Biochimie (1993) 75, 433–441

Conformational studies of hairpin sequences from the ColE1 cruciform

NB Blatt, SE Osborne, RJ Cain, GD Glick*

Department of Chemistry, University of Michigan, Ann Arbor, MI 48109, USA

(Received 5 January 1993; accepted 11 January 1993)

Summary — Inverted repeat sequences derived from the ColE1 cruciform were investigated by nuclear magnetic resonance (NMR) and UV spectroscopy. It was shown that 15 different sequences exist as stable hairpin structures over a range of buffer conditions and DNA concentrations. Experiments with six oligomers (**1–6**) containing the native stem sequence and five base loops, found that the two hairpins with the wild-type loops (**1–2**) served as upper and lower bounds for the thermodynamic stability of all the other sequences. NMR experiments, including rotational correlation time measurements and NOESY spectra, were then performed on **1**, the most stable hairpin sequence to begin to uncover a structural basis of its stability.

DNA / hairpin stem-loop / nuclear magnetic resonance(NMR) / cruciform / thermodynamic stability

Introduction

DNA is structurally more complex than the archetypal right-handed helix proposed by Watson and Crick. Of the many DNA geometries that have been observed, considerable attention has focused on cruciform structures because they are thought to play an important role in the control of transcription, translation, and other biological functions ([1–6] and references therein). The biological importance attributed to cruciforms has motivated the study of short palindromic, or near-palindromic sequences for understanding the more complex systems *in vivo*.

Nearly any self-complementary DNA sequence can form a hairpin structure. In completely palindromic oligomers and those with mismatches up to two base-pairs, an equilibrium between a hairpin and a duplex with an internal loop usually exists [7–16]. However, a conformational transition between the hairpin and duplex forms occurs in sequences that contain three unpaired bases [7, 16, 17]. Once a sequence has four or more unpaired bases, it generally adopts only a hairpin stem-loop structure [5–7, 10, 14, 16–19]. In these oligomers the duplex structure is presumably disfavored because it is thermodynamically less stable to have multiple base-pair mismatches in a bulge than it is to form a hairpin structure.

*Correspondence and reprints

The main factor in determining the thermal stability of a given hairpin is its stem composition and length. Longer stems with high G·C content tend to afford the most stable hairpins. One study, for example, reported that for a six base-pair stem consisting of all G·C base-pairs, the substitution to one A·T base-pair resulted in a ten-degree drop in the melting temperature [10]. Loop composition is considered to be only a minor factor in the overall thermal stability of hairpin structures. For example, recent studies by Benight and Breslauer comparing hairpins with T_4, G_4, C_4, and A_4 loops presented by the same stem found only a three to four degree difference in the melting temperature [5, 18]. However, loop composition is significant when thermodynamic stability is considered. Antao and Tinoco [20] found that different sequences presented on the same stem can produce up to a three kcal/mol difference in the $\Delta G°$ for the hairpin to coil transition.

The relationship between loop size and hairpin stability has led to some confusion. Early NMR studies suggested that DNA hairpins are most stable when they possess loops four to five bases long [7, 16, 19]. Yet other laboratories have found that three base loops are more stable than either four or five base loops, and that hairpins with even two base loops have been detected [12–14, 17]. A possible explanation for this discrepancy was proposed based on the particular sequences under study [14]. For a given sequence, a G·C base-pair adjacent to the loop supports the formation of a two or three base loop, whereas an A·T

base-pair located in the same position can accommodate a loop with four to five bases [7, 21, 22]. However, a complication to this picture has recently emerged in studies which find that alternate base pairs can form within a hairpin loop [23]. A pattern underlying these various results should become clearer as more hairpin sequences are examined.

We have investigated the thermodynamics of hairpin formation using sequences derived from the well known ColE1 cruciform (fig 1) [24]. Specifically, we were interested to determine if inverted repeats from bona fide, highly characterized, cruciforms adopt particularly stable (or unstable) hairpin structures. Two, 21 base sequences from the cruciform (sequences **1–2**) as well as other sequences containing either modified loops (**3–7**) or stems (**8–15**) were investigated at moderate sodium ion concentrations by UV spectroscopy to assess the effects of both stem and loop sequence on hairpin stability. NMR studies of **1**, the most stable of our hairpin sequences, were performed to investigate the basis of its stability.

Materials and methods

Preparation of the DNA

Sequences **1–16** were prepared on a Milligen Cyclone Plus DNA synthesizer using phosphoramidite chemistry following the manufacturer's suggested protocols. The crude DNA was purified by reversed-phase HPLC on a Vydac C_4 column (eluting with acetonitrile–0.1 M triethylammonium acetate, pH 6.6). The fractions containing product were pooled and acidified to pH 3 to remove the 5'-trityl groups. The DNA was then concentrated *in vacuo* and re-purified by HPLC as described above. The triethylammonium counterions on the phosphate groups were exchanged for sodium by centrifugal dialysis against sodium chloride (4 M) followed by dialysis against water. The sequences gave the expected ratio of constituent monomers upon enzymatic degradation with phosphodiesterase and alkaline phosphatase, and were verified chemically by the method of Maxam and Gilbert [25].

