Recognition of GC base pairs by triplex forming oligonucleotides containing nucleosides derived from 2-aminopyridine

Sarah A. Cassidy, Peter Slickers1,+, John O. Trent1, Daniel C. Capaldi2, Peter D. Roselt2, Colin B. Reese2, Stephen Neidle1 and Keith R. Fox*

Division of Biochemistry and Molecular Biology, School of Biological Sciences, University of Southampton, Bassett Crescent East, Southampton SO16 7PX, UK, 1CRC Biomolecular Structure Unit, The Institute for Cancer Research, Cotswold Road, Sutton, Surrey SM2 5NG, UK and 2Department of Chemistry, King’s College London, Strand, London WC2R 2LS, UK

Received September 19, 1997; Revised and Accepted November 3, 1997

ABSTRACT

We have attempted to alleviate the pH dependency of triplex recognition of guanine by using intermolecular triplexes containing 2-aminomethylene-5-(2-deoxy-β-D-ribofuranosyl)pyridine (AP) as an analogue of 2-deoxycytidine (dC). We find that for the β-anomer of AP, the complex between (AP)6T6 and the target site G6A6T6C6 is stable, generating a clear DNase I footprint at oligonucleotide concentrations as low as 0.25 µM at pH 5.0, in contrast to 50 µM C6T6 which has no effect on the cleavage pattern. This complex is still stable at pH 6.5 producing a footprint with 1 µM oligonucleotide. Oligonucleotides containing the α-anomer of AP are much less effective than the β-anomer, though in some instances they are more stable than the unmodified oligonucleotides. The results of molecular dynamics studies on a range of AP-containing triplexes has rationalized the observed stability behaviour in terms of hydrogen-bonding behaviour.

INTRODUCTION

The formation of DNA triple helices offers the possibility of designing sequence specific DNA binding agents, which may have therapeutic and experimental uses (1–4). The third strand oligonucleotide binds in the major groove of polypurine tracks of duplex DNA and is held in place by specific hydrogen bonds to substituents on the bases (1,2). Two types of triple helix have been characterised which differ according to the orientation of the third strand. Pyrimidine-rich oligonucleotides bind parallel to the duplex purine strand forming T•AT and C•GC triplets (5–7) while purine-rich oligonucleotides bind in an anti-parallel orientation and are characterised by G•GC, A•AT and T•AT triplets (8–11).

One limitation to the formation of parallel (Y•RY) triplexes is that conditions of low pH are necessary for protonation of the third strand cytosine in the C•GC triplet. The free base has a pKₐ of ~4.35, though this may be elevated within triplex forming oligonucleotides which may be stable at a pH up to 6.0, depending on the number and location of the cytosine residues. Contiguous C•GC triplets are especially unstable (12), presumably on account of the proximity of the charged bases. A number of cytosine analogues and mimics have been synthesised in attempts to overcome this restriction. 5-Methylcytosine, which has a slightly higher pKₐ value than cytosine (13,14) generates triplexes which are more stable at a higher pH, but are still not formed under physiological conditions (12). This increase in stability may be a consequence of the extra spine of methyl groups within the DNA major groove (15) rather than from changes in the pKₐ. A carbocyclic analogue of 5-methylcytosine also has a higher pKₐ than cytosine (4.80 instead of 4.35) and has been shown to enhance triplex formation at elevated pHs (16). 6-Oxocytosine (17), pseudocytosine (18) and 8-oxoadenine (19,20) have also been shown to recognize GC base pairs in a pH-independent fashion, though they not have been widely used in triple-forming oligonucleotides. Other promising dC analogues are β- and α-amino-5-(2-deoxy-β-D-ribofuranosyl)pyridines (β- and α-AP) (Fig. 1B and C) which have pKₐs of 5.93 and 6.16 respectively (21). Psoralen-linked oligonucleotides containing these modifications have been successfully used to target a portion of the aromatase gene (21). These initial studies with AP-containing oligonucleotides suggested that both β- and α-anomers of this nucleoside analogue could form stable triplets. The possibility of forming complexes with α-anomers was especially interesting since α-oligonucleotides are more resistant to serum nucleases. However these studies used psoralen-linked oligonucleotides which form DNA cross-links after UV-irradiation and might promote the formation of less stable complexes. Initial modelling studies also confirmed the possibility of forming a triplet with α-AP, though this consisted of a single α-AP•GC triplet within a block of T•AT.

