Nucleic Acids Research, Vol. 18, No. 6



# Perturbation of DNA hairpins containing the EcoRI recognition site by hairpin loops of varying size and composition: physical (NMR and UV) and enzymatic (EcoRI) studies

Markus W.Germann*, Bernd W.Kalisch[1], Peter Lundberg, Hans J.Vogel and Johan H.van de Sande[1,*]

Departments of Biological Sciences and [1]Medical Biochemistry, The University of Calgary, Calgary, Alberta, T2N 4N1, Canada

Received December 5, 1989; Revised and Accepted February 20, 1990

## ABSTRACT

**We have investigated loop-induced structural perturbation of the stem structure in hairpins d(GAATTCX$_n$GAATTC) (X = A, T and n = 3, 4, 5 and 6) that contain an EcoRI restriction site in close proximity to the hairpin loop. Oligonucleotides containing either a $T_3$ or a $A_3$ loop were not hydrolyzed by the restriction enzyme and also showed only weak binding to EcoRI in the absence of the cofactor $Mg^{2+}$. In contrast, hairpins with larger loops are hydrolyzed by the enzyme at the scission site next to the loop although the substrate with a A4 loop is significantly more resistant than the oligonucleotide containing a T4 loop. The hairpin structures with 3 loop residues were found to be thermally most stable while larger hairpin loops resulted in structures with lower melting temperatures. The T-loop hairpins are thermally more stable than the hairpins containing the same number of A residues in the loop. As judged from proton NMR spectroscopy and the thermodynamic data, the base pair closest to the hairpin loop did form in all cases studied. The hairpin loops did, however, affect the conformation of the stem structure of the hairpins. From $^{31}P$ and $^{1}H$ NMR spectroscopy we conclude that the perturbation of the stem structure is stronger for smaller hairpin loops and that the extent of the perturbation is limited to 2-3 base pairs for hairpins with $T_3$ or $A_4$ loops. Our results demonstrate that hairpin loops modulate the conformation of the stem residues close to the loop and that this in turn reduces the substrate activity for DNA sequence specific proteins.**

## INTRODUCTION

DNA hairpin structures have been studied using biophysical and biochemical techniques (1-4). Interest in this class of alternate DNA structures was sparked by the observation that palindromic sequences, which can potentially form hairpin structures, are found near functional loci such as regulation or promoter sites, origin of replication and transcription termination sites. Under torsional stress, palindromic sequences can lower the superhelical free energy by extruding into a cruciform as has been amply demonstrated for a number of these sequences (1,5,6). These observations suggest that DNA hairpin structures may play an important role in gene expression or provide potential binding sites for drugs and proteins. The stability and structures of short DNA hairpins with various stem sequences have been investigated by biochemical, optical and NMR techniques, (3,4,7-18).

Based on systematic NMR studies of hairpin structures containing $T_n$ loops and on model building and energy minimization calculations, a loop building principle for DNA and RNA has been proposed. This principle indicates that an optimal loop size, as judged from the enthalpy, is reached at 4 to 5 residues for DNA or 6 to 7 residues for RNA (2,10,19,20). The essence of this model is that base stacking interactions within the loop and the loop stem interface are the determinants of the stability and size of a hairpin loop structure. Further investigations on these hairpin structures have indicated that the two loop thymines closest to the stem may form a T-T wobble pair (21,22). Other studies on hairpin structures however have provided evidence that hairpin loops can consist of as few of two or three nucleotides (11,15). In addition, loop structures which do not conform to the proposed loop building principle have been observed (4,23,24).

Most of the studies reported to date have been carried out on hairpin structures that contain T loops and consequently little is known about the influence of the loop composition. In a recent report, Senior et al., (25) investigated the effect of different four membered loops on the stability and structure of hairpins of the series d(CGAACGX$_4$CGTTCG) with X=A, C, G and T. The stem structure resembled B-DNA in all these hairpins, with pyrimidine hairpin loops being more stable than purine loops. We have initiated a similar study on the effect of hairpin loop size and composition on the hairpin d(GAATTCX$_n$CGAATTC)

* To whom correspondence should be addressed

(X=A, T and n=3, 4, 5 and 6). In particular, we investigated possible loop induced perturbations of the hairpin stem structure and whether these perturbations are of a short range nature. These questions were addressed by using DNA hairpin structures containing a EcoRI restriction site in close proximity to the hairpin loop. This enabled us to study the interaction of EcoRI with DNA hairpins of various loop sizes and composition and allowed us to correlate the biological activity with the thermodynamic properties and the structural data obtained from NMR studies.

## MATERIALS AND METHODS

### Synthesis and purification

Oligodeoxynucleotides were synthesized and purified as described previously (26). Molar extinction coefficients of the pure products were estimated from the extinction coefficients of the mononucleotides at 90°C in 5 M $NaClO_4$ (27).

### Gel electrophoresis

Analysis of native oligodeoxynucleotides was carried out on a 15% polyacrylamide gel run under non-denaturing conditions in TBM buffer (10 mM $MgCl_2$, 90 mM Tris-borate, pH 8.3). Bands were visualized by UV-shadowing. Products of the digestion of 5′ end labelled oligonucleotides by EcoRI were analyzed by denaturing gel electrophoresis on 15% polyacrylamide gels (8 M urea, 90 mM Tris-borate, 5 mM EDTA, pH 8.3). Kodak-X-Omat 5 film was used for autoradiography.

### End labeling

The 5′ hydroxyl ends of the oligonucleotides were labeled with a two fold excess of [$\gamma$-$^{32}$P] ATP and $T_4$ polynucleotide kinase in 10 mM $MgCl_2$, 5 mM DTT, 50 mM Tris, pH 7.6 (28). The reaction mixture was incubated at 37°C for 30 minutes. Reactions were terminated by heat inactivation of the enzyme (2 min at 90°C) and the end labeled oligonucleotides were separated from unincorporated ATP by Sephadex G-25 gel filtration.

### Restriction enzyme reactions

EcoRI was obtained from Bethesda Research Laboratories or Pharmacia. Oligonucleotides were digested with EcoRI (2–5 units/picomole substrate) under standard conditions of 100 mM NaCl, 10 mM $MgCl_2$, 100 $\mu$g/ml BSA, 1 mM $\beta$-mercapto-ethanol, 10 mM Tris, pH 7.5.

### Filter binding

The affinities of the restriction endonuclease EcoRI for DNA hairpin were determined at 20°C using 0.45 $\mu$m Millipore cellulose nitrate filters (HAWP 025 00). The filters were hydrated in buffer (10 mM Tris, pH 7.2, 50 mM NaCl, 10 mM EDTA). Immediately following the binding of the DNA oligonucleotide complex 2 ml of buffer was passed through each filter. The nitrocellulose filters were dried at 50°C and the bound oligonucleotide was quantitated by liquid scintillation counting in 5 ml Scintiverse II (Fisher Scientific). Relative binding constants for T3L, T4L, T5L, A3L, A4L and A6L hairpins were obtained from competition experiments in which 5′ phosphorylated competitor and 5′ $^{32}$P end labeled T4L were incubated with a limiting amount of EcoRI. The ratios of the binding constants were calculated using equation [1].

$$\frac{K_2}{K_1} = \frac{(cpm° - cpm)}{cpm} \cdot \frac{[Hp_1]}{[Hp_2]} \qquad [1]$$

where $K_1/K_2$ is the ratio of the binding constant, cpm and cpm° indicates the radioactivity bound in presence and absence of the competitor respectively. The strand concentration of the competitor whose affinity is to be determined is given by $Hp_2$, while that of T4L is $Hp_1$.