Sequence No.	Designator	Sequence
ColE1 wild-type stems:		
1	CATTT	AGCAATCC CATTT GGATTGCT
2	AAATG	AGCAATCC AAATG GGATTGCT
3	A-loop	AGCAATCC AAAAA GGATTGCT
4	C-loop	AGCAATCC CCCCC GGATTGCT
5	G-loop	AGCAATCC GGGGG GGATTGCT
6	T-loop	AGCAATCC TTTTT GGATTGCT
7	T_6-loop	AGCAATCC TTTTTT GGATTGCT
modified stems:		
8		ATCAATCC CATTT GGATTGAT
9		AACAATCC CATTT GGATTGTT
10		AGCAATCT CATTT AGATTGCT
11		AGCAATCA CATTT TGATTGCT
12		ATCAATCC TTTTT GGATTGAT
13		AACAATCC TTTTT GGATTGTT
14		AGCAATCT TTTTT AGATTGCT
15		AGCAATCA TTTTT TGATTGCT
16	Duplex	AGCAATCC GGATTGCT

ColE1 Cruciform:

5'-AAA GTCCTAGCAATCCAAATGGGATTGCTAGGAC CAA-3'
3'-TTT CAGGATCGTTAGGTTTACCCTAACGATCCTG GTT-5'

Fig 1. Hairpin sequences studied from the ColE1 cruciform.

UV spectroscopy

Optical melting experiments were conducted using a Cary 3 spectrophotometer interfaced to a Varian Peltier and Compu-Add 212 computer following standard protocols [26]. For all experiments, the absorbance was monitored at 260 nm while the temperature was increased linearly at 1.0°C/min (no difference in the melting curves was observed when the heating rate was 0.5°C/min). Prior to ramping, all solutions were heated to 90°C and then allowed to slowly return to the starting temperature. For each sodium ion concentration, melting experiments were performed in 1.0-cm path-length cuvettes over a 25-fold DNA concentration range (300 nM - 8 μM, DNA molar concentrations are based upon the extrapolated value of the upper baseline at 25°C using an extinction coefficient calculated from nearest-neighbor interactions [26]). An additional DNA concentration (80 μM) was also performed in 0.1-cm path-length cuvettes. Ten absorbance measurements were collected at equal intervals for every 1°C. Five different buffers were used for the optical melting experiments. Samples were prepared either by dialysis against the appropriate buffer or by dissolving an aliquot of DNA that was stored in water into the desired buffer solution. The four buffers with the largest sodium ion concentration contained NaCl (1.0, 0.10, 0.010, or 0.0010 M), EDTA (0.0010 M), and NaH_2PO_4 (0.010 M). An additional lower salt buffer was used in order to evaluate sequences **1** and **2** at a millimolar sodium ion concentration. This buffer consisted of NaCl (1.0 mM), EDTA (0.1 mM), and NaH_2PO_4 (1.0 mM). All five buffers were adjusted to pH 7.0 with NaOH (1.0 M).

Analysis of the melting profiles

Hairpin to coil transitions were analyzed using the non-linear least squares method as described by Puglisi and Tinoco [26] and Xodo *et al* [6]. The analyses were performed on an IBM ES/9000-720 computer running the BMDP statistical software package (program 3R). The parameter estimates and variances from the non-linear regression program were converted into $\Delta H°$, $\Delta S°$, and T_m values using standard statistical formulas [27].

Duplex to coil transitions were analyzed following the method of Marky and Breslauer [28]. Upper and lower baselines were fit to the melting profiles and then the absorbance *versus* temperature profiles were converted into fraction of initial state (often referred to as α or *f*) *versus* temperature profiles. The melting temperature is then defined as the temperature at which a $\alpha = 0.5$. The thermodynamic parameters for these transitions were then determined from a plot of the concentration dependence of the melting temperature. Based on experimental reproducibility, the variance in the thermodynamic parameters reported is ± 0.5°C for the T_m values, ± 0.2 kcal/mol for $\Delta G°$ (298 K), and ± 5% for $\Delta H°$ and $\Delta S°$. These variances are consistent with what was reported in other recent studies [20].

NMR spectroscopy

NMR experiments were performed on Bruker AMX-600, AMX-500, and GE GN 500 NMR spectrometers. DNA

samples of **1** (2 mM in single strands) were prepared in H_2O/D_2O (90/10) containing NaCl (50 mM) and NaH_2PO_4 (10 mM) and adjusted to pH 7.0 with NaOH (1 M). Two-dimensional spectra were measured at 10°C with TPPI phase cycling [29]. The intensity of the solvent peak was suppressed either by presaturation or by the jump and return method [30]. Phase sensitive NOESY spectra (τ_m = 100, 200 ms) were acquired as 2048 points in T2 and 512 points in T1 with a spectral width of 12 500 MHz [31]. The data in each time domain were apodized using a 60° shifted sinebell, then zero-filled to 8192 points. Digital resolution of 1.5 Hz/point in each dimension was obtained after zero-filling. Phase sensitive TOCSY spectra (τ = 52 ms) were acquired as 2048 points in T2 and 512 points in T1 [32]. A 90° shifted sinebell was used for apodization in each dimension. The T1 vectors were then zero-filled to 2048 points before transformation to a square matrix. For melting studies, a series of spectra were measured at five-degree intervals over the temperature range 10–85°C. To determine the rotational correlation time of **1**, truncated driven nuclear Overhauser effect experiments were performed at 500 MHz using a sample that was 7 mM in DNA [33]. Data processing and analysis of all spectra were carried out on a Silicon Graphics 4D/35 workstation using the FELIX v2.0 software package.

Results

UV spectroscopy

In all buffer and DNA concentrations, sequences **1–15** show a single cooperative melting transition. Figure 2 shows representative melting curves in all five buffers for sequence **2**. These curves for all the hairpins are reproducible upon renaturation, and no hysterisis is observed. The melting temperatures of each sequence are independent of DNA concentration (table I). These data indicate that the observed transitions for sequences **1–15** are from hairpin to coil.

The percentage difference between the fitted and actual melting profiles for sequences **1–15** is within 0.25% for each data point. The excellent agreement between these curves indicates that each of these sequences thermally denature in a two-state fashion. The van't Hoff enthalpy and entropy for the hairpin to coil transition of **1–15** are listed in table I. Due to the high statistical correlation between $\Delta H°$ and $\Delta S°$ in the data analysis, the variance in these two parameters is slightly larger than for the melting temperature [34, 35]. Within this variance the values of $\Delta H°$ and $\Delta S°$ for each sequence are independent of DNA concentration (data not shown).