*To whom correspondence should be addressed. Tel: +44 1703 594374; Fax: +44 1703 594459; Email:krf1@soton.ac.uk

*Visiting scientist from The Institute of Molecular Biotechnology, Beutenbergstrasse II, D-07745 Jena, Germany
triplets. Recent studies have also shown that oligonucleotides containing AP and its 3-methyl derivative bind with higher affinity to duplex DNA targets between pH 6 and 8 than 5-methylcytosine containing oligonucleotides (22,23). This difference was especially pronounced for sequences containing contiguous cytosines.

In this paper we use DNase I footprinting to examine triple helix formation at the sequence G6A6•T6C6 using a series of modified oligonucleotides based on C6T6, in which the cytosine are replaced with 2-aminopyridine residues. Oligonucleotides containing either the β- or α-anomers of 2-AP were tested. We have previously shown that unmodified CT-containing oligonucleotides show little interaction with this target site, even at pH 5.5. We find that oligonucleotides containing the β-anomer form stable complexes at higher pH than cytosine-containing oligonucleotides. We also report on results from more extensive molecular modelling studies on sequences analogous to those examined by footprinting.

**MATERIALS AND METHODS**

**Oligonucleotides**

C6T6 was purchased from Oswel DNA Service. Modified oligonucleotides containing 2′-deoxy-5-methylcytidine (MeC) and α- and β-AP were synthesized as previously reported (21) from the appropriate phosphoramidite building blocks in a Applied Biosystems 381A DNA synthesizer. The amino functions in the α- and β-AP derived phosphoramidites were protected with dichloroacetyl groups, and dichloroacetic anhydride was used instead of acetic anhydride in the capping steps. The abbreviations used for these oligonucleotides, which were designed to interact with the target site G6A6•T6C6, are shown in Figure 1D. The integrity of the oligonucleotides was confirmed at various times by labelling with [γ-32P]ATP using polynucleotide kinase, and examining the products on denaturing polyacrylamide gels. In each case >90% of the material ran as the intact 12mers. The naphthylquinoline triplex-binding ligand (24–26) was a generous gift from Dr L. Strekowski, Department of Chemistry, Georgia State University.

**Plasmids and DNA fragments**

The preparation of plasmid pGA1 has been previously described (27). This contains the sequence G6A6•T6C6 cloned into the BamH1 site of pUC19. A radiolabelled restriction fragment containing this sequence was prepared as previously reported by labelling with [α-32P]ATP using reverse transcriptase and cutting again with EcoRI. This procedure labels the pyrimidine-containing strand of the triplex target site. The fragment of interest was resolved from the remainder of the plasmid DNA on an 8% non-denaturing polyacrylamide gel. This was eluted from the gel slice, precipitated the remainder of the plasmid DNA on an 8% non-denaturing polyacrylamide gel. This was eluted from the gel slice, precipitated with 0.1 mM EDTA.

