATA] and [T5]. This behaviour is a compelling piece of evidence that changing the conformation of the stem helix from B to Z influences the structure of the loop.

Comparisons between hairpins in the B- and Z-DNA conformations have been previously reported. It has been found that the d(CGCGCGCGCGTTTTCGCGCGCGCG) hairpin in the B-form shows loop-to-stem connectivities at both sides of the loop (between G_{10}-T_{11} and T_{14}-C_{15}), while in the Z-DNA form, in 6 M $NaClO_4$, only the connectivity of the 5′side is seen (14). The authors suggested that the stereochemistry of the T_4 loop should depend on the conformation of the stem helix. A qualitatively similar result was obtained for the d(CGCGCGTTTTCGCGCG) in 4 M $NaClO_4$ (15). In this case no loop-stem connectivities from either sides of the loop were found. By contrast, when the hairpin is in the B-DNA form it shows a connectivity between T_7 and G_6 (13). Thus, these data lend support to the notion that a loop built on a B-DNA helix adopts a different structure with respect to that built on a Z-DNA helix.

The crystal structure of hairpin d(CGCGCGTTTTCGCGCG) has been resolved (18). It was found that the stem of the hairpin was in the Z-DNA conformation and that the four thymidines were rotated away: only T_8 and T_9 stack on one another. Hence, both crystallography and NMR showed that a loop built on a Z-DNA helix is rather extended and not, as found for B-DNA, inward stacked (16). Unlike what is formed in B-DNA, the minor groove in Z-DNA is not uniform and the interstrand phosphate-phosphate distance varies from 12-to-15 A (18), since the helical repeat is a dinucleotide. Therefore, it is not surprising that B-DNA and Z-DNA form different loops.

Following the rationale of correlating the thermostabilities with hairpin structures, the data of Table II suggest that **1** is less stable than **2** and **3** (−2.9 vs −4.7 and −5.1 kca/mol in free energy) as it forms a four-membered loop by breaking the last G-C pair. This interpretation is consistent with the lower enthalpic content of **1** (42 kcal/mol) with respect to that of **2** (50 kcal/mol) and **3** (51 kcal/mol).

ACKNOWLEDGMENTS

This work was supported by the Italian Ministry of Education and the Netherlands Organization for Advancement of Pure Research.

REFERENCES

1. Xodo,L.E., Manzini,G., Quadrifoglio,F., van der Marel, G.A.& van Boom, J.H. (1988) Biochemistry 27, 6321–6326.
2. Xodo,L.E., Manzini,G., Quadrifoglio,F., van der Marel, G.A. & van Boom, J.H. (1988) Biochemistry 27, 6327–6331.
3. Rosenberg, M.& Court, D.(1979) Annu. Rev. Genet. 13, 319–351.
4. Sinden, R.R. & Pettijohn, D.E.(1984) J.Biol.Chem.259, 6593–6600.