### Absorption spectroscopy

Ultraviolet absorption spectra and thermal denaturation profiles were recorded with a Varian 2280 spectrophotometer equipped with temperature regulated cuvette holders.

### Thermal denaturation analysis

Thermally induced helix-coil transitions were monitored at 268 nm, using a temperature gradient of 0.5 °C/min. Absorbance readings were collected at 0.5°C intervals and corrected for volume expansion. A concerted two state model in which the absorbances of helix and coil form were allowed to vary linearly with the temperature was employed to analyze the data (29). The measured absorbance was fitted to [2] using a non linear least squares Simplex procedure:

$$A_T = A_H(1-\Theta) + A_C\Theta \qquad [2]$$

where $A_T$ is the absorbance of the oligonucleotide as a function of the temperature and $A_H$ and $A_C$ are the absorbances of helix and coil forms respectively. The degree of transition $\Theta$ is given by $1/(1+K)$ where the helix coil equilibrium constant (K) at temperature T was obtained from the van't Hoff equation using the van't Hoff enthalpy $\Delta H°vH$ and the the melting temperature $T_M$ ($\Theta = 0.5$) Entropies are calculated from $\Delta S° = \Delta H°vH/T_M$, where R is the gas constant, and $T_M$ is the melting temperature (Kelvin).

### Nuclear magnetic resonance spectroscopy

All measurements were done on a Bruker AM-400 NMR spectrometer equipped with a wide bore 9.4 Tesla magnet and an Aspect 3000 computer.

Proton decoupled $^{31}$P NMR data were recorded using a 10 mm broad band probe tuned to phosphorus (162 MHz). A bi-level composite pulse scheme (WALTZ-16) was employed for proton decoupling (30). $^{31}$P NMR samples were typically made up of 0.2 mM strands in 1.7 ml water (10% $D_2O$) containing 50 mM Tris-hydroxymethyl-aminomethane pH 7.0, 100 mM NaCl, 0.1 mM EDTA. The $^{31}$P NMR acquisition parameters were as follows: pulse length 9 $\mu$s (50°), spectral width 6000 Hz, acquisition time 1.4 s, relaxation delay 3.0 s. Approximately 1000 scans were collected in 8 K word data sets. Chemical shifts were referenced to 85% phosphoric acid as an external standard at 25°C and were given a negative sign when upfield from this standard.

Imino proton spectra were recorded using the jump and return pulse sequence for water suppression (31). The parameters were as follows: the carrier was set at the water resonance, the spectral width was 10000 Hz and the acquisition time was 0.81 s. The delay between the two pulses in the pulse train was 100 $\mu$s. Data were collected in 16 K word data sets. Samples containing 40–90 $A_{260nm}$ units (ca. 0.3–1 mM in strands) were dissolved in 0.3 ml buffer (100 mM NaCl, 10 mM sodium phosphate, 0.1 mM EDTA, pH 5.7 containing 10% $D_2O$).

Nuclear Overhauser effects between imino protons were

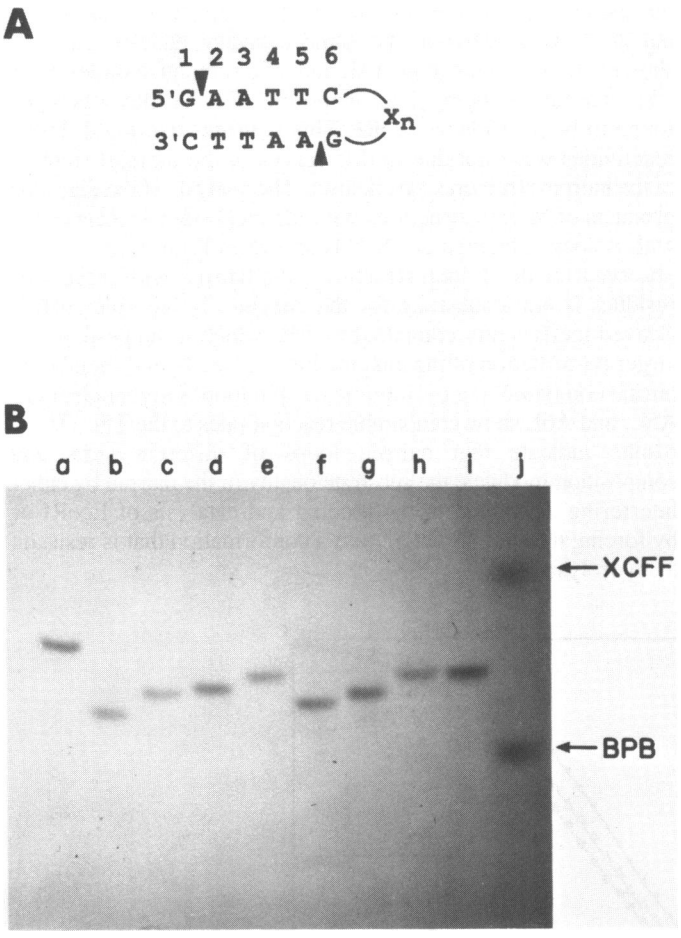


**Figure 1.** (A) DNA hairpin substrates investigated in this study. The hairpin consists of either T or A residues with n= 3, 4, 5 and 6 which are referred to as T3L, T4L, T5L, A3L, A4L, A5L and A6L. The base pairs are labeled 1 through 6 starting at the free end. (B) Gel electrophoretic analysis of hairpin oligonucleotides run under native conditions. Oligonucleotides (0.2 $A_{260nm}$) were applied in 10 $\mu$l running buffer containing 10% sucrose. Lane a, d(CGCGCGCGCG); Lane b, A3L; Lane c, A4L; Lane d, A5L; Lane e, A6L; Lane f, T3L; Lane g, T4L; Lane h, T5L; Lane i, d(GTACGTAC). BPB, bromophenol blue; XCFF, xylene cyanol.