Sequence **16**, the stem duplex of hairpins **1–15**, showed a single cooperative melting transition that was dependent upon the strand concentration (fig 3). The melting temperature for this duplex was significantly lower than that of the hairpins, and therefore, melting profiles were only obtained in the buffers with the three highest sodium ion concentrations. No hysterisis is observed with the melting of this stem duplex.

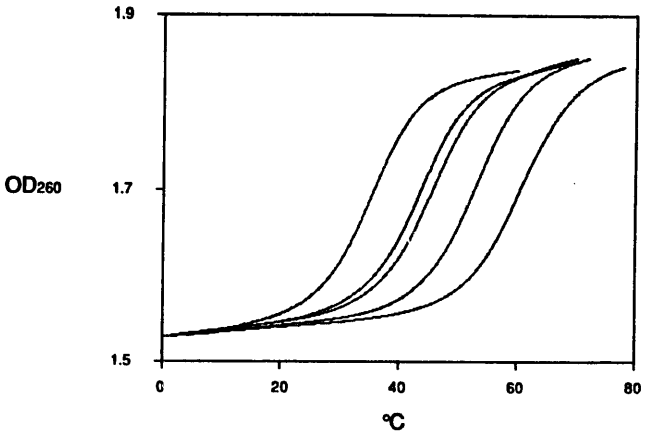

Fig 2. Representative melting profiles of sequence **2** at all five sodium ion concentrations (2.5 mM, 12 mM, 21 mM, 0.11 M, and 1.0 M). DNA concentration is 8 μM in single strands. Similar profiles were obtained for the other hairpins studied.

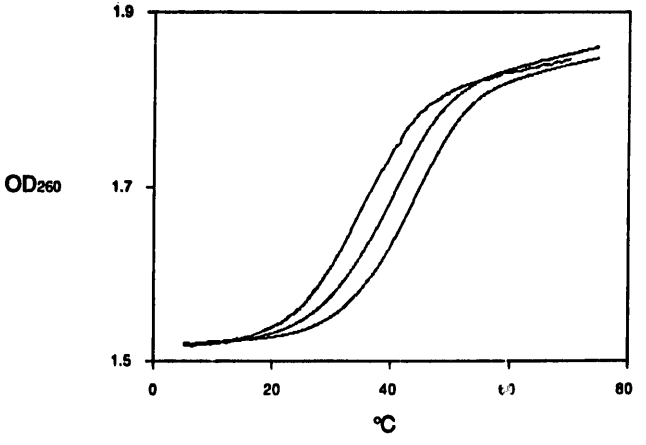

Fig 3. Representative melting profiles demonstrating the concentration dependence of the melting of stem duplex (**16**). DNA concentrations are 1.25, 5.0, and 12.5 μM in single strands. The melting buffer contains 1.0 M NaCl, 0.010 M NaH_2PO_4, 0.0010 M EDTA (pH 7).

The thermodynamic parameters describing the melting of this sequence can be found in table I.

The melting of the stem dupex (**16**) along with hairpins **1** and **2** showed significant dependence upon the sodium ion concentration. Figures 4 and 5 show the linear relationship between the log of the sodium ion concentration *versus* the melting temperature for each of the sequences. The salt dependence of the melting temperature reflects the fact that the single stranded melted DNA has a lower charge density than the double stranded stem duplex. The linear relationship found for the three sequences is characteristic of the

Table I. Thermodynamic parameters for the melting of sequences **1–16**. The variance associated with each value is discussed in the *Materials and methods* section. An additional sequence was also examined that had the wild-type stem and four base thymidine loop. This compound did not show monophasic cooperative melting and was not investigated further.

Sequence no	*[Na⁺] (M)*	*T_m (°C)*	*$\Delta S°$ (cal/mol•K)*	*$\Delta H°$ (kcal/mol)*	*$\Delta G°$ 298 K (kcal/mol)*
1	1.0	81.4	165	58.6	9.3
2	1.0	80.7	146	51.8	8.2
16	1.0		149	53.3	8.8
1	0.11	73.3	184	63.8	8.9
2	0.11	73.2	153	53.1	7.4
3	0.11	73.2	159	55.1	7.7
4	0.11	71.4	177	60.9	8.2
5	0.11	73.3	167	57.9	8.1
6	0.11	75.0	171	59.6	8.5
7	0.11	71.4	186	64.1	8.6
8	0.11	65.5	159	53.8	6.4
9	0.11	67.0	172	58.6	7.2
10	0.11	66.5	176	59.8	7.3
11	0.11	66.7	176	59.9	7.3
12	0.11	66.8	169	57.3	7.0
13	0.11	68.3	174	59.4	7.5
14	0.11	67.8	179	61.0	7.7
15	0.11	68.1	187	63.7	8.1
16	0.11		170	58.4	7.7
1	0.021	65.2	183	62.0	7.4
2	0.021	65.5	153	51.7	6.2
16	0.021		176	58.7	6.0
1	0.012	63.2	188	63.2	7.2
2	0.012	63.4	156	52.6	6.0
1	0.0025	54.2	182	59.5	5.3
2	0.0025	54.5	161	52.8	4.8

melting of DNA duplexes and hairpin structures [15, 26, 36, 37].