**DNase I footprinting**

DNase I footprinting was performed as previously described (24–26). An aliquot (1.5 µl) of labelled DNA (∼10 nM, dissolved in 10 mM Tris, pH 7.5 containing 0.1 mM EDTA) was mixed with 1.5 µl oligonucleotide, dissolved in an appropriate buffer as indicated in the text. All buffers contained 5 mM MgCl2, necessary to stabilise triple helix formation. This mixture was left to equilibrate at room temperature for at least 1 h before adding 2 µl DNase I (0.1 U/ml dissolved in 20 mM NaCl, 2 mM MgCl2, 2 mM MgCl2). The reaction was stopped after 1 min by adding 4 µl of formamide containing 10 mM EDTA and 0.1% (w/v) bromophenol blue. Samples were boiled for 3 min and cooled rapidly on ice before loading onto 11% polyacrylamide gels containing 8 M urea. Gels of 40 cm length were run at 1500 V for about 2 h, then fixed in 10% (v/v) acetic acid, transferred to Whatmann 3MM paper and dried at 80°C. Dried gels were exposed to autoradiography at −70°C using an intensifying screen. Bands in the autoradiographs were assigned by comparison with Maxam–Gilbert marker lanes specific for G+A.

**Unrestrained molecular dynamics simulations**

Initial oligonucleotide models were constructed with G6A6•T6C6 duplexes in an idealised B-DNA conformation. The bases of the third strand were placed in the major groove of the duplex, connected together with sugar-phosphate backbones, and then energy minimised in vacuo with a distance dependent dielectric constant and with distance restraints to force Hoogsteen hydrogen bonding of the bases. The five dodecamer triplex models constructed had third strand of sequence β6T6 α6T6 αβ3T6 β3α7T6 and (αβ)3T6.

---

**Figure 1.** (A) Proposed hydrogen bonding schemes between a protonated 2-aminopyridine (AP) nucleoside and a GC base pair (dR, deoxyribose). (B) The β-anomer of AP [2-amino-5-(2-deoxy-β-D-ribofuranosyl)pyridine]. (C) The α-anomer of AP. (D) Sequences of the oligonucleotides used in this work. β, β-anomer of AP; α, α-anomer of AP.
Molecular dynamics (MD) simulations were performed with the AMBER 4.1 programs and force-field (30). Each triple helix was solvated in a box sized $65 \times 45 \times 45$ Å containing 98 Na$^+$ ions, 71 Cl$^-$ ions and 3898 water molecules. This is approximately equivalent to a NaCl concentration of 0.9 M. A cut-off radius of 10 Å was used for van der Waals interactions. The particle mesh Ewald approach was used to model electrostatic interactions (31). Atom-centred partial charges for $\alpha$- and $\beta$-2-aminopyridine nucleosides were determined with MOPAC/ESP (29). Missing force-field parameters were chosen in analogy to existing ones. MD simulations were run with constant temperature (300 K), constant pressure (1 bar) and a time step of 2 fs with SHAKE applied to all bond lengths. All ions were placed at least 5 Å away from any triplex atoms. The MD protocol involved an initial short minimisation (100 steps), followed by 2000 steps of MD at 1 K with DNA atoms fixed to allow the water and ions to relax, followed by 2000 steps with all atoms free to move. Then the temperature was raised linearly from 1 to 300 K within 20 ps. During this initial phase distance restraints were applied to the hydrogen bonds of the terminal base triplets. The systems were further equilibrated for 20 ps under production conditions. The production runs involved 200 ps simulations without any restraints.

Coordinates were saved every 2 ps, and averaged structures were calculated for the last 50 ps of each run. Several properties including the distances between hydrogens and acceptors of the hydrogen bonds between first and third strand were plotted over time. Hydrogen bonds were classified as established, if the distance between hydrogen and acceptor was lower than 2.5 Å. By analogy to pyrimidine nucleotides we define the torsion angle $\chi$ about the glycosidic bond of AP nucleotides by the atoms O4', C1', C5 and C6.