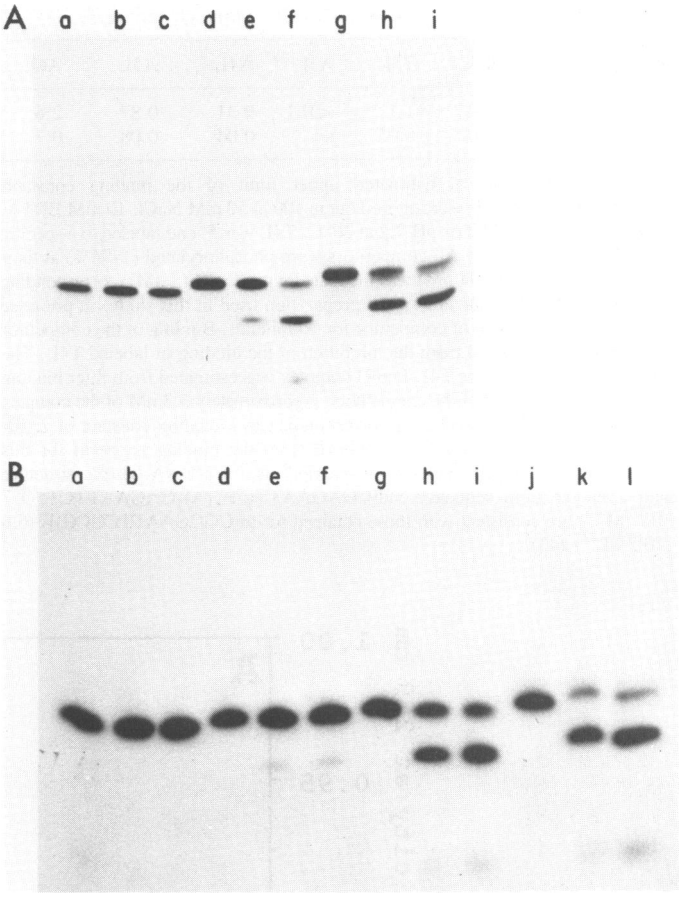


**Figure 2.** Denaturing gel electrophoretic analysis of the fragments produced by the hydrolysis of the DNA hairpins dGAATTC$X_n$GAATTC with EcoRI. DNA-hairpins (2 pmoles, 40000 cpm) were incubated with 10 units of EcoRI at 15°C for the times indicated below. (A) T-loop hairpins. Lane a, 5′ end labeled T3L. Lane b, T3L + EcoRI, 30 minutes. Lane c, T3L + EcoRI, 60 minutes, Lane d, 5′ end labeled T4L. Lane e, T4L + EcoRI, 30 minutes. Lane f, T4L + EcoRI, 60 minutes. Lane g, 5′ end labeled T5L. Lane h, T5L + EcoRI, 30 minutes. Lane i, T5L + EcoRI, 60 minutes. (B) A-loop hairpins. Lane a, 5′ end labeled A3L. Lane b, A3L + EcoRI, 30 minutes. Lane c, A3L + EcoRI, 60 minutes. Lane d, 5′ end labeled A4L. Lane e, A4L + EcoRI, 30 minutes. Lane f, A4L + EcoRI, 60 minutes. Lane g, 5′ end labeled A5L. Lane h, A5L + EcoRI, 30 minutes. Lane i, A5L + EcoRI, 60 minutes. Lane j, 5′ end labeled A6L. Lane k, A6L + EcoRI, 30 minutes. Lane l, A6L + EcoRI, 60 minutes.

measured at −3 and 15°C respectively by collecting sets of 8 or 16 scans with the decoupler alternately on and off resonance. Typically, a total of 800 scans were collected for each individual experiment. Imino protons were irradiated for 0.25−0.5 s with sufficient power to obtain 80% saturation. The relaxation delay between the acquisitions was 4s, and the acquisition time was 0.41 s. A 5 Hz exponential line broadening function was applied to the FID prior to Fourier transformation to improve the signal noise ratio of the spectra. $^1$H chemical shifts are referenced to internal DSS.

## RESULTS

### Structures

The general sequence of the oligonucleotides used in this study is shown schematically in Figure 1A. The oligonucleotides were analyzed by gel electrophoresis under denaturing and native conditions. Under native conditions (Figure 1B), the oligonucleotides migrated as single bands with electrophoretic mobilities that correspond to a duplex with 6 to 8 base pairs, demonstrating that in all cases only intramolecular hairpin structures were formed. Each hairpin contains 2 G-A phosphodiester bonds ( ), that can potentially be hydrolyzed by EcoRI.

### Dependence of the EcoRI catalyzed hydrolysis on the loop size and composition

The product distribution of the EcoRI catalyzed hydrolysis of the 5′ end labelled hairpins was analyzed by denaturing polyacrylamide gel electrophoresis (Figure 2). Strand cleavage at the G-A phosphodiester bond near the 5′ terminus results in the release of pG, while for strand cleavage near the hairpin loop larger molecular weight products are expected, allowing for a ready distinction of the preferred cutting site. The hairpin with 3 T loop residues is a poor substrate for EcoRI at 15°C (Figure 2A). Digestion of T4L and T5L hairpins results in the appearance of a 11-mer and 12-mer respectively, while very little pG is produced demonstrating preferential cleavage of the G-A

**Table I:** Relative binding constants for EcoRI-hairpin complexes.

| Substrate | T3L | T4L | T5L | A3L [2] | A4L | A5L | A6L |
|---|---|---|---|---|---|---|---|
| $K_2/K_1$ | 0.14 | 1.0 | 1.1 | <0.1 | 0.31 | 0.87 | 2.6 |
| St. dev [1] | 0.06 | 0.02 | 0.05 | nd | 0.04 | 0.08 | 0.7 |

[1] Standard deviation [2] Estimated upper limit of the binding constant. Competition experiments were carried out in 100 μl 50 mM NaCl, 10 mM EDTA, 50 μg/ml BSA, 10 mM Tris pH 7.2 at 20°C. T4L was 5′ end labeled to a specific activity of 2.3 Ci/mmol. All competitors were phosphorylated (>96%) at their 5′ ends. Labeled T4L (94 nM) was incubated with EcoRI (5 nM) (corresponding to 10.9 μl of the 10 unit/ μl enzyme preparation used in this study) in presence of an equimolar amount of competitor for 30 minutes. Binding of the competitor to EcoRI was monitored from the inhibition of the binding of labeled T4L. The association constant of the T4L-EcoRI complex was estimated from filter binding experiments in absence of the competitors. Approximately 3.3 nM of the complex is formed under these conditions corresponding to a binding constant of $2 \cdot 10^7$ $M^{-1}$. A similar value was also obtained from the binding isotherm for this hairpin. Control binding experiments carried out for a DNA hairpin structure with a longer stem sequence 5′d(CGAGAATTCTC(A)$_5$GAGAATTCT ($1.7 \cdot 10^{10}$ $M^{-1}$) are consistent with those obtained for [d(CGCGAATTCGCG)]$_2$ ($6.6 \cdot 10^{10}$ $M^{-1}$) (43).

phosphodiester bond next to the hairpin loop. Reactions carried out at 37°C resulted in the same cleavage pattern (data not shown). In particular, both T4L and T5L oligonucleotides show a similar rate of hydrolysis, while the T3L hairpin was again found to be resistant to EcoRI. This indicates that the different reactivities were not due to differences in the thermal stability of the hairpin structures (see below). The analysis of the digestion products of hairpin structures with identical stem sequences but with A loops obtained at 15°C is shown in Figure 2B. As was observed for the T loop structures, the hairpin with three loop residues is not a substrate for the enzyme. In addition A4L is cleaved inefficiently compared to T4L which is suggestive of a larger perturbation of the enzyme binding site by a 4 membered purine compared to a pyrimidine hairpin loop. Oligonucleotides A5L, and A6L show comparable reaction rates to the T5L. These results indicate that hairpin loops of different sizes and composition modulate the substrate quality of the hairpin by either interfering directly with the binding and catalysis of EcoRI or by forcing the stem structure into a conformation that is resistant to hydrolysis.