NMR spectroscopy

Sequential assignment of the proton resonances in **1** was performed as outlined by Wüthrich [38]. We have unambiguously assigned the non-exchangeable base protons, the ribose H1', H2'/2", H3', H4', 25 of the 42 H5'/5" protons, and the hydrogen bonded imino protons. The individual resonance assignments will be reported elsewhere (Glick *et al*, unpublished results). The pattern of NOE cross-peaks observed between the base protons and H1' and 2'/2" sugar protons is consistent with a B-DNA helix extending throughout the stem. Figure 6 shows an expanded contour plot of the NOESY connectivity between thymidine methyl protons and base protons. In a B-DNA duplex the thy-

midine methyl protons exhibit a strong NOE with its own H6 proton and with the H6/8 proton of the preceding base. The weak intensity of the sequential NOE cross-peaks of the loop methyls (T11, T12, and T13) clearly demonstrates that the B-form duplex is disrupted at these positions and is consistent with a hairpin structure. The NMR spectra indicate that base stacking of the stem duplex continues into the loop in the 3' direction, but spectral overlap prevents the complete determination of its extent.

Figure 7 shows the observed NOE cross-peaks that correlate protons on opposite sides of the duplex. Since these NOE measurements were conducted with presaturation to reduce the intensity of the solvent peak, all labile protons exposed to solvent were attenuated. No other imino resonances were observed, even when the jump and return method was used for solvent suppression. This finding indicates that the stem imino protons are appreciably less solvent exposed than the T11, T12, T13, or T21 imino protons. The NOE pattern demonstrates that a stable duplex extends from the G2–C20 base-pair through the C8–G14 base-pair, while the terminal base-pair is frayed.

The ^{1}H NMR spectrum of **1** is independent of DNA concentration over a wide range of conditions (45 μM to 7 mM in single strands). Multiple resonances for a given proton, indicative of multiple conformational states, are not detected. For example, the ^{1}H NMR spectrum of **1** at 10°C contains seven sharp methyl resonances between 1.3 and 1.8 ppm. While this finding is also consistent with a hairpin that is in rapid equilibrium with a duplex structure, this prospect is unlikely since the hairpin to coil transition for similar oligomers is slow on the NMR time scale [8, 39]. Figure 8 shows the NMR melting profiles using the thymidine methyl protons as reporter groups. These plots indicate that **1** melts in a two-state manner and are in qualitative agreement with the UV melting profiles ($\Delta T_m \sim 5$°C).

To unambiguously differentiate between the hairpin and bulged duplex forms, we determined the overall rotation correlation time of **1** through a series of truncated driven NOE spectra [33]. Following the procedure of Clore and Gronenborn [40], the build-up

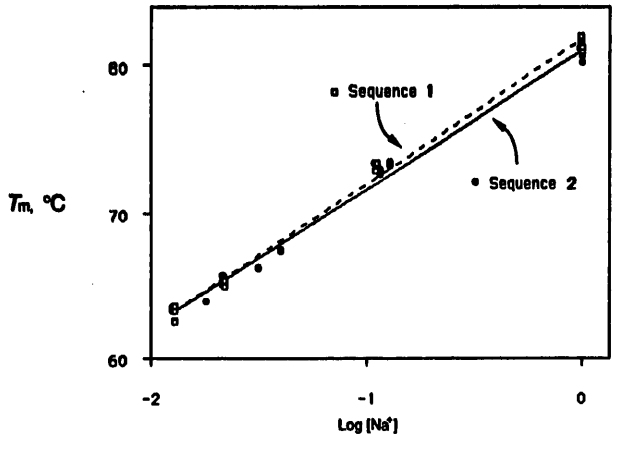

Fig 4. Plot of T_m *versus* the log sodium ion concentration for sequences **1** and **2**. The dashed line indicates the linear least-squares fit for sequence **1** (r = 0.996), the solid line for sequence **2** (r =0.995).

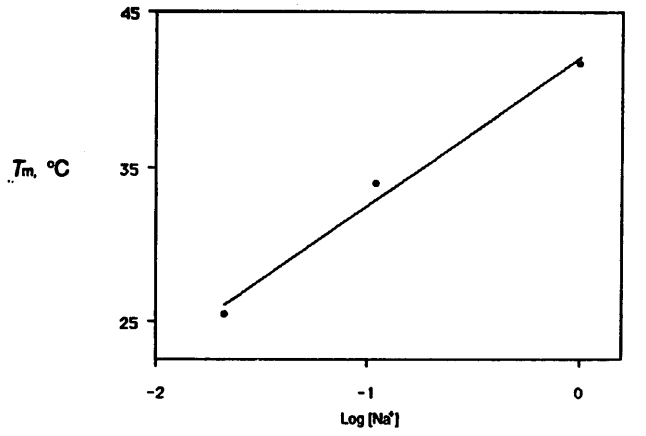

Fig 5. Plot of T_m *versus* the log sodium ion concentration for sequences **16** (r =0.994). The melting temperatures are based on a total strand concentration of 10 μM.

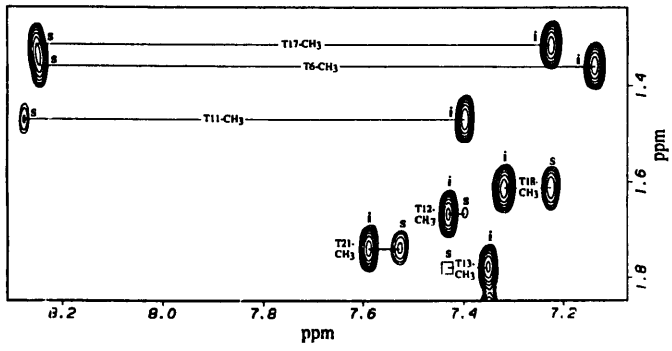

Fig 6. Expanded NOESY contour plot showing the connectivity between the thymidine methyl protons (vertical axis) and the aromatic base protons (horizontal axis). Each thymidine methyl exhibits a sequential NOE (*s*) with the H6/8 proton of the previous base and an intranucleotide NOE (i) with the thymidine H6 proton of the same base. The horizontal lines connect the NOEs of each thymidine methyl group. Note that the loop methyls (T11, T12, and T13) exhibit only weak internucleotide NOEs indicating that the B-form helix does not extend through these bases.

