Structural flexibility of a DNA hairpin located in the long terminal repeat of the Drosophila 1731 retrotransposon

A. Slama-Schwok*, E. Brossalina, Y. Demchenko1, M. Best-Belpomme2 and V. Vlassov1

INTRODUCTION

Hairpins offer binding sites for specific proteins because they adopt distinct structures (1 and references therein, 2–7). RNA hairpins are generally conserved and often possess a regulatory role (8–11). They are involved in protein translation, in mRNA splicing (12–14,15 and references therein, 16), in retroviral packaging and reverse transcription (17–20).

The existence and function of DNA hairpins are less well documented. They could act at the transcription and replication level (21–23,24 and references therein). DNA hairpins and cruciforms are determinants for topoisomerase II recognition and cleavage (25–28). Hairpins may be associated with methylation processes occurring at the onset of the fragile X syndrome (29–30).

In the present work, we characterize the structure of a DNA element, located in the U3 part of the long terminal repeat (LTR) of the Drosophila 1731 retrotransposon (31). This sequence, named Bc (c stands for coding strand), belongs to the core promoter. It contains a limited palindrome that may form a hairpin (Scheme 1). Bc constitutes the binding site for two nuclear single-stranded binding factors, NssBF and p11 (31–33). Contact points between NssBF and the 1731 LTR are distributed in two distinct domains, 6 and 15 nt long, separated by the 5 nt GAACAC (31). The sequence of these 5 nt can be changed without affecting NssBF binding, but the distance between the two domains has to be maintained for efficient protein binding (32).

In vitro transcription of the 1731 promoter is repressed by Bc binding protein(s) (31–33). Bc could thus be implicated in transcriptional repression (32). Bc and its complementary strand constitute one of the two major sites of topoisomerase II cleavage (34). The architectural state of DNA plays a role in the regulation of transcription of numerous genes (35,36).

The aim of the present work was to determine the most probable structures of the DNA binding site of the factor NssBF, in the light of our previous results (31–34). Melting temperature (Tm) experiments comparing the native sequence with mutated ones were performed. Bc structures were also investigated using probing techniques and fluorescence measurements using a macrocyclic bis-acridine as a hairpin-selective probe (37–39).

The secondary structures derived from our data mainly involve two hairpins in equilibrium at pH 6.0. In contrast, the Bc complementary strand adopts a single hairpin. Since Bc is implicated in repression of transcription via binding of two specific factors, its structural flexibility could be associated with this process.

ABSTRACT

The structure of the DNA binding site of the Nuclear single-stranded Binding Factor (NssBF), located in the long terminal repeat of the Drosophila 1731 retrotransposon, was investigated by melting temperature experiments, chemical probing and fluorescence measurements using a macrocyclic bis-acridine. The most probable structure of this element, named Bc, mainly involves two hairpins in equilibrium at pH 6.0 at low concentration. The hairpins differ in their apical loop size: 4 and 8 nt. The structural flexibility of Bc probably derives from the three consecutive CATA repeats complementary to the GTAT nucleotides of the palindrom. In contrast, the Bc complementary strand adopts a single hairpin. Since Bc is implicated in repression of transcription via binding of two specific factors, its structural flexibility could be associated with this process.

MATERIALS AND METHODS

Materials

The oligonucleotides, referred to as Bc (c for coding strand), BcI, BcII, H1, H2, Bc2, Bc3, Bc5, Bc6, Bc7 and Bnc (nc for non-coding strand), possess the following sequences, the dashed lines representing deletions (Scheme 1 and Fig. 1).

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bc</td>
<td>5′-TAT GTA ATT TGT TAT GAG AAC ATA CAT ACA TAC ACA TGC A-3′</td>
</tr>
<tr>
<td>BcI</td>
<td>5′-TAT GATT TTT TTA GAG AAC ATA CTG A-GA CTA AGT ACA CAC A-3′</td>
</tr>
<tr>
<td>BcII</td>
<td>5′-TAT GAT ATT TTC TTA GAG AAC ATA CTG A-GA CTA AGT ACA CAC A-3′</td>
</tr>
<tr>
<td>H1</td>
<td>5′-TC TTA GAG AAC ATA GAAG ACAA TTC ACAA TAC ACA TGC A-3′</td>
</tr>
<tr>
<td>Bc2</td>
<td>5′-TG TAT GAG AAC ATCA CAA CTG A-3′</td>
</tr>
<tr>
<td>Bc3</td>
<td>5′-TG TAT GAG AAC ATCA CAA CTG A-3′</td>
</tr>
<tr>
<td>Bc4</td>
<td>5′-TG TAT GAG AAC ATCA CAA CTG A-3′</td>
</tr>
<tr>
<td>Bc5</td>
<td>5′-TG TAT GAG AAC ATCA CAA CTG A-3′</td>
</tr>
<tr>
<td>Bc6</td>
<td>5′-TG TAT GAG AAC ATCA CAA CTG A-3′</td>
</tr>
<tr>
<td>Bc7</td>
<td>5′-TG TAT GAG AAC ATCA CAA CTG A-3′</td>
</tr>
<tr>
<td>Bnc</td>
<td>5′-TG TAT GAG AAC ATCA CAA CTG A-3′</td>
</tr>
</tbody>
</table>