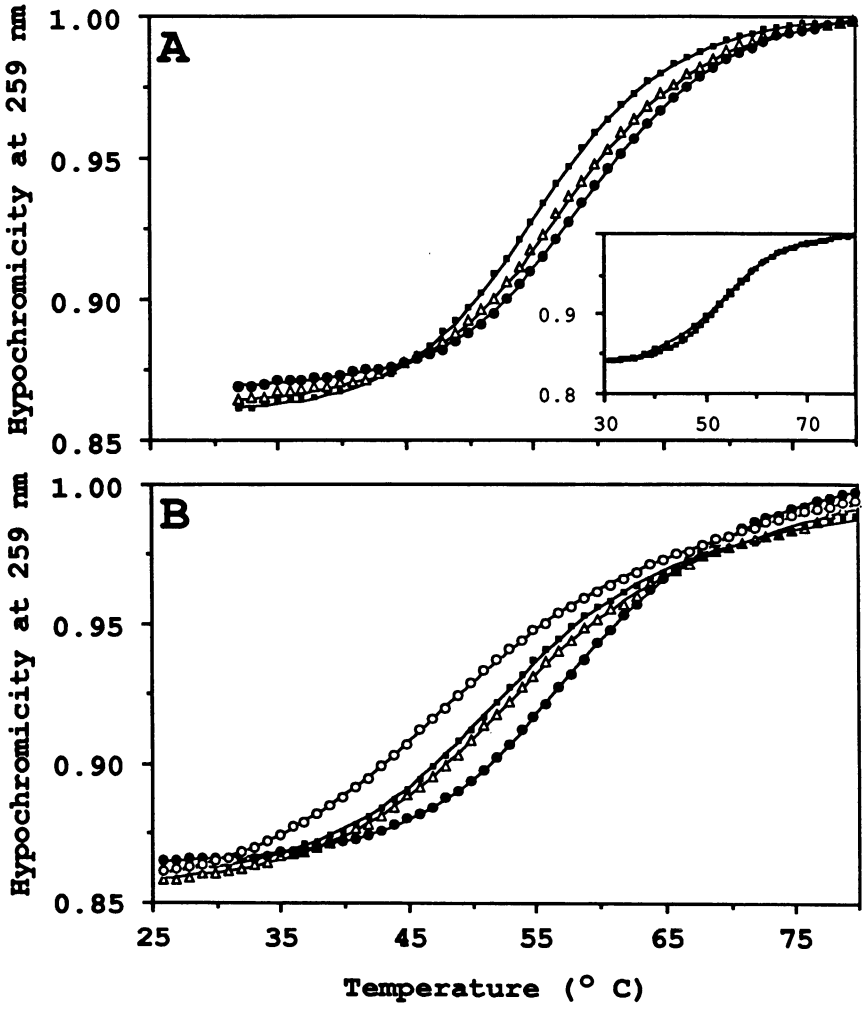


**Figure 3.** Ultraviolet melting profiles of d(GAATTCX$_n$GAATTC), recorded at 259 nm in 200 mM NaCl, 0.1 mM EDTA, 50 mM sodium phosphate pH 7.0. The oligonucleotide strand concentration was approximately 3.3 μM. (A) T3L (●), T4L (△), T5L (■). Inset: Concentration independence of the melting of T3L 3.3 μM (□) and 79 μM (■)in EcoRI buffer (100 mM NaCl, 10 mM $MgCl_2$,100 μg/ml BSA, 10 mM Tris pH 7.2) monitored at 261 nm. The melting temperatures were 53.6°C and 53.2°C for the low and high concentration sample respectively. Enthalpies were estimated from a two state model which assumes a flat lower and upper base line and are 153 $kJmol^{-1}$ and 156 $kJmol^{-1}$ respectively. (B) Buffer as in A, A3L (●), A4L (△), A5L (■), A6L (○). The solid lines represent the best fit obtained from nonlinear least squares curve fitting to a two state model (see Material and Methods)

**Table II:** Thermodynamic data for the helix coil transition of d(GAATTCX$_n$GAATTC).

| | X = T | | | X = A | | | |
|---|---|---|---|---|---|---|---|
| n | 3 | 4 | 5 | 3 | 4 | 5 | 6 |
| $T_M$ | 58.4 | 56.6 | 55.3 | 56.6 | 51.4 | 50.8 | 46.2 |
| $\Delta H°vH$ | 166 | 161 | 168 | 148 | 137 | 135 | 117 |
| $\Delta S°$ | 0.50 | 0.49 | 0.51 | 0.45 | 0.42 | 0.42 | 0.37 |

Thermodynamic data are the average of several independent measurements and were determined in 200 mM NaCl, 0.1 mM EDTA, 50 mM sodium phosphate pH 7.0. The oligonucleotide strand concentration was approximately 6 $\mu$M. The van't Hoff enthalpy and entropy were obtained from a two state analysis of the melting profile and are given in kJmol$^{-1}$ and kJmol$^{-1}$K$^{-1}$ respectively. The standard deviations are 0.4–0.8 °C, 2.3–9.0 kJmol$^{-1}$ and 0.02–0.04 kJmol$^{-1}$K$^{-1}$ for $T_M$, enthalpies and entropies respectively.

## Binding of EcoRI to DNA hairpin structures

The observation that the T3L and the A3L hairpins are cleaved very inefficiently by EcoRI could indicate that in these cases the presence of the loop structure affects the binding of the substrate to the enzyme. In order to investigate this possibility we determined the affinity of EcoRI for the hairpin substrates by nitrocellulose filter binding in absence of $Mg^{2+}$. Initial filter binding experiments showed that the binding constants are relatively low for the hairpin structures, preventing us from recording complete binding isotherms for all hairpin structures. Based on the filter binding experiments the binding constant of T4L was estimated to be $2 \cdot 10^7$ M$^{-1}$. Extending the incubation time to 60 minutes did not produce an increase in binding, demonstrating that equilibrium was reached under our reaction conditions. The affinity constant for T4L is in agreement with that reported for the hexamer [d(GAATTC)]$_2$ ($8 \cdot 10^6$ M$^{-1}$) which corresponds to the stem sequence of the hairpins studied here (32). Binding constants for the other hairpins were obtained from competition experiments in which 5′ phosphorylated (unlabeled) competitor and 5′ $^{32}$P end labeled T4L were incubated with a limiting amount of EcoRI. T3L is the least effective inhibitor in the T-loop series with an association constant that is 10 fold lower than for T4L, while T5L shows a similar affinity as T4L (Table I). In the A-loop series a similar trend is observed in that increasing the loop size results in a stronger binding to EcoRI.

## Thermodynamic properties of the EcoRI hairpin structures

The resistance of hairpin structures with 3 loop residues to cleavage by EcoRI could be due to the small hairpin loop which does not allow the formation of the adjacent base pair 6. This would result in a structure that is thermally significantly less stable and also has a lower enthalpy than hairpins with larger loops (33). In order to investigate this possibility we have recorded UV denaturation profiles of the T- and A-loop hairpins (Figure 3). All helix coil transitions were fully reversible, appeared monophasic and, as demonstrated for d(GAATTCT$_3$GAATTC), were independent of the strand concentration (Figure 3A, inset). This behavior is consistent with intramolecularly formed hairpins, but not with dimeric structures. For the relatively small enthalpies measured here, the melting temperature at a 24 fold higher strand concentration would be shifted by more than 10°C if the melting occurred directly from a dimeric structure. From the thermodynamic data, based on a two state model, it follows that hairpins containing T loops are thermally and enthalpically more stable than corresponding A-loop hairpins (Table II). The melting temperature for both T and A loop hairpins increases upon decreasing the loop size; the most stable hairpin being the one with 3 loop residues in either hairpin series. The enthalpy for the helix-coil transition is constant for the T-loop series and decreases for the A-loop series upon enlarging the loop size. Thus, our results demonstrate that for the hairpins investigated here, a small loop does not lead to a destabilized structure. The observation that the enthalpy is constant for the T-loop series strongly argues against the hypothesis that the innermost CG base pair is disrupted, since that should result in a lower enthalpy (10). In case of the A-loop hairpins the enthalpy is consistently lower than for the corresponding T-loop hairpins which could indicate that the base pair next to the loop did not form. However, increasing the number of A residues in the loop did not result in enthalpically more stable structures as might be expected if this base pair is disrupted for small A- loops In contrast, we observe that the enthalpy of the A loop hairpins decreases markedly for larger loop sizes which could indicate that an increase of the number of A residues may prevent the hairpin loop from forming a compact and stable structure.