rates for the H6-H5 inter-proton vectors of C3, C7, and C8 (spectral overlap prevented analysis of other resonances) were measured at low mixing times when spin diffusion is minimal and relaxation rate is approximately equal to the dipolar relaxation rate (1):

$$\sigma_{ij} = \frac{\gamma^4 h^2}{40\pi^2 r_{ij}^6}\left(\tau_{app} - \frac{6\tau_{app}}{1 + 4\omega^2\tau_{app}^2}\right) \quad (1)$$

where σ_{ij} is the cross-relaxation rate, γ is the gyromagnetic ratio for a proton, r_{ij} is the interproton distance, and τ_{app} is the apparent correlation time for the interproton vector. Since the H6-H5 inter-proton distance is fixed at 2.46 Å, the apparent rotational correlation can be calculated from the NOE build-up rate. Each vector gave a rotational correlation time for sequence **1** of 5.8 ± 0.6 ns (the equivalence of the correlation times suggests that **1** acts as a rigid isotropic rotor).

To place the value of the overall rotational correlation time for **1** into context, it is helpful to compare it to values previously determined for DNA duplexes of varying size. Duplexes that are 20 base pairs long are expected to be hydrodynamically similar to the bulged duplex form of **1** and have been shown to have rotational correlation times between 9.2 and 11.1 ns at 10°C [41, 42]. In addition, DNA duplexes with 12 base pairs should be hydrodynamically slightly larger than the hairpin form of **1**, and have been shown to have a rotational correlation time of 6.1 ns at 8°C [43, 44]. Taken together, these results support the conclusion that **1** adopts a hairpin structure at concentrations up to 7 mM.

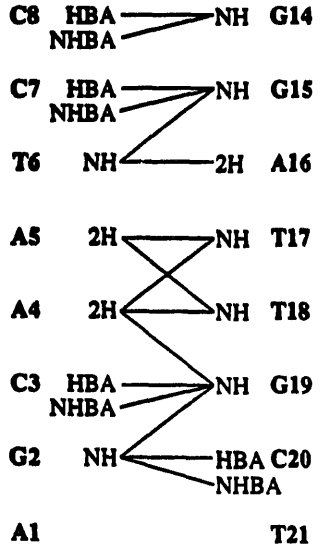

Fig 7. A schematic representation of the NOE connectivities observed between protons on opposite sides of the stem-duplex. NOEs were taken from the 600 MHz NOESY described in the text. HBA = hydrogen bonded 4-amino proton of cytosine; NHBA = non-hydrogen bonded 4-amino proton of cytosine; NH = imino proton; and 2H = 2H proton of adenosine.

Discussion

Our UV data indicate that sequences **1–15** exist exclusively as hairpin structures over a range of DNA concentrations at moderate sodium ion concentrations ($[Na^+]$ = 0.1 M). In addition, these data indicate that **1** and **2** are stable hairpins at both higher and lower sodium ion concentrations ($[Na^+]$ = 1.0 M, 10 mM, 3 mM). Sequences **1–15** migrate as 14-mers on nondenaturing polyacrylamide gels over a similar range of conditions providing further evidence that they adopt hairpin structures (data not shown). Even at the high DNA concentrations used for NMR measurements of **1** (2–7 mM in single strands), no evidence for duplex formation is observed.

The T_m values indicate that **1–6** are among the most thermally stable hairpin structures that have been reported [5–18, 20]. The ΔH° and ΔS° for hairpin **1** are greater than the other sequences, whereas ΔH° and ΔS° for sequence **2** are significantly smaller than for the other hairpins. This finding is of particular interest since hairpins **1** and **2** form both halves of the ColE1 colE1 cruciform. These differences can be explained from the shape of melting curves during the transition. As figure 9 shows, sequence **1** has a much sharper melting transition relative to sequence **2**. The slope of the curve at the melting temperature is related to the van't Hoff transition enthalpy [28]. A steeper slope at the transition temperature for sequence **1** indicates a larger ΔH° given the same melting temperature. This difference will also correspond to a difference in the entropy terms given that ΔH° = TΔS° at the melting temperature.

By comparing the thermodynamic data for sequences **1–7** to the melting of the stem duplex (**16**), it is found that some of the loop sequences are stabilizing whereas others are destabilizing. Considering ΔG° (298 K), the AAATG loop is destabilizing, whereas the CATTT, C_5, G_5, T_5, and T_6 loops stabilize the hairpin structure relative to the duplex stem. The A_5 loop, however, does not add to, or decrease the stability of **16**. NMR melting studies of **1** can provide insight into the loop conformation of this hairpin and may even suggest why sequence **1** is a 'stabilizing loop'. Of the seven thymidines in **1**, five show cooperative sigmoidal melting profiles that indicate the tran-