They were purchased from Eurogentec or Genosys and purified by HPLC. The samples were ethanol precipitated, then desalted on a Sephadex G-10 or 25 column. The chemicals used were of the highest commercial purity. All aqueous solutions utilized MilliQ water (Millipore). The macrocyclic bis-acridine was a gift of Prof. J.-M. Lehn (37).
Scheme 1. (Upper) Bc sequence with the palindrome positions indicated by arrows and the nucleotides corresponding to the NssBF binding site shown in red. (Middle) Possible structures of Bc; the numbering arbitrarily begins from the first guanine at the 5′ extremity.

Melting temperatures

The $T_m$ measurements were performed with a Kontron 942 spectrophotometer. The temperature within the cell was measured using a temperature sensor. The melting profiles of the oligonucleotides were monitored at 270 and 260 nm and corrected for baseline fluctuations by subtracting the absorbance at 350 nm. Cells of variable path length were used in order not to exceed 1.7 OD. The samples were heated to 70°C for 30 min then gently cooled to 10°C and maintained at this temperature overnight for equilibration before measurement. Deconvolution of the melting derivatives was performed using Origin 3.5 software.

Probing experiments

Samples of 1 µM Bc, labelled with $^{32}$P at its 5′-end, were incubated for 40 min at 20°C in the presence of 10% DEPC. Incubation was performed in 30 mM cacodylate buffer, pH 6.0, in the presence of 0.5 mM EDTA (unstructured Bc) or in 100 mM KCl or 10 mM MgCl$_2$ as indicated (structured forms). After the reaction, DEPC was extracted with 0.5 ml diethyl ether and Bc was precipitated in ethanol in the presence of 1 µg/ml DNA carrier and 0.3 M sodium acetate, pH 4.8. The modified Bc was then treated with piperidine (1 M piperidine for 10 min at 100°C). The sample was then precipitated in ethanol and analysed on a 20% denaturing polyacrylamide gel. The reaction of Bc with permanganate was performed in the same buffer for 10 min at 20°C or 15 min at 5°C, as indicated, using [Bc] = 1 µM, [KMnO$_4$] = 15 mg/l. After incubation, the same treatment as above was performed. A similar procedure was used for Bnc but the reaction with permanganate was for 5 min at 20°C. The cleavage reaction with S1 nuclease (1 U/µl) was done under the same conditions.

Fluorescence measurements

The fluorescence measurements were performed with a Fluoromax instrument (Spex) equipped with a Hamamatsu 931 photomultiplier and a thermostated cell holder. The data have been corrected for the photomultiplier response and for variation in absorption at the excitation wavelength. Repeated fluorescence spectra of bound to oligomers induce spectral modifications under continuous 360 nm irradiation due to a photochemical reaction, which was avoided by taking short integration times and small excitation path lengths $l = 0.4$ cm.

RESULTS

The Bc oligonucleotide forms hairpins

Melting profile of Bc at pH 6.0, 100 mM KCl. The melting profiles of Bc are reversible at very slow heating rates (4–6°C/h). The curve derivatives display a main broad peak at 38.5 ± 0.5°C that remains unchanged at a 150-fold increase in strand concentration; [Bc] = 0.3–53.5 µM (Fig. 1 and Table 1). A comparison of curves obtained at 1 and 53 µM reveals an additional lower transition that depends on Bc concentration. Since the data suggested that Bc might adopt more than one structure, the derivatives have been deconvoluted, assuming a gaussian shape to each peak. Typical calculations are presented in Figure 1, curve 1. They yield a low temperature concentration-dependent peak and two concentration-independent maxima at 37.2 ± 1.2 and 41.9 ± 1.0 (Table 1A). The low temperature transition is attributed to an intermolecular conversion from a double-stranded structure, called a bulged duplex, to two (identical) single-stranded hairpins carrying a loop of sequence L, depicted in Scheme 2.

Scheme 2. The main derivative peak at 38.5°C, fitted by two gaussian curves, does not depend on Bc concentration and thus melting corresponds to intramolecular folding/unfolding processes. The two intramolecular structures of Bc may correspond to two hairpins. Examination of the Bc sequence reveals that it contains a limited palindrome and three consecutive CA TA repeats complementary to the GTA T nucleotides, which allow several hairpin structures. These are divided in two families: a base pairing limited to the palindrome (Scheme 1A) defines the first one. The second one is characterized by more base pairs than those of the palindrome (Scheme 1B). Melting temperature experiments with mutated oligonucleotides were performed to determine the most probable structures of Bc.
Figure 1. (Left) Melting derivatives of Bc and related oligonucleotides. The absorbances were monitored at 260 nm, in 30 mM cacodylate buffer, pH 6.0, 100 mM KCl. Curve 1, \([\text{Bc}] = 0.8 \mu M\); curve 2, \([\text{Bc3}] = 1.1 \mu M\); curves 3 and 4, \([\text{H1}] = [\text{H2}] = 4.2 \mu M\), respectively. The fit of curves 1 and 2 by three gaussian peaks is shown as a full line and the individual gaussian peaks as dashed lines. (Right) Deletions relative to native Bc yielding the related oligonucleotides Bc1, Bc3, Bc5 and BcII, Bc6, Bc7, presented on Bc structures I and II respectively (top), and H1 and H2, Bc2, drawn on Bc hairpins 1 and 2 respectively (bottom). The double arrows represent base pair permutations. This schematic representation does not intend to show the structures of the mutated oligomers.