## NMR properties of EcoRI hairpin substrates

In order to explain the different substrate behavior of the hairpins on a structural basis we have selected the EcoRI resistant T3L and the EcoRI substrates T4L and A4L for further study by proton and phosphorus NMR in the absence of $Mg^{2+}$. The extent of base pairing was investigated by recording NMR spectra in $H_2O$. Imino protons corresponding to the T hairpin loop resonate at 10.5–11.5 ppm, while imino protons of the CG and AT base pairs resonate at 12.5–13 ppm and 13.6–14.3 ppm respectively (Figure 4 A,B,C). The observation of two CG resonances for all oligonucleotides demonstrates that the innermost CG base pair forms in all three oligonucleotides, contrary to what was suggested by the thermodynamic data for the A4L hairpin. The spectra recorded at 5°C demonstrate that there are differences between T3L, T4L and A4L oligonucleotides. In the T3L hairpin (Figure 4A) all 6 imino protons are resolved while for the T4L (Figure 4B) and A4L hairpin (Figure 4C) AT resonances overlap. A preliminary assignment of the resonances was obtained by recording the imino proton spectra at increasing temperature. The intensity of the imino proton resonances at the open end will be reduced because of faster exchange with the solvent and can thus be identified. Upon increasing the temperature the T imino protons of the hairpin loop and the resonances from base pairs 1 and 2 decrease markedly while the others remain relatively unaffected. The remaining resonances cannot be assigned using this method because further increases in temperature result in a simultaneous loss of the remaining resonances. Spectra recorded at pH 7 are consistent with our partial assignment since the same protons that were found to be most temperature sensitive were also those most broadened by the higher pH. The partial assignment based on

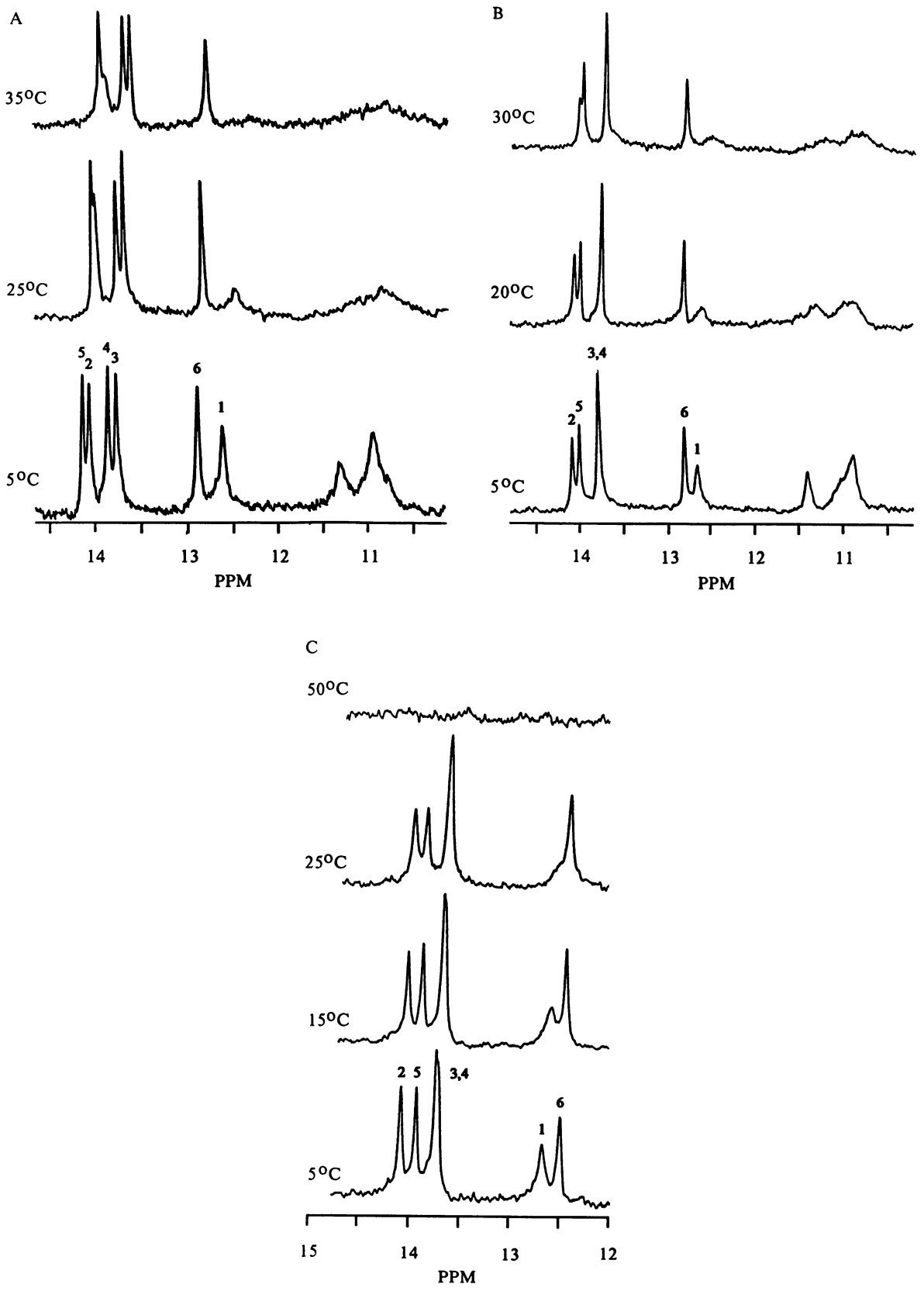


**Figure 4.** Temperature dependence of the 400 MHz $^1$H imino proton spectra of d(GAATTCX$_n$GAATTC). (A) T3L, (B) T4L and (C) A4L in 200 mM NaCl, 0.1 mM EDTA, 50 mM phosphate, pH 5.7 in $H_2O$.