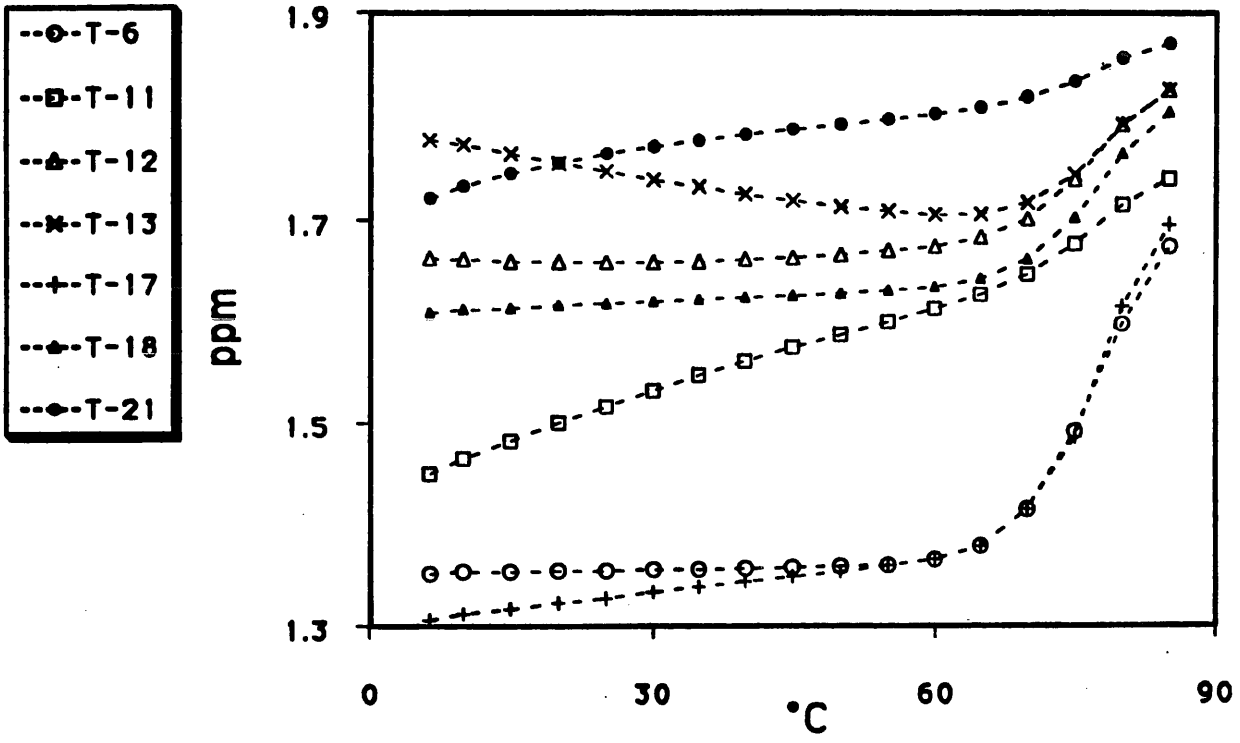

Fig 8. Variation in chemical shift with temperature of the thymidine methyl protons of sequence **1**.

sition from a greater to a lesser solvent-shielded state (fig 8). Significantly, two of these five thymidines (T12, T13) are located in the loop suggesting that they are engaged in favorable stacking interactions.

The thermodynamic stability of sequences **3–6** are in general agreement with previous studies of hairpins with homonucleotide loops [5, 18]. The order of increasing stability ($\Delta G°$ at 298 K) is A_5-loop < G_5-loop < C_5-loop < T_5-loop, and can be generalized to purines < pyrimidines. Interestingly, sequence **1** which has the CATTT loop is more stable than the T_5-loop, while sequence **2** with a loop consisting of AAATG is less stable than the A_5-loop. This finding is surprising since based on the order purines < pyrimidines one might have predicted that the CATTT loop would be less stable than the C_5- or T_5-loops because a purine was substituted for pyrimidine, when in fact the opposite is true. These results suggest that model studies using homonucleotide loops cannot adequately represent the range of stabilities that can be found in all possible hairpin loops.

A trend can be seen comparing the stability of hairpins with modified stems. Substituting an A·T (or T·A) base pair for a G·C base pair at the end of the stem decreases the stability more than making a substitution at the base pair adjacent to the loop. This is observed for both hairpins **8–11** containing the CATTT loop, and **12–15** possessing a T_5-loop. This result can be understood by recalling the cooperative nature of hairpin melting. A substitution to a weaker base pair near the end of the stem duplex means that it will be easier to begin the cooperative 'unpairing' of the base pairs required for melting to occur [10, 19]. In contrast, a weaker base pair adjacent to the loop will

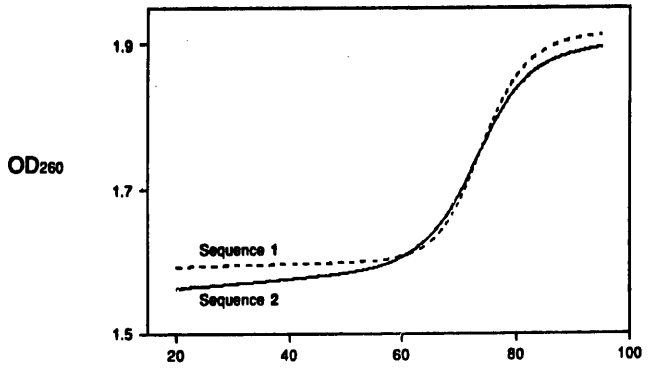

Fig 9. Comparison of representative melting profiles of sequences **1** and **2**. DNA concentration is 8 μM in single strands. The melting buffer contains 0.10 M NaCl, 0.010 M NaH_2PO_4, 0.0010 M EDTA (pH 7).

not have the same effect because the loop itself can protect the base pair from fraying and initiating the cooperative denaturation.

It was mentioned above that for the hairpins with the wild type stem structure, the CATTT loop is more stable than the T_5-loop. Interestingly, if sequences **8** and **12**, **9** and **13**, **10** and **14**, or **11** and **15** are compared, then the opposite pattern is found. For each of these modified stems, the T_5-loop is more stable than the CATTT loop. These findings are in general agreement with the hairpin 'folding principle' put forward by Hassnoot *et al* [16], and they suggest that hairpin stability can depend on both the loop and stem sequence.