Table 1A. Melting temperatures of Bc and Bnc as a function of concentration

<table>
<thead>
<tr>
<th>Oligo</th>
<th>[Oligo] (µM)</th>
<th>(T_m) (±0.5°C)</th>
<th>(\Delta H_{VH}) (kcal/mol)</th>
<th>(\Delta S_{VH}) (cal/mol K)</th>
<th>Deconvolution into three gaussian peaks ((\chi^2))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bc(\text{ex})</td>
<td>0.31</td>
<td>38.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.85</td>
<td>38.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.3</td>
<td>38.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.2</td>
<td>38.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9.0</td>
<td>38.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>18.1</td>
<td>38.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>53.5</td>
<td>39.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bc(\text{b,d})</td>
<td>1.3</td>
<td>47.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bnc(\text{c,e})</td>
<td>1.1</td>
<td>41.0</td>
<td>(23.5 ± 0.7)</td>
<td>(74.8 ± 2.0)</td>
<td>35.6, 48.0 (2.4 × 10⁻⁹)</td>
</tr>
<tr>
<td></td>
<td>4.4</td>
<td>41.4</td>
<td>(23.8 ± 0.7)</td>
<td>(75.4 ± 2.0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>19.5</td>
<td>41.8</td>
<td>(23.2 ± 0.7)</td>
<td>(73.6 ± 2.0)</td>
<td></td>
</tr>
</tbody>
</table>

\(T_m\) of mutated sequences relative to Bc. Nine mutated sequences relative to wild-type Bc were designed, named BcI, BcII, H1, H2, Bc2, Bc3, Bc5, Bc6 and Bc7 (Fig. 1). BcI is an analogue of Bc, in which 3 bp from the second and third CATA repeats are permuted in order to selectively stabilize the eight-membered loop hairpin (structure I or 1). Similarly, BcII differs from the native Bc by permutations in the first and second repeats and only allows four-membered loop hairpin (structure II or 2) formation.
H1 and H2 derive from BcI and BcII by deletion of the extremities, thus modelling hairpins 1 and 2 respectively. Bc2 differs from hairpin 2 by deletion of the dangling ends (fragment 8–23 of Bc). Bc3, deleted in the 5′-end compared with Bc, corresponds to the binding site of NssBF (31). A21, which is one of the important residues for protein binding, is deleted in Bc5 compared with Bc (41). The deletion of A_{4}T_{7}T_{7} in Bc6 relative to Bc was designed since these residues are inserted within the second TATG motif, complementary to the CA TA repeats. Bc7 derives from Bc6 by deletion of the dangling 3′-end. The melting derivatives of the mutated sequences are compared with that of Bc in Figure 1 and Table 1B.

Table 1B. Melting temperatures of the mutated sequences BcI, BcII, Bc2, H1, H2, Bc3, Bc5, Bc6 and Bc7