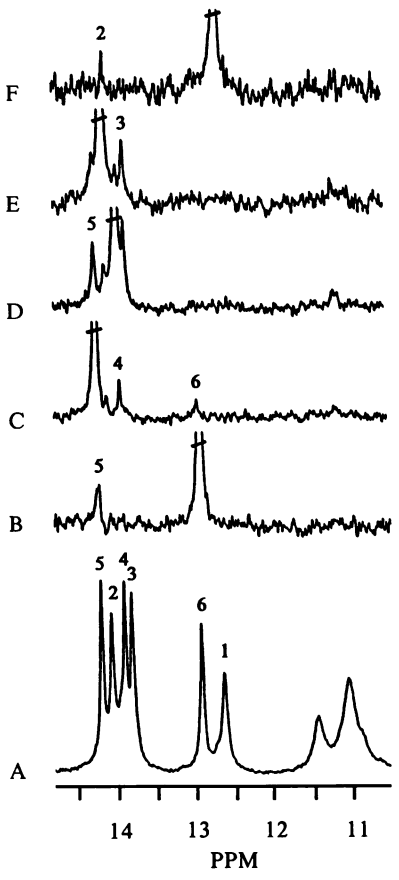


**Figure 5.** Assignment on the imino protons of d(GAATTCT$_3$GAATTC) by one dimensional NOE. (A) Imino proton spectrum of d(GAATTCT$_3$GAATTC) in 200 mM NaCl, 0.1 mM EDTA, 50 mM phosphate pH 5.7 at 5°C together with the NOE difference spectra obtained on irradiation of (B) GC imino proton resonance at 12.94 ppm (6). (C) AT imino proton resonance at 14.22 ppm (5). (D) AT imino proton resonance at 13.91 ppm (4). (E) AT imino proton resonance at 14.08 ppm (2). (F) GC imino proton resonance at 12.66 ppm (1).

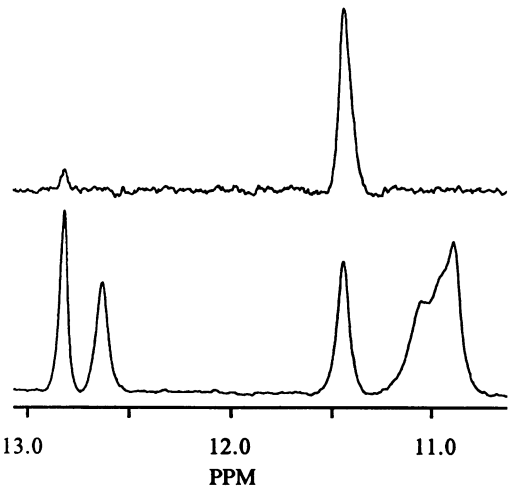


**Figure 6.** Assignment base pair 1 of d(GAATTCT$_4$GAATTC) by one dimensional NOE. Top: Imino proton spectrum of d(GAATTCT$_4$GAATTC) in 200 mM NaCl, 0.1 mM EDTA, 50 mM phosphate pH 5.7 at −3°C to slow down the exchange of the loop T imino protons. Bottom: NOE difference spectra obtained on irradiation of the T-loop imino proton at 11.45 ppm.

**Table III:** Imino proton chemical shifts of d(GAATTCX$_n$GAATTC).

| Imino proton | $T_3$-loop | $T_4$-loop | $A_4$-loop | Dimer [1] |
|---|---|---|---|---|
| 1 | 12.66 | 12.64 | 12.65 | 12.75 |
| 2 | 14.08 | 14.08 | 14.06 | 13.97 |
| 3 | 13.82 | (13.81) | 13.70 | 13.86 |
| 4 | 13.91 | (13.81) | 13.73 | – |
| 5 | 14.22 | 14.00 | 13.94 | – |
| 6 | 12.94 | 12.73 | 12.47 | – |
| loop−T a | 11.45 | 11.44 | – | – |
| loop−T b | ~11.0 | ~11.1 | – | – |
| loop−T c | ~11.0 | ~11.0 | – | – |
| loop−T d | – | ~10.9 | – | – |

Chemical shifts were measured at 5°C in 200 mM NaCl, 0.1 mM EDTA, 50 mM phosphate at pH 5.7. Resonances in brackets could not be assigned unambiguously due to overlap. [1] Imino proton chemical shift of CGCG(1)A(2)A(3)TTCGCG, 2.5 mM EDTA, 0.1 M phosphate, pH 7.5 at −5°C, (44).

the thermolability was confirmed and extended by measuring the nuclear Overhauser enhancement between the imino protons, here shown for T3L (Figure 5). In this experiment, the imino proton of one base pair is irradiated and changes in the intensity of the other imino protons indicate that they are close in space. For both the T3L and T4L hairpin we observed a NOE from the T-loop imino protons to one of the CG base pairs which consequently was assigned as the innermost CG base pair 6 (shown for the single downfield shifted T-loop imino proton of T4L, Figure 6). By elimination, the remaining CG base pair in each structure was assigned as 1. Irradiating base pair 6 produced an NOE connectivity to a single AT base pair 5, but did not show a NOE back to the T-loop imino protons. This is due to the faster hydrogen exchange of these latter residues which is also manifest in their greater pH sensitivity. Irradiation of the AT 5 resonance produced NOE's to CG 6 and to AT 4. The latter showed strong NOE's back to AT 5 and to a central AT base pair 3. The only remaining AT imino proton resonance which, by elimination must be 2, showed the expected NOE to CG residue 1. The assignment of imino proton 3 was also confirmed from its NOE to AT base pair 2. The smaller peaks flanking the irradiated resonance 2 are due to nonselective saturation and do not signify that these protons are close in space. In case of the T4L oligonucleotide, the imino protons of the central two AT base pairs have the same chemical shift and therefore could not be assigned.

The assignment for the imino protons of the A4L hairpin could not be obtained using the hairpin loop to identify the base pair 6. In this case the CG base pair at 12.65 ppm was chosen as the starting point for the assignment. This base pair shows both the thermal and pH sensitivity expected for a terminal base pair and consequently was assigned as 1. This is also supported by the observation that its chemical shift (12.65 ppm) is very similar to that of base pair 1 in T3L (12.66 ppm) and T4L (12.64 ppm). The remaining imino proton resonances in this oligonucleotide were assigned based on the nuclear Overhauser enhancement between each other and base pair 1. Inspection of Table III shows that the chemical shift of the imino protons 1 and 2 are similar for all three oligonucleotides. This indicates that the loop effects are limited to three to four base pairs. The effect of the different hairpin loops is felt strongly by the innermost CG base pair which results in a upfield shift of 0.21 ppm for T4L compared to T3L. The same base pair in the A4L hairpin is downfield shifted by 0.5 ppm.

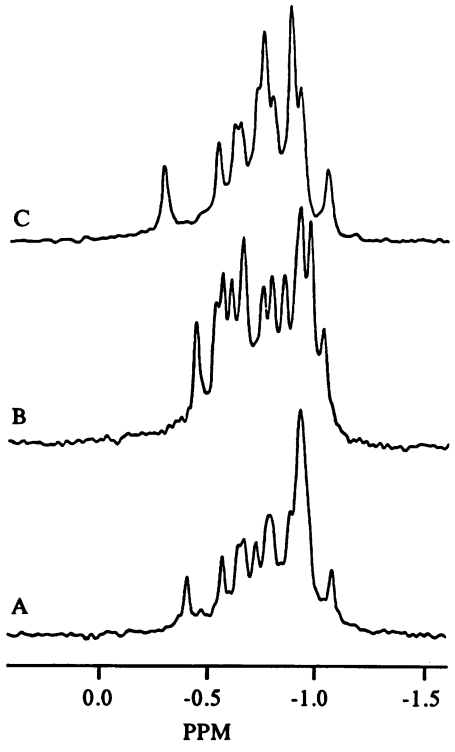


**Figure 7.** Proton decoupled $^{31}P$ NMR spectra (162 MHz) of d(GAATTCX$_n$GAATTC) of (A) T3L, (B) T4L and (C) A4L in 200 mM NaCl, 0.1 mM EDTA, 10 mM Tris, pH 7.0, recorded at 25°C.