While it may not be possible to accurately predict the stability of a hairpin sequence, whether a given oligomer adopts a hairpin conformation depends upon the interaction between the stability of the stem duplex and the number of base-pair mismatches in the putative loop region. If base-pairing in the stem is extremely stable, then a bulged duplex structure can form independent of the number of mismatches in the loop. An example of this comes from a sequence with a loop consisting of five adenines attached to an alternating C·G six base-pair stem that showed biphasic melting curves even at a sodium ion concentration of 0.5 mM [14]. If the stem duplex is not enthalpically stabilized by a sufficient number of G·C base-pairs then a stable hairpin structure may not form. For example, a six base stem consisting of alternating A·T base-pairs with a five base cytosine loop has a T_m of only 28°C in 0.1 M NaCl buffer [14]. The sequences based on the stem duplex reported in this work are 50% G·C in the stem duplex (and the entire stem region of the ColE1 cruciform is 53.8% G·C).

Conclusion

We have investigated the thermodynamics of hairpin formation using 15 sequences derived from ColE1 cruciform to assess the effects of both stem and loop sequence on hairpin stability. We find that hairpins **1** and **2**, which form both halves of the cruciform, serve as bounds for the stability of hairpins in which the loop sequence has been changed to all A, T, G, or C. The melting data using modified stem duplexes suggest that changes in the stem sequence both proximal and distal to the loop have significant effects on the relative stability of the given hairpin. This finding may complicate any future attempts at predicting the thermodynamic stability of a hairpin based solely upon its sequence. As cruciform structures are increasingly implicated in biological processes, our studies reinforce the need for studies that address the stability of the constituent hairpin structures [45].

Acknowledgments

This research was supported in part by a grant from the National Institutes of Health (GM 46831). We thank J Warner of the University of Michigan Center for Statistical Consultation and Research for assistance with the UV data analysis and ERP Zuiderweg for assistance with the correlation time measurements.

References

1 Horwitz MSZ, Loeb LA (1988) An *E coli* promoter that regulates transcription by DNA superhelix-induced cruciform extrusion. *Science* 241, 703–705
2 Wells RD, Hanrey SC (eds) (1988) *Unusual DNA Structures, Proceedings of the First Gulf Shores Symposium.* Springer-Verlag, Berlin
3 Wells RD (1988) Unusual DNA structures. *J Biol Chem* 263, 1095–1098
4 Müller UR, Fitch WM (1982) Evolutionary selection for perfect hairpin structures in viral DNAs. *Nature (Lond)* 298, 582–585
5 Paner TM, Amaratunga M, Doktycz MJ, Benight AS (1990) Analysis of melting transitions of the DNA hairpins formed from the oligomer sequences d[GGATAC-$(X)_4$GTATCC] (X = A,T,G,C). *Biopolymers* 29, 1715–1734
6 Xodo LE, Manzini G, Quadrifoglio F, van der Marel G, van Boom JH (1989) Hairpin structures in synthetic oligodeoxynucleotides: sequence effects on the duplex-to-hairpin transition. *Biochimie* 71, 793–803
7 Haasnoot CAG, den Hartog JHJ, de Rooij JFM, van Boom JH, Altona C (1980) Loopstructures in synthetic oligodeoxynucleotides. *Nucleic Acids Res* 8, 169–181
8 Summers MF, Byrd RA, Gallo KA, Samson CJ, Zon G, Egan W (1985) Nuclear magnetic resonance and circular dichroism studies of a duplex – single-stranded hairpin loop equilibrium for the oligodeoxyribonucleotide sequence d(CGCGATTCGCG). *Nuleic Acids Res* 13, 6375–6386
9 Roy S, Weinstein S, Borah B, Nickol J, Apella E, Sussman JL, Miller M, Shindo H, Cohen JS (1986) Mechanism of oligonucleotide loop formation in solution. *Biochemistry* 25, 7417–7423
10 Pramanik P, Kanhouwa N, Kan L (1988) Hairpin and duplex formation in DNA fragments CCAATTTTGG, CCAATTTTTTGG, and CCAATTTTTGG: a proton NMR study. *Biochemistry* 27, 3024–3031
11 Xodo LE, Manzini G, Quadrifoglio F, van der Marel GA, van Boom JH (1988) The duplex-hairpin conformational transition of d(CGCGCGATCGCGCG) and d(CGCGCG-TACGCGCG): a thermodynamic and kinetic study. *J Biomol Struct Dyn* 6, 139–152
12 Orbons LPM, van der Marel GA, van Boom JH, Altona C (1986) Hairpin and duplex formation of the DNA octamer d(m^5C-G-m^5C-T-G-m^5C-G) in solution. An NMR study. *Nucleic Acids Res* 14, 4187–4196
13 Orbons LPM, van der Marel GA, van Boom JH, Altona C (1987) An NMR study of the polymorphous behavior of the mismatched DNA octamer d(m^5C-G-m^5C-G-A-m^5C-G) in solution. *Eur J Biochem* 170, 225–239
14 Xodo LE, Manzini G, Quadrifoglio F, van der Marel GA, van Boom JH (1988) Oligodeoxynucleotide folding in solution: loop size and stability of B-hairpins. *Biochemistry* 27, 6321–6326
15 Marky LA, Blumenfeld KS, Kozlowski S, Breslauer KJ (1983) Salt-dependent conformational transitions in the