5. Scheffler, I.E., Elson, E.L.& Baldwin, R.L. (1968) J.Mol.Biol. 36, 291–304.
6. Gralla, J.& Crothers,D.M.(1973) J.Mol.Biol. 73, 497–511.
7. Marky, L.A., Blumenfeld,K.S., Kozlowski,S. & Breslauer, K.J.(1983) Biopolymers 22, 1247–1257.
8. Groebe, D. & Uhlenbeck, O. (1988) Nucl.Acids Res. 16, 11725–11735.
9. Haasnoot, C.A.G., Hilbers, C.W., van der Marel, G.A., van Boom, J.H., Singh, U.C., Pattabiraman, N.& Kollman, P.A. (1986) J. Biomol. Struct. Dynam. 3, 843–857.
10. Wemmer, D.E., Chou, S.H., Hare, D.R.& Reid, B.R. (1985) Nucl. Acids, Res. 13, 3755–3772.
11. Summers, M.F., Byrd, R.A., Gallo, R.A., Samson, C.J., Zon, G. & Egan, W. (1985) Nucl. Acids Res. 13, 6375–6396.
12. Orbons, L.P.M., van der Marel, G.A., van Boom, J.H. & Altona, C. (1986) Nucl. Acids. Res. 14, 4187–4195.
13. Ikuta,S., Chattopadhyaya, R., Ito, H., Dickerson, R.E. & Kearns, D.R. (1986) Biochemistry 25, 4840–4849.
14. Wolk, S., Hardin, C.C., Germann, M.W., van de Sande, J.H. & Tinoco,I.Jr.(1988) Biochemistry 27, 6960–6967.
15. Benight,A.S., Wang, Y., Amaratunga, M., Chattopadhyaya, R., Henderson,J., Hanlon, S. & Ikuta, S. (1989) Biochemistry 28, 3323–3332.
16. Blummers, M.J.J., Walters, J.A.L.I., Haasnoot, C.A.G., Aelen, J.M.A., van Boom, J.H. & Hilbers, C.W. (1989) Biochemistry 28, 7491–7498.
17. Cheong,C., Varani,G.& Tinoco,I.Jr.(1990) Nature 346, 680–682.
18. Chattopadhyaya, R., Grzeskowiak,K. & Dickerson, R.E. (1988) Nature (London) 334, 175–179.
19. Baumann, U., Frank, R. & Blocker, H. (1986) Eur.J.Biochem. 161, 409–413.
20. Drew, H.R. (1984) J.Mol. Biol. 176, 535–557.
21. Xodo, L.E.,Manzini, G., Quadrifoglio, F., van der Marel, G.A. & van Boom,J.H.(1989) Biochimie 71, 793–803.
22. van Boom, J.H., van der Marel,G.A., Wille, G., Hoyng, C.F.(1982) in: Chemical and Enzymatic Synthesis of gene Fragments (H.Gassen and H.Lang, eds. 53–70.
23. Cantor, C.R & Schimmel, P.R. (1980) Biophysical Chemistry, Part II, W.H. Freeman and Company , San Francisco
24. Bevington,P.R.(1969) Data Reduction and Error Analysis for the Physical Sciences, McGraw-Hill, New York.
25. Klump, H.H. & Loffler, R. (1985) Biol.Chem. Hopper-Seyler 366, 345–353.
26. McConaughy, R.L., Laird, C.D & McCarthy, B.J. (1969) Biochemistry 8, 3289–3295.
27. Riazance,J.H., Baase,W.A., Johnson,C.,Hall,K., Cruz,P. & Tinoco,I.Jr.(1985) Nucl.AcidsRes. 13, 4983–89.
28. Rokita, S.E. & Romero-Fredes Lorraine (1989) Biochemistry 28, 9674–9679.
29. Senior, M.M., Jones, R.A. & Breslauer, K.J.(1988) Proc. Natl. Acad. USA 85, 6242–6246. Orbons et al Thesis
30. Amaratunga, M., Pancoska,P., Paner,T.M. & Benight,A.S. (1990) Nucl. Acids Res. 18, 577–582.
31. Paner, M.P., Amaratunga, M., Doktycz, M.J. & Benight, A.S. (1990) Biopolymers 29, 1715–1734.
32. Heus, H.A., van Kimmenade, J.M.A., van Knippenberg, P.H. & Hinz, H.-J. (1983) Nucl.Acids Res. 11, 203–210.
33. Groebe, D. & Uhlenbeck, O.(1989) Biochemistry 28, 742–747.
34. Tuerk, C., Gauss, P., Thermes, C., Groebe, D.R., Gayle, M., Guild, N., Stormo, G., D'Aubenton-Carafa, Y., Uhlenbeck, O.C.,Tinoco, I.Jr., Brody, E.N. & Gold, L. (1988) Proc. Natl. Acad. Sci. USA 85, 1364–1368.
35. Ikuta, S. & Wang, Y-S. (1989) Nucl. Acids Res. 17, 4131–4144.
36. Manzini,G., Xodo,L., Quadrifoglio,F., van Boom,J.H. & van der Marel,G.A.(1987) J.Biomol.Struct.Dyn. 4, 651–662.