Figure 7 shows the phosphorus NMR spectra of the three oligonucleotides under native conditions. The resonances of the phosphodiester backbone of all oligonucleotides are centered around −0.8 ppm, which is characteristic for right handed oligo and polynucleotides (34). Resonances that are resolved from the main group are likely due to phosphates from the loop region, since $^{31}P$ chemical shifts are known to be sensitive to the backbone torsion angles (35). The spectra obtained for the oligonucleotides differ significantly. In the case of the T3L and T4L hairpin we find that approximately 8 of the 15 phosphodiester bonds have a similar chemical shift while a comparison of A4L with either T3L or T4L shows approximately 6 matches. Thus, this qualitative analysis indicates that the phosphodiester backbones in the stem part of these hairpins differ, since the differences observed cannot be simply accounted for by the different phosphodiester linkages in the loop structures.

## DISCUSSION

The existence of hairpin conformation of the oligonucleotides was confirmed by both gel electrophoresis under native conditions and UV melting experiments. The oligonucleotide T3L was not hydrolyzed by EcoRI, in contrast to T4L and T5L which showed similar reactivity. Similarly, in the A loop series, A3L was not a substrate for the enzyme. For the hairpins that are hydrolyzed by EcoRI, strand scission occurs preferentially at the GpA next to the hairpin loop, thereby demonstrating asymmetric cleavage of the enzyme recognition site. Asymmetric cleavage is also observed for EcoRI restriction sites that are flanked on one side by AT and by GC base pairs on the other side. Preferential hydrolysis occurs at the GpA phosphodiester bond adjacent to the AT rich flanking region (36). EcoRI recognition sites in which either a CT or CA mismatch or an AT base pair is substituted for one of the CG base pairs are cleaved predominantly at the nonperturbed cleavage site (37). The EcoRI enzyme is known to make two ionic contacts per subunit outside the recognition sequence which cannot be made at the 5′ end in the hairpins used in this study. This provides also a rational for the lower affinity of these hairpins for EcoRI compared to longer oligonucleotides (vide infra). In addition, since the oligonucleotides are small, end effects have to be considered (38,39). The combination of these effects renders the 5′ site relatively resistant to EcoRI and cleavage first occurs near the hairpin loop. Once cleaved, the resulting bimolecular structure has little stability under the reaction conditions and dissociates readily, thereby preventing any further reaction at the 5′ site.

Filter binding experiments show that the enzyme binds T4L and T5L with similar affinity but shows significantly reduced binding to T3L. For the A loop series the A3L binds weakly to the enzyme, while the A4L, A5L, and A6L show increasingly stronger interactions. Thus, the reactivity observed with the endonuclease EcoRI correlates with the strength of interaction between the DNA and the enzyme. The hairpin-EcoRI binding constant can be influenced by several factors. 1) The constrained and rigid loops (see below) may physically interfere with the binding of the enzyme. 2) A hairpin loop modulates the conformation of the adjacent base pairs resulting in a decrease of the affinity. 3) Hairpins containing larger loops are also more negatively charged which may aid the binding. The latter effect is probably small since both T4L and T5L have similar affinity. The observation that A4L, A5L, and A6L bind the enzyme with comparable affinity as T4L indicates that in these cases the A-loops do not directly interfere with the binding event.

The finding that both T3L and A3L are resistant to cleavage and also show weak interaction with the enzyme could indicate that the CG base pair adjacent to the hairpin loop did not form. However, as illustrated by thermodynamic data, the hairpins with the smallest loop structures are thermally and enthalpically most stable, suggesting that this base pair formed. Confirmation for this was obtained from the imino proton spectra of these hairpins which provided direct evidence for base pair 6 in T3L, T4L, A3L (data not shown) and A4L. This CG base pair did not have reduced stability, because upon increasing the temperature melting occurred exclusively from the free end. Thus, the hairpin loop did not provide an extra melting site, in contrast to earlier observations on other T-loop hairpins (9). In addition, the thermodynamic data of the hairpins investigated here differ from those reported previously, Haasnoot et al. (10), suggesting that the loop structures of these hairpin series differ. Both T3L and T4L show similar T-loop imino proton spectra and in both cases the loop imino protons have marginal thermal stability which is not compatible with a potential T-T wobble pair formed on top of the stem helix. Furthermore, for a T-T wobble pair the distance between the imino protons is short thereby giving rise to a strong NOE (21). Yet we could not observe an NOE between the downfield shifted loop-T imino proton which is a neighbor of the CG base pair 6 and the other loop T imino protons at (10.9−11.1 ppm). The fact that the hairpin loops in our study were closed by a CG as opposed to a less stable AT base pair may be the reason why loops with 3 residues showed the observed high thermal stability and also result in different loop structures due to the different stacking interactions (CG versus an AT base

pair) at the loop stem interface. It is noteworthy that the both the 2 and 3 nucleotide loop hairpin structures studied by Orbons (15) and Summers et al. (11) respectively also contain a CG base pair next to the loop.

Our results demonstrate that small hairpin loops perturb the base pairs in its vicinity. However, the chemical shifts of the imino protons 1, 2 (and 3 for the T-loop hairpins) are similar and thus the perturbation caused by the loop structure is not of a long range nature but is limited to the 3 base pairs nearest the loop. The observed chemical shift differences for the imino protons that are not adjacent to the hairpin loop are too large to be accounted for by ring current shifts of the loops and suggest that the hairpin loops studied here modulate the stem structure of the hairpins (40). This conclusion is also supported from the $^{31}P$ NMR spectra which suggest that the perturbation extends 2−3 base pairs into the stem region.

Filterbinding, UV-melting and NMR spectroscopic studies were carried out in the absence of $Mg^{2+}$, which is required for catalytic activity of the enzyme. Potential effects of this counter-ion on the hairpin structure have therefore to be considered. Circular dichroism as well as UV-melting measurements, performed in the presence or absence of the cofactor indicate that $Mg^{2+}$ does not appreciably affect the hairpin structure. The crystalline complex of EcoRI and the tridecamer 5'd(TCGCGAATTCGCG) is catalytically competent in that if the cofactor is diffused into the cocrystal, catalysis occurs, thereby illustrating the importance of structural information obtained in the absence of $Mg^{2+}$ (41).

The thermodynamic data for T4L and A4L are consistent with those reported for hairpins with 4 nucleotide A and T-loops that contain a GC base pair next to the loop (25). In conclusion, our results firstly indicate that hairpin loops can modulate the conformation of the stem structure as evidenced by spectroscopic and enzymatic methods and secondly that the base pair next to the hairpin loop strongly influences hairpin loop formation. Both $^{31}P$ and $^{1}H$ NMR indicate that the perturbation caused by hairpin loops is limited to approximately 3 base pairs, but does not result in a drastically altered stem structure since additional NMR data (42) indicate that the stem structure of T3L (and T4L) is still in a conformation that is compatible with B-DNA parameters (see also Senior et al.(25)). Our data suggests that the reduced binding of EcoRI to T3L and also its resistance to hydrolysis derives from the loop induced alteration of the stem structure.

## ACKNOWLEDGMENTS

This work was supported by the Medical Research Council of Canada and the Alberta Heritage Foundation for Medical Research.