<table>
<thead>
<tr>
<th>Sequence</th>
<th>[Oligo] (µM)</th>
<th>T_{m} (±0.5 °C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BcF-e</td>
<td>0.8</td>
<td>29.3</td>
</tr>
<tr>
<td>2.7</td>
<td>30.0</td>
<td></td>
</tr>
<tr>
<td>5.1</td>
<td>29.8</td>
<td></td>
</tr>
<tr>
<td>16.4</td>
<td>29.4</td>
<td></td>
</tr>
<tr>
<td>BcI-e</td>
<td>1.3</td>
<td>47.4</td>
</tr>
<tr>
<td>BcII-e</td>
<td>0.5</td>
<td>33.1</td>
</tr>
<tr>
<td>0.8</td>
<td>33.2</td>
<td></td>
</tr>
<tr>
<td>5.0</td>
<td>33.5</td>
<td></td>
</tr>
<tr>
<td>14.0</td>
<td>33.0</td>
<td></td>
</tr>
<tr>
<td>BcII-e</td>
<td>1.3</td>
<td>47.5</td>
</tr>
<tr>
<td>1.0</td>
<td>49.4</td>
<td></td>
</tr>
<tr>
<td>4.2</td>
<td>48.9</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>49.2</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>49.1</td>
<td></td>
</tr>
<tr>
<td>H1-e</td>
<td>1.0</td>
<td>34.8</td>
</tr>
<tr>
<td>4.2</td>
<td>34.0</td>
<td></td>
</tr>
<tr>
<td>9.4</td>
<td>34.1</td>
<td></td>
</tr>
<tr>
<td>34.0</td>
<td>33.8</td>
<td></td>
</tr>
<tr>
<td>H2-e</td>
<td>1.0</td>
<td>41.0</td>
</tr>
<tr>
<td>4.2</td>
<td>40.0</td>
<td></td>
</tr>
<tr>
<td>7.3</td>
<td>40.1</td>
<td></td>
</tr>
<tr>
<td>17.0</td>
<td>40.1</td>
<td></td>
</tr>
<tr>
<td>Bc3-bc</td>
<td>1.0</td>
<td>38.5</td>
</tr>
<tr>
<td>2.4</td>
<td>38.6</td>
<td></td>
</tr>
<tr>
<td>9.6</td>
<td>38.9</td>
<td></td>
</tr>
<tr>
<td>23.0</td>
<td>38.7</td>
<td></td>
</tr>
<tr>
<td>Bc3-bd</td>
<td>1.3</td>
<td>50.8</td>
</tr>
<tr>
<td>Bc5-e</td>
<td>1.1</td>
<td>36.43</td>
</tr>
<tr>
<td>7.0</td>
<td>36.44</td>
<td></td>
</tr>
<tr>
<td>Bc6-e</td>
<td>0.3</td>
<td>59.8</td>
</tr>
<tr>
<td>1.3</td>
<td>59.6</td>
<td></td>
</tr>
<tr>
<td>5.7</td>
<td>60.3</td>
<td></td>
</tr>
<tr>
<td>17.4</td>
<td>59.4</td>
<td></td>
</tr>
<tr>
<td>Bc6-d-e</td>
<td>1.3</td>
<td>66.0</td>
</tr>
<tr>
<td>0.5</td>
<td>62.3</td>
<td></td>
</tr>
<tr>
<td>1.3</td>
<td>62.2</td>
<td></td>
</tr>
<tr>
<td>5.0</td>
<td>62.7</td>
<td></td>
</tr>
</tbody>
</table>

*a* Calculated according to a two-state model.

*b* The data were obtained using a heating rate of 6 °C/h.

*c* 30 mM cacodylate buffer, pH 6.0, 100 mM KCl.

*d* 50 mM cacodylate buffer, pH 6.0, [MgCl₂] = 10 mM.

*e* The data were obtained using a heating rate of 12 °C/h.

The Bc3 melting curve derivative, presented in Figure 1, curve 2, is as broad as that of Bc, with a main peak at 38.6 ± 0.5 °C and a shoulder around 42–44 °C. The Bc3 derivative can be deconvoluted into two concentration-independent peaks at 37.2 ± 1.1 and 42.4 ± 0.9 °C, as for Bc, together with a concentration-dependent gaussian curve at lower temperature. Bc3 differs from Bc by a 5′-end deletion that does not allow base pairing in the lower stem. The similarity in T_{m} values of Bc and Bc3 supports the assumption that the sum of the BcI and BcII derivatives should be lower than that of Bc.

Bc1, BcII, H1 and H2 present a single concentration-independent transition. Their melting temperatures are listed in Table 1B. The single concentration-independent peaks of these hairpins contrast with the two peaks of Bc. This result is anticipated from the design of Bc1, BcII, H1 and H2, each of which is expected to mimic a single hairpin (Fig. 1).

We now hypothesize that, at low concentration, Bc mainly adopts both hairpins 1 and 2, as suggested by the deconvolution and Bc3 results (Fig. 1). This assumption implies that the sum of BcI and BcII melting derivatives should mimic that of Bc. However, the base permutations within the mutated sequences could alter their stability compared with hairpins 1 and 2 (below). Figure 2A shows that the sum of the BcI and BcII derivatives can model that of Bc, assuming a 5 °C relative translation (the use of the sum of BcI and BcII derives from the averaged weight of each gaussian curve obtained by deconvolution). Data at higher Bc concentrations can be similarly fitted by the sum of BcI and BcII provided a 5 °C relative translation is applied (data not shown).

Bc2 melts at 7 °C higher than Bc hairpin 2. The same 7 °C gap exists between H2 and BcII and yet these hairpins differ by an identical deletion of the dangling ends. Thus the melting temperature difference between Bc2 and Bc hairpin 2 is likely to arise from the dangling ends deletion.

H2 melting (compared with hairpin 2) probably results from a balance between a decreased stability due to base permutations and an increased stability resulting from the dangling ends deletion. Thus the H2 T_{m} at 40.3 ± 0.5 °C fits Bc hairpin 2, peaking at 41.9 ± 1.0 °C (Fig. 1, curve 4). The same reasoning holds for hairpin 1: deletion of the dangling ends in H1 increases the T_{m} by 4 °C compared with BcI and the base pairs permutations in H1 relative to hairpin 1 are probably destabilizing. Thus, the H1 T_{m} at 34.5 ± 0.5 °C is quite close to the calculated T_{m} of hairpin 1 at 37.2 ± 1.2 °C (Fig. 1, curve 3). Additional evidence for the validity of the analysis is given by the correct fit of the Bc derivative by the sum of H1 and H2 (data not shown).