RESULTS

Comparison of $\alpha$-AP, $\beta$-AP, dC and 5-Me-C

We have examined the ability of 2-AP to act as a cytosine mimic for recognising GC base pairs in a pH-independent fashion by studying triple helix formation at the target site $G_6A_6\cdot T_6C_6$. We chose this target site since we have previously shown that the cytosine-containing oligonucleotide $C_5T_5$ does not form a particularly stable triplex, presumably as a result of the run of contiguous C$^+\cdot$GC triplets. This triplex only forms at relatively high oligonucleotide concentrations, at low pH (<5.5) and low temperatures (4°C). Figure 2A shows DNase I digestion of a fragment containing this target site, and compares the binding of the various AP-containing oligonucleotides with $C_6T_6$ and 5MeC6T6 at pH 5.0. It can be seen that, at this third strand
the binding of the $\alpha$-anomers and the unmodified oligonucleotides. The results of this experiment are shown in Figure 3A. It can be seen that, although $C_5T_6$ does not affect the digestion pattern, all the oligonucleotides containing $5\text{MeC}$, $\alpha$-AP or $\beta$-AP generate clear footprints at the target site. From these results it appears that AP-modified oligonucleotides containing a mixture of $\alpha$- and $\beta$-nucleosides [i.e. $\alpha_6$, $\alpha(\beta_3)\alpha_3\beta_3$ and $\beta_\alpha\gamma_\delta$] can also bind to this target site and that their binding is stronger than the unmodified oligonucleotide. Previous studies with TC-containing oligonucleotides containing all $\alpha$-anomeric residues have shown that these bind in the opposite orientation to the natural $\alpha$-anomers (32). Since $T*\cdot\alpha\beta$ triplets can form in either parallel or anti-parallel configuration it is possible that the correct target site for $\alpha_6$ should be $A_6G_6\cdot\alpha\betaC_6T_6$ (generating an anti-parallel structure) rather than $G_6A_6\cdot\alpha\betaT_6C_6$. We tested this possibility by examining the interaction of 50 $\mu$M of the modified oligonucleotides with a fragment containing the target site $A_6G_6\cdot\alpha\betaC_6T_6$. No interaction of any of the oligonucleotides with this fragment could be detected at pH 5.0 (not shown).

Previous studies have shown that the binding of similar CT-containing oligonucleotides is enhanced by decreasing the temperature to 4°C (28). We have therefore repeated these experiments at 4°C, in order to further compare the stability of the various AP-modified oligonucleotides; the results are shown in Figure 3B. It can be seen that under these conditions even 10 $\mu$M of the unmodified $C_5T_6$ produces a DNase I footprint at the target site. Footprints are also generated with $5\text{MeC}_5C_6T_6$, $\alpha\beta\gamma\delta$ and $\beta\alpha\gamma\delta$. However the binding of $\alpha_6$ and $\alpha(\beta_3)$ appears to be weaker, and cleavage products are still evident within the target site. These observations, which are considered in more detail in the Discussion, suggest that the $\beta$-anomer of AP is more effective than the $\alpha$-anomer.

Interaction with $\beta_6$

The results presented above clearly show that $\beta_6$ generates the best complex and that this is still stable at pH 6.5, even though it generates a block of six contiguous AP•GC triplets. We have investigated the binding of $\beta_6$ to $G_6A_6\cdot\alpha\betaC_6T_6$ in more detail by examining concentration dependence of the footprints at pH 5.0 and 6.5. The results are presented in Figure 5. It can be seen that at pH 5.0, $\beta_6$ produces a footprint which persists to concentrations as low as 0.25 $\mu$M (Fig. 4A). Since 50 $\mu$M of the cytosine-containing oligonucleotide shows no interaction under the same conditions (Fig. 3A), it appears that the unmodified oligonucleotide enhances the oligonucleotide binding affinity by >200-fold. A similar experiment at pH 6.5 is presented in Figure 4B. In this case, $\beta_6$ generates a DNase I footprints down to a concentration of 1 $\mu$M. This represents an increase in stability (compared to the unmodified oligonucleotide at pH 5.0) of at least 50-fold.