## ABBREVIATIONS

NMR, nuclear magnetic resonance; Tris, tris-hydroxymethyl-aminomethane; EDTA, ethylenediamine tetraacetic acid; UV, ultraviolet; DTT, dithiothreitol; BSA, bovine serum albumin; FID, free induction decay; DSS, 2,2-dimethylsila-pentane-5-sulfonate; NOE, nuclear Overhauser effect.

## REFERENCES

1. Lilley, D. M. J. (1981), Nucleic Acid Res. 9, 1271−1289.
2. Hilbers, C. W., Haasnoot, C. A. G., de Bruin, Joordens, J. J. M., van der Marel, G. A. & van Boom, J. A. (1985), Biochimie 67, 685−695.
3. Ikuta, S., Chattopadhyaya, R., Ito, H., Dickerson, R. E. & Kearns, D. R., (1986), Biochemistry 25, 4840−4849.
4. Hare, D. R. & Reid, B. R. (1986), Biochemistry 25, 5341−5350.
5. Panayotatos, N. & Wells, R. D. (1981), Nature 289, 466−470.
6. Mizuuchi, K., Mizuuchi, M. & Gellert, M. (1982), J. Mol. Biol. 156, 229−243.
7. Scheffler, I. E., Elliot, E. L. & Baldwin R. L. (1968), J. Mol. Biol. 36, 291−304.
8. Haasnoot, C. A. G., den Hartog, J. H. J., de Rooij, J. F. M., van Boom, J. H. & Altona, C. (1979), Nature 281, 235−236.
9. Haasnoot, C. A. G., den Hartog, J. H. J., de Rooij, J. F. M., van Boom, J. H. & Altona, C. (1980), Nucleic Acids Res. 8, 169−181.
10. 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), Journal of Biomolecular Structure and Dynamics 3, 843−857.
11. Summers, M. F., Byrd, R. A., Gallo, K. A., Samsom, C. J., Zon, G. & Egan, W. (1985), Nucleic Acids Res. 13, 6375−6386.
12. Germann, M. W., Schoenwaelder, K.-H. & van de Sande, J. H. (1985), Biochemistry 24, 5698−5702.
13. Baumann, U., Frank, R. & Blöcker, H. (1986), Eur. J. Biochem. 161, 409−413.
14. Roy, S., Weinstein, S., Borah, B. Nichol, J., Appella, E., Susman, J. C., Miller, M., Shindo, H., & Cohen, J. S. (1986), Biochemistry 25, 7417−7423.
15. Orbons, L. P. M. (1987), PhD. Thesis University of Leiden.
16. Gupta, G., Sarma, M., Sarma, M. H., Bald, R., Engelke, U., Uei, S. L., Gessner, R. & Erdman V. A. (1987), Biochemistry 26, 7715−7723.
17. Rinkel, L. J., van der Marel, G. A., van Boom, J. H. & Altona, C. (1987), Eur. J. Biochem. 163, 287−296.
18. Pramanik, P., Kanhouwa, N. & Kan, L.-S. (1988), Biochemistry 27, 3024−3031.
19. Haasnoot, C. A. G., de Bruin, S. H., Berendsen, R. G., Janssen, H. G. J. M., Binnendijk, T. J. J., Hilbers, C. W., van der Marel, G. A. & van Boom, J. A. (1983), Journal of Biomolecular Structure and Dynamics 1, 115−129.
20. Tinoco Jr., I., Borer, P. N., Dengeler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. & Gralla, J. (1973), Nature 246, 40−41.
21. Blommers, M. J. J., Haasnoot, C. A. G., Hilbers, C. W., van Boom, J. H. & van der Marel, G. A., (1987), Structure and Dynamics of Biolpolymers, NATO ASI Ser, E 133, 78−91.
22. Blommers, M. J. J., Walter, J. A. L. I., Haasnoot, C. A. G., Aelen, J. M. A., van der Marel, G. A., van Boom, J. H. & Hilbers, C. W., (1989), Biochemistry 28, 7491−7498.
23. Wolk, S. K., Hardin, C. C., Germann, M. W., van de Sannde, J. H. &Tinoco Jr., I. (1988), Biochemistry 27, 6960 −6967.
24. Chattopadhyaya, R., Ikuta, S., Brzeskowiak, K. & Dickerson, R. E. (1988), Nature 334, 175−179.
25. Senior, M. M., Jones, R. A. & Breslauer, K. J. (1988), Proc. Natl. Acad. Sci. USA 85, 6242−6246.
26. Germann, M. W., Vogel, H. J., Pon, R. T. & van de Sande, J. H. (1989), Biochemistry 28, 6220−6228.
27. Germann, M. W. (1989) PhD. Thesis, The University of Calgary.
28. Chaconas, G. & van de Sande, J. H. (1980), Methods Enzymol. 65, 75−85.
29. Pörschke, D. & Jung, M. (1982), Nucleic Acids Res. 10, 6163−6176.
30. Freeman, R.(1988), A handbook of nuclear magnetic resonance. pp 21−26.
31. Plateau, P. & Guéron, M. (1982), J. Am. Chem. Soc.104, 7310−7311.
32. Goppelt, M., Pingould, A., Maass, G., Mayer, H., Köster, H. & Frank, R. (1980), Eur. J. Biochem. 104, 101−107.
33. Haasnoot, C. A. G., de Bruin, S. H., Hilbers, C. W., van der Marel, G. A. & van Boom, J. H. (1985), Proc. Int. Symp. Biomol. Struct. Interactions, Suppl. J. Biosci. 8, 767−780.
34. Chen, C.-W. & Cohen, J. S. (1984), in Phosphorus-31 NMR Principles and Applications (Gorenstein, D. Ed.) pp 233−263, Academic Press, Orlando.
35. Gorenstein, D. G., Schroeder, S. A., Fu, J. M., Metz, J. T., Roongta, V. & Jones, C.R. (1988), Biochemistry 27, 7223−7237.
36. Alves, J., Pingould, A., Haupt, W., Langowski, J., Peters., F., Maass, G. & Wolff, C. (1984), Eur. J. Biochem. 140, 83−92.
37. Germann, M. W., Kalisch, B. W. & van de Sande J. H. in preparation
38. Conolly, B. A. & Eckstein, F. (1984), Biochemistry 23, 5523−5527.
39. Otting, G., Grütter, R., Leupin, W., Miniganti, C., Ganesh, K. N., Sporat, B. S., Gait, M. J. & Wüthrich, K. (1987), Eur. J. Biochem. 166, 215−220.
40. Giessner-Prettre, C., Pullman, B., Borer, P. N., Kan, L.-S. Ts'o P. O. P. (1976), Biopolymers 15, 2277−2286.
41. McClarin, J.A., Frederick, C.A., Wang, B-C., Greene, P., Boyer, H.W. Grable, J. & Rosenberg, J.M. (1986), Science, 234, 1526−1541.

42. Zhou, N., Germann, M. W., van de Sande J. H. & Vogel, H. J. in preparation.
43. Jen-Jacobson, L., Kurpiewski, M., Lesser, D., Grable, J., Boyer, H. W., Rosenberg, J. M., & Greene, P. J. (1983), J. Biol. Chem. 258, 14638–14646.
44. Patel., D.J., Kozlowski, S. A., Marky., L. A., Broka, C., Rice., J. A., Itakura, K. & Breslauer, K. J. (1982), Biochemistry 21, 428–436.