The derivative of the Bc5 melting curve exhibits two broad peaks, both independent of Bc5 concentration, reflecting the existence of two hairpin conformers. These data emphasize the structural flexibility of Bc5 and, indirectly, that of native Bc.

Bc6 and Bc7 melting profiles exhibit more cooperativity compared with that of Bc. The concentration-independent transition of Bc6 is 23 °C higher than that of Bc. Only hairpins of family B could be changed by this A_{4}T_{7}T_{7} deletion: hairpins I or II would yield a larger contiguous base paired stem in Bc6.
Mg$^{2+}$, which is not observed at 100 mM KCl, points to a change (Table 1). This difference between the two sequences at 10 mM Mg$^{2+}$; this assumption was cross-checked using probing experiments.

The probing experiments presented below were performed to confirm this hypothesis.

**Effect of Mg$^{2+}$ on Bc, BcI, BcII and Bc3 melting.** The structures adopted by Bc may depend on the ionic strength; i.e. Mg$^{2+}$ could stabilize the extended structures of family B. Melting data for Bc, BcI and BcII are shown in Figure 2B. For BcI and BcII the main transition is observed at 47.5 \(\pm\) 0.7 \(^\circ\)C, whereas the Bc derivative peaks at 47.6 \(\pm\) 0.7 \(^\circ\)C. These derivatives all present a more pronounced low lying transition than at 100 mM KCl. The sum of BcI and BcII reproduces the \(T_m\) of the Bc main and low transitions without any relative translation (Discussion).

The \(T_m\) of native Bc is 47.6 \(^\circ\)C, whereas the \(T_m\) of Bc3 is 50.8 \(^\circ\)C (Table 1). This difference between the two sequences at 10 mM Mg$^{2+}$, which is not observed at 100 mM KCl, points to a change in Bc structure. The lower \(T_m\) value of Bc could be compatible with the lower stem and bulge formation, which is avoided in Bc3 due to its 5'-end deletion relative to Bc. The energetic cost of the bulge is expected to be large, which may be consistent with the 18 \(^\circ\)C difference between the \(T_m\) of Bc and Bc6.

The \(T_m\) data strongly suggest that Bc could adopt both hairpins I and II at 10 mM Mg$^{2+}$; this assumption was cross-checked using probing experiments.

**Probing experiments**

Probing of Bc at 100 mM KCl, pH 6.0. Probing of Bc was performed using the chemicals diethylpyrocarbonate (DEPC) and potassium permanganate as well as the enzyme nuclease S1. These experiments were done under the same ionic conditions as used in the \(T_m\) studies. Typical gels are presented in Figure 3 and quantified in Figure 4, comparing folded Bc at 100 mM KCl (lanes 3 and 6) with unstructured Bc in the presence of 0.5 mM EDTA (lanes 2 and 5). Independent \(T_m\) experiments confirmed that Bc is unstructured at 20 \(^\circ\)C under the latter ionic conditions (\(T_m\) = 6 \(^\circ\)C; data not shown). Nucleotides G9, A11, G13 and A23 are strongly protected against DEPC cleavage (Fig. 4A). Residues A14, G15, A16 and A17 display enhanced reactivity to DEPC (compare lanes 2 and 3). In contrast, adenine A21 is partially protected against DEPC, whereas A25 and A27 are somewhat more reactive than A23. Residues A3 and A4 have similar reactivity in folded and unfolded Bc. In the folded molecule, T1, T2 and T5–T7 display enhanced permanganate cleavage. T8 is partially protected, probably due to its location in the stem of the structure. The single-stranded nature of the 5'-part of Bc is confirmed by efficient cleavage by nuclease S1 at G1 (Fig. 3C and data not shown).

From these data, two structures for Bc can be proposed, depicted in Scheme 1A: a hairpin with a 6 bp stem and an 8 nt loop (hairpin 1) and a structure with a similar stem but a 4 nt loop (hairpin 2). The results show that these structures are in equilibrium, indicated by the partial protection of A21, A25 and A27 to a greater extent and strong protection of the region G9–G13. The structural flexibility of Bc probably occurs via slippage of one extremity by one CATA repeat, relative to the other end, which allows, on average, the formation of 6 bp restricted to the upper stem. This slippage is probably a dynamic process. The data cannot exclude transient contributions of
alternative structures involving a more extended base pairing, like hairpins I and II. The data obtained at pH 6.0, 100 mM KCl, strongly suggest that Bc involves a dynamic equilibrium between two hairpins of family A. The question then arises, could Mg$^{2+}$ stabilize extended structures of family B?