self-complementary deoxydodecanucleotide d(CGCGAATTCGCG): evidence for hairpin formation. *Biopolymers* 22, 1247–1257
16 Haasnoot CAG, Hilbers CW, van der Marel GA, van Boom JH, Singh UC, Pattabiraman N, Kollman PA (1986) On loopfolding in nucleic acid hairpin-type structures. *J Biomol Struct Dyn* 3, 843–857
17 Germann MW, Kalisch BW, Lundberg P, Vogel HJ, van de Sande JH (1990) Perturbation of DNA hairpins containing the *Eco*RI recognition site by hairpin loops of varying size and composition: physical (NMR and UV) and enzymatic (*Eco*RI) studies. *Nucleic Acids Res* 18, 1489–1498
18 Senior MM, Jones RA, Breslauer KJ (1988) Influence of loop residues on the relative stabilities of DNA hairpin structures. *Proc Natl Acad Sci USA* 85, 6242–6246
19 Haasnoot CAG, de Bruin SH, Berendsen RG, Janssen HGJM, Binnendijk TJJ, Hilbers CW, van der Marel GA, van Boom JH (1983) Structure, kinetics, and thermodynamics of DNA hairpin fragments in solution. *J Biomol Struct Dyn* 1, 115–129
20 Antao VP, Tinoco I (1992) Thermodynamic parameters for loop formation in RNA and DNA hairpin tetraloops. *Nucleic Acids Res* 20, 819–824
21 Raghunathan G, Jernigan RL, Miles HT, Sasisekharan V (1991) Conformational feasibility of a hairpin with two purines in the loop. 5'-GGTACIAGTACC-3'. *Biochemistry* 30, 782–788
22 Orbons LPM, van Beuzekom AA, Altona C (1987) Conformational and model-building studies of the hairpin form of the mismatched DNA octamer d(m^{5}C-G-m^{5}C-T-G-m^{5}C-G). *J Biomol Struct Dyn* 4, 965–987
23 Hilbers CW, Blommers MJJ, van de Ven FJM, van Boom JH, van der Marel GA (1991) High resolution NMR studies of DNA hairpins with four nucleotides in the loop region. *Nucleosides Nucleotides* 10, 61–80
24 Lilley DMJ (1989) Structural isomerization in DNA: the formation of cruciform structures in supercoiled DNA molecules. *Chem Soc Rev* 18, 53–83
25 Maxam AM, Gilbert W (1980) Sequencing end-labeled DNA with base-specific chemical changes. *Methods Enzymol* 65, 497–559
26 Puglisi JD, Tinoco I (1989) Absorbance melting curves of RNA. *Methods Enzymol* 180, 304–325
27 Kish L (1965) *Survey Sampling*. J Wiley, New York
28 Marky LA, Breslauer KJ (1987) Calculating thermodynamic data for the transition of any molecularity from equilibrium melting curves. *Biopolymers* 26, 1601–1620
29 Marion D, Wüthrich K (1983) Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurement of ^{1}H-^{1}H spin-spin coupling constants in proteins. *Biochem Biophys Res Comm* 113, 967–974
30 Plateau P, Gueron M (1982) Exchangeable proton NMR without baseline distortion, using new strong-pulse sequences. *J Am Chem Soc* 104, 7310–7311
31 States DJ, Habenkorn RA, Ruben DJ (1982) A two-dimensional nuclear Overhauser experiment with pure absorption phase in four quadrants. *J Mag Res* 48, 286–292
32 Braunschweiler L, Ernst RR (1983) Coherence transfer by isotropic mixing application to proton spectroscopy. *J Magn Res* 53, 521–528
33 Wagner G, Wüthrich K (1979) Truncated driven nuclear Overhauser effect (TOE). A new technique for studies of selective ^{1}H-^{1}H Overhauser effects in the presence of spin diffusion. *J Magn Res* 33, 675–680
34 Nadeau JG, Gilham PT (1985) Anomalous hairpin formation in an oligodeoxyribonucleotide. *Nucleic Acids Res* 13, 8259–8274
35 Bower M, Summers MF, Kell B, Hoskins J, Zon G, Wilson WD (1987) Synthesis and characterization of oligodeoxyribonucleotides containing terminal phosphates. NMR, UV spectroscopic, and thermodynamic analysis of duplex formation of [d(pGGAATTCC)]$_2$, [d(GGAATTCCp)]$_2$, and [d(pGGAATTCCp)]$_2$. *Nucleic Acids Res* 15, 3531–3547
36 Rentzeperis D, Kharakoz DP, Marky LA (1991) Coupling of sequential transitions in a DNA double hairpin: energetics, ion binding, and hydration. *Biochemistry* 30, 6276–6283
37 Erie DA, Jones RA, Olson WK, Sinha NK, Breslauer KJ (1989) Melting behavior of a covalently closed, single-stranded, circular DNA. *Biochemistry* 28, 268–273
38 Wüthrich K (1986) *NMR of Proteins and Nucleic Acids*. J Wiley, New York
39 Altona C (1987) Versatile oligonucleotides: B DNA, Z DNA, and DNA hairpins as seen in aqueous solution by two-dimensional NMR. *Nucleosides Nucleotides* 6, 157–172
40 Clore GM, Gronenborn AM (1984) Internal mobility in a double-stranded B DNA hexamer and undecamer. A time-dependent proton-proton nuclear Overhauser enhancement study. *FEBS Lett* 172, 219–225
41 Lane AN (1989) NMR assignments and temperature-dependent conformational transitions of a mutant trp operator-promoter in solution. *Biochem J* 259, 715–724
42 Lane AN, Lefevre JF, Jardetzky O (1986) A method for evaluating correlation times for tumbling and internal motion in macromolecules using cross-relation rate constants from NMR spectra. *J Magn Res* 66, 201–218
43 Borden KLB, Jenkins TC, Skelly JV, Brown T, Lane AN (1992) Conformational properties of the G·G mismatch in d(CGCGAATTGGCG)$_2$ determined by NMR. *Biochemistry* 31, 5411–5422
44 Cantor CR, Schimmel PR (1980) *Biophysical Chemistry Part II: Techniques for the study of biological structure and function*. WH Freeman and Co, New York, 549–565
45 Bianchi ME, Beltrame M, Paonessa G (1989) Specific recognition of cruciform DNA by nuclear protein HMG1. *Science* 243, 1056–1059