**Probing of Bc in the presence of 10 mM MgCl$_2$ at pH 6.0.** A comparison of Bc probing by DEPC and KMnO$_4$ at 100 mM KCl and 10 mM Mg$^{2+}$ is shown in Figure 4. These results differ from those obtained at 100 mM KCl in the 5'-end of the sequence. Cleavage of nucleotide T2 is very weak compared with that at T5 and T6 (Fig. 4B), in contrast to that obtained at 100 mM KCl, where all the thymine residues T1–T7 are equally cleaved. Similarly, A3 and A4 are both efficiently cleaved at 100 mM KCl, whereas A3 is somewhat protected compared with A4 at 10 mM MgCl$_2$ (Figs 3C and D and 4A).

The data in Figure 4 support the existence of a lower stem and a bulge at 10 mM Mg$^{2+}$, pH 6.0. The partial protection of A19, T20 and A21 at 20°C is concomitant with a strong protection of A23, A25 and A27. The latter residues are indeed located in the upper stem of hairpin I and in the lower stem of hairpin II. These data strongly suggest that, in the presence of magnesium ions at pH 6.0, Bc adopts two predominant structures in equilibrium, hairpins I and II, depicted in Scheme 1B.

**Probing Bnc in the presence of 10 mM Mg$^{2+}$.** Typical probing experiments of the Bc complementary strand, Bnc, using permanganate, are shown in Figure 5A. These experiments test whether Bnc could adopt a single hairpin or two hairpins in equilibrium, similarly to Bc. The enhanced cleavage at residues T23, T24 and T26, compared with the control lane, is most obvious. The 5'-end residues T7, T9 and T11 follow a similar trend. In contrast, T13, T15, T17, T19, T21 and T29 are relatively protected. From these data, a single structure of Bnc can be proposed, depicted in Figure 5B. This hairpin carries a four-membered loop, shown by the strong modification at T23, T24 and T26. The Bnc stem is defined by marked protection of the base paired nucleotides T15, T19, T21 and T29. The less pronounced protection of T13 and T17 relative to T15, T19, T21 and T29 is consistent with their location, being the first base pair of the lower and upper stems, respectively.

**Probing Bc loops by macrocyclic bis-acridine 1 fluorescence**

The macrocycle bis-acridine 1 selectively binds to model hairpins compared with other DNA structures such as a double helix or
Figure 4. (A) Bc cleavage by DEPC in 10 mM Mg$^{2+}$ or in 100 mM KCl. Quantification of the data of Figure 3 and DEPC probing performed in the presence of Mg$^{2+}$ at 20°C (data not shown). The ordinates are a white to black scale; 0 and 256 pixels correspond to 100% white and 100% black respectively. The abscissa shows the position of a given nucleotide from the first G at the 5'-end. Filled and open circles, data obtained at 100 mM KCl and 10 mM MgCl$_2$ respectively. One representative blank at pH 6.0, 0.5 mM EDTA is shown as filled squares. (B) Comparison of Bc modification with KMnO$_4$ in the presence of Mg$^{2+}$ and 100 mM KCl. Quantification of the probing data of Figure 3 (100 mM KCl) and modifications performed in 10 mM MgCl$_2$ for 15 min, filled and open diamonds respectively.

'unstructured' single-stranded oligomers (38). Furthermore, this compound presents interesting fluorescence properties that discriminate between single-stranded and double-stranded oligomers (38,39). The macrocycle is thus expected to bind preferentially to loop sites as compared with other single-stranded subunits and should allow probing of the structural flexibility of Bc by comparison with mutated hairpins. The fluorescence spectra of the bis-acridine 1 bound to Bc, BcI and BcII are shown in Figure 6 (a 1:1 Bc–macrocycle complex has been shown by independent measurements of the stoichiometry; data not shown). The broad and unstructured maximum of free 1 is modified in the presence of Bc: the spectra present two vibronic bands, the higher one peaking at 442 nm, in agreement with a single-stranded, adenine-rich environment (38). The macrocycle bound to either Bc3 or Bc5 presents the same fluorescence yield as in the presence of Bc (Φ/Φ$_0$ = 1.40 ± 0.04). In contrast, the spectrum of 1 bound to BcII displays a higher intensity than that to Bc (Φ/Φ$_0$ = 1.75 ± 0.06), but comparable within experimental error to that of Bc6 and Bc7 (Φ/Φ$_0$ = 1.80 ± 0.06 and 1.85 ± 0.06, respectively). The relative yield of 1 complexed to BcI or H1 is lower than that of Bc (Φ/Φ$_0$ = 0.94 ± 0.06). These data are obtained only at low concentrations of BcII, Bc6 and Bc7 (0.4–1.0 µM); at higher concentrations (2.5–9.0 µM), the spectra present typical shoulders at 472 nm, in agreement with the bulged duplex structures adopted by these sequences (38,39; data not shown). The similar yield of the first 1 equivalent, whether bound to either BcI or H1, shows that the macrocycle first binds to the loop, the only common single-stranded site on these hairpins. Why should the fluorescence yields in the presence of Bc be different from those with BcII, Bc6, Bc7, BcI and H1? BcII is designed as a mimic of hairpin II, whereas BcI models hairpin I. The apparent discrepancy between 1 relative fluorescence yield on Bc and on the mutated sequences is explained if the macrocycle binds and probes both 4 and 8 nt hairpins loops. Indeed, Φ/Φ$_0$ (Bc) = 1.40 is intermediate between Φ/Φ$_0$ (BcI) = 0.94 and Φ/Φ$_0$ (BcII) = 1.75 and the macrocycle spectrum bound to Bc can be modeled by the sum of 1 complexed to BcI and BcII (Fig. 6).

DISCUSSION

Bc adopts two hairpin structures at pH 6.0

The $T_m$ of Bc is 38.5 ± 0.5°C. The corresponding peak derivative is concentration-independent in the range 0.3–53.5 µM, which indicates intramolecular folding (Table 1). Bc migrates as a single band, faster than the duplex formed by Bc and its complementary strand; this holds in the range [Bc] = 10$^{-10}$–10$^{-6}$ M (data not shown). Thus Bc folding is intramolecular. There is evidence of a distinct intermolecular transition at high concentrations or in the...
presence of magnesium ions (Fig. 2B and Table 1), indicating that Bc does form a bulged duplex under certain experimental conditions. Bc3 bulged duplex always melts at a lower temperature than that of Bc, consistent with the difference in sequence lengths (Fig. 1 and data not shown). Bc6 and Bc7, which possess the longest contiguous stem, form a bulged duplex at low strand concentration. The thermodynamic parameters of the Bc7 and Bc duplex to hairpin interconversions were calculated from the $T_m$ concentration dependence: $\Delta H_{VH}(Bc7) = 68 \pm 8 \text{ kcal/mol}$ and $\Delta S_{VH}(Bc7) = 196 \pm 25 \text{ eu}$; $\Delta H_{VH}(Bc) = 86 \pm 6 \text{ kcal/mol}$ and $\Delta S_{VH}(Bc) = 262 \pm 15 \text{ eu}$ (the latter are calculated from the deconvoluted low transition as a function of concentration; Table 1).

The $T_m$ curve derivative of Bc shows a main concentration-independent peak at 38.5°C, together with a shoulder around 42–44°C (Fig. 1, curve 1). This hints that more than one hairpin exists. Bc3, which lacks the 5′ extremity compared with Bc, gives similar results (Fig. 1, curve 2). Both profiles are best fitted by three gaussian curves, two of which are obtained at the same temperature, 37.2 and 42°C, for both sequences. Thus we hypothesize that, at 100 mM KCl and low concentration, Bc mainly adopts two hairpins of family A among the possible structures depicted in Scheme 1. BcI and BcII $T_m$ should model Bc melting if Bc adopts both hairpins 1 and 2. However, permutations of their stem compared with the native structures 1 and 2 could modify their $T_m$. [Calculation of the stem free energy difference due to permutations is only $\Delta G = 0.1 \text{ kcal/mol}$ (40), but the loop free energy may also differ.] Indeed, comparison of the Bc2 and H2 $T_m$ points to a 9°C decrease owing to base permutations. This is consistent with the need for a relative shift in the sum of the BcI and BcII curves to model the Bc derivative (Fig. 2A). Thus the Bc sequence is likely to adopt mainly hairpins 1 and 2 in equal proportions according to the averaged weight of the two (concentration-independent) gaussian curves. These structures slowly interconvert, probably on the timescale of seconds to minutes, in order to account for the very slow heating rates needed for fully reversible profiles.

Further supporting evidence for the two Bc conformers is given by the partial protection from DEPC cleavage of A21, consistent with an equilibrium between 4 and 8 nt loop hairpins. Moreover, the strong protection of A23 is expected, since this base is located in the stem of the two structures of Scheme 1A. The equilibrium probably involves slippage of four bases, corresponding to one CATA repeat. Both probing and melting results at 100 mM KCl support the structural flexibility of Bc that is suggested from inspection of its sequence.

### Probing Bc loops with the macrocyclic bis-acridine 1

The fluorescence maximum of the bound macrocycle to Bc is characteristic of a single-stranded, adenine-rich site (38). Probing of Bc by the macrocycle can be deduced from the fluorescence yields obtained with Bc compared with mutated sequences. The same 1 fluorescence yield, measured in the presence of BcI and H1, implies that the first macrocycle binds to the only common single-stranded site of these sequences: the loop. The fluorescence spectrum of 1 on Bc can be modelled by the sum of the bound macrocycle to BcI and BcII. This shows that the macrocycle binds to both 4 and 8 nt loops of the native hairpins, in agreement with the equilibrium between two Bc hairpins determined by the melting and probing experiments. The lower fluorescence yield of bound 1 to BcI relative to that observed with BcII probably reflects less base stacking in the 4 nt loop than in the 8 nt one. A correlation between the macrocycle fluorescence yield and the extent of base stacking has been previously noted to account for the higher yield of 1 bound to an A3 hairpin loop relative to that on the trimer A3 (38). Therefore, the variation in the macrocycle fluorescence yields allows probing of the structural difference between the hairpin I and II.
loops, suggesting that macrocycle 1 could be a structural probe for complex sequences bearing loops subunits.

Existence of a bulge in Bc

The melting curves of Bc6 and Bc7 compared with that of Bc first suggest the possible formation of bulged hairpins. These structures (Scheme 1B) indeed exist, stabilized by Mg$^{2+}$, shown by a partial protection of A3 and a strong one of T2. These two residues are located in the lower stem. Extensive cleavage of the bulge residues T5 and T6 is also observed. The equilibrium between hairpins I and II is suggested by partial protection of residues A19, T20 and A21 combined with strong protection of A23, A25 and A27. The melting temperature results, showing a great similarity between Bc and Bc + BcII derivative profiles (Fig. 2B) support this attribution. The lower $T_m$ of Bc compared with that of Bc3 is indeed consistent with a high energetic cost of the bulge.

Bc hairpins I and II possess a four-membered bulge. The occurrence of such a bulge is frequent in RNA hairpins. Bulges induce bending of DNA and RNA duplexes (41–49) in a direction away from the bulge-containing strand, providing an opening in the opposite groove. Bulge-induced kinks have been suggested to facilitate recognition of the opposite strand sequence by cognate proteins. The larger recognition site of NssBF, (TACA)$_3$, lies in the 3′-part of Bc, precisely opposite the bulge (31,32). Moreover, Bc3 does not bind NssBF, although it corresponds to the binding site of this protein on Bc (31,32). Deletion of the 5′-end in Bc3 relative to Bc does not allow formation of a bulge and a lower stem. This modification of Bc3 overall structure compared with that of Bc might be responsible for the lack of NssBF binding on Bc3.

Bnc adopts a single hairpin structure

The Bc complementary strand, Bnc, also forms a hairpin, which melts at a higher temperature than Bc ($T_m = 41.4 \pm 1.0^\circ$C) (Table 1A). This may result from the change in the stem closing pair from GC in Bc to CG in Bnc or from a difference in adenine/thymine content of their loops (50 and references therein). Probing data indicate that Bnc adopts a single hairpin structure bearing a four-membered loop TCTT, shown by the similar and enhanced cleavage pattern of T23, T24 and T26 (Fig. 5). The dangling 3′-end residues, T7, T9 and T11, follow a similar trend. Protection of T15, T17, T21 and T29 indicates the stem residues. The results show that Bnc does not possess the structural flexibility of its complementary sequence.

Biological relevance of Bc flexibility to the 1731 retrotransposon

1731 is a mobile retrovirus-like element, framed by two LTRs of U3-R-U5 architecture, both being unidirectional promoters (31–34). The Bc sequence is located in the U3 region of both LTRs, between positions −110 and −73 in the 5′ LTR, thus just upstream of the transcription start (+1). The structural flexibility of Bc may be associated with its function. The regulatory transcription factors NssBF and p11 only bind to the Bc strand (31–33). DNA hairpins have been shown to be involved in transcription and replication processes (21–24) and in topoisomerase II recognition and cleavage (25–28). Moreover, one of the main cleavage sites of topoisomerase II is precisely located in the Bc sequence within the whole LTR (34), indicating that this enzyme is probably required for Bc extrusion and function. Thus Bc slippage suggests that structural and topological parameters contribute to modulate the rate of transcription and transposition of 1731 retrotransposon (51).

CONCLUSION

The binding site of the transcription factor NssBF, named Bc, mainly adopts two hairpin structures in slow equilibrium at pH 6.0. The predominant structures of Bc, derived from probing experiments, consist of 4 and 8 nt loop hairpins, but these techniques cannot discount minor or transient contributions of alternative structures. Mutated hairpins carrying 4 and 8 nt loops can mimic its melting profile. A macrocyclic bis-acridine has been used to probe the structural flexibility of Bc compared with 4 and 8 nt loop model hairpins. The slow equilibrium between these hairpins probably involves slippage of 4 nt, originating from the presence of three CATA repeats complementary to the GTAT nucleotides in the Bc sequence.

The presence of magnesium ions affects Bc structures probed at pH 6.0. Chemical probing shows modifications at the lower stem level, consistent with two bulged hairpins in equilibrium. A comparison of Bc melting with that of a mutated sequence lacking the 5′-end is also consistent with lower stem and bulge formation in the presence of magnesium ions.

In contrast, the Bc complementary strand adopts a single hairpin structure carrying a 4 nt loop. Two transcriptional factors bind only to the Bc strand within the whole retrotransposon 1731 LTR, suggesting that Bc structural flexibility could be associated with transcription.

ACKNOWLEDGEMENTS

The authors thank Prof. Lehn and Drs Vigneron and Teulade-Fichou for the gift of the macrocyclic bis-acridine. Dr Ouali and Prof. Westhof are acknowledged for helpful discussions, Drs Buckingham and Hui Bon Hoa for their critical reading of the manuscript and T. Bataille for the drawings. This work was supported by INTAS grant 95-IN-RU-653.

REFERENCES