Interaction with a naphthylquinoline triplex-binding ligand

Several studies have shown that a naphthylquinoline triplex-binding ligand can enhance the stability of complexes containing $T\cdot\alpha\beta$ triplets, reducing the oligonucleotide concentration required to generate a footprint by as much as 200-fold. We were therefore interested to see whether this compound could stabilise the complexes formed with oligonucleotides containing $\alpha$- and $\beta$-AP. Figure 5A examines the effect of 10 $\mu$M of this compound on the interaction of 10 $\mu$M AP-modified oligonucleotides with $G_6A_6\cdot\alpha\betaC_6T_6$, at pH 5.0. In the absence of the ligand, the only

![Figure 3.](image-url)
oligonucleotide which shows any interaction at 10 µM is β₆ (Fig. 2). In the presence of the ligand clear footprints are evident with the oligonucleotides containing cytosine and 5-methylcytosine residues and for β₆, (αβ)₃ and α₃β₃. No interaction is detected with α₆ or β₃α₃, supporting the previous observation that the β-anomers are more effective than the α-anomers. Since the complexes generated with AP-containing oligonucleotides should be less pH dependent than those involving cytosine or 5-methylcytosine residues, we have investigated whether the ligand could also stabilise these complexes at a higher pH. The results are presented in Figure 5B, showing the effect of 10 µM naphthylquinoline compound on the binding of 10 µM oligonucleotides to G₆A₆•T₆C₆ at pH 6.5. It can be seen that only β₆ generates a clear footprint. The binding of this oligonucleotide is not surprising since 10 µM β₆ can itself generate a footprint at pH 6.5 (Fig. 2). The ligand is unable to induce triplex formation with the remaining oligonucleotides, though some weak interaction with α₆β₃ may be indicated by a reduction in cleavage intensity around the triplex target site.

The results of these footprinting experiments are summarised in Table 1.

**Figure 4.** Interaction of oligonucleotide β₆ with a DNA fragment containing the triplex target site G₆A₆•T₆C₆. Oligonucleotide concentrations (µM) are shown at the top of each gel lane. Reactions were performed at 20°C in 50 mM sodium acetate pH 5.0 containing 5 mM MgCl₂ (A) or 20 mM PIPES pH 6.5 containing 10 mM NaCl and 5 mM MgCl₂ (B). The brackets show the position and length of the triplex target site. Con indicates digestion of the DNA in the absence of added oligonucleotide. The track labelled GA is a Maxam–Gilbert marker specific for G+A.

**Figure 5.** Effect of 10 µM naphthylquinoline triplex-binding ligand on the interaction between the AP-containing oligonucleotides (10 µM) and a fragment containing the target site G₆A₆•T₆C₆. The experiments were performed in (A) 50 mM sodium acetate pH 5.0 containing 5 mM MgCl₂ or (B) 20 mM PIPES pH 6.5 containing 10 mM NaCl and 5 mM MgCl₂. Con indicates digestion of the DNA in the absence of added oligonucleotide. The bracket shows the position and length of the target site. The tracks labelled GA are Maxam–Gilbert marker lanes specific for G+A.

**Molecular modelling**

The MD simulations resulted in stable structures for all five α/β combinations generated. (A full analysis of the structures in terms of their energetics and conformations will be presented elsewhere.) None of the structures showed any significant distortions from acceptable geometric features. All are B-type helices with an average helical twist of 30–32°, and an average rise of 3.4–3.5 Å. Stereo plots of the averaged structure of β₆T₆ and α₆T₆ are presented in Figures 6 and 7, respectively.

Two types of hydrogen bonding schemes are found between AP bases and guanine, canonical Hoogsteen hydrogen bonds and a bifurcated arrangement, where both H1 and H21 of AP are hydrogen-bonded to guanine O6. These two types are illustrated in Figure 1A. The 5’-terminal triplets are only loosely hydrogen bonded and do open during the simulation in all but the β₆α₂β₁T₁ complex. The α₆T₁₆ triplex has the least number of triplets in a full Hoogsteen arrangement. It is notable that the α/β interface in the α₃β₃T₆ and β₃α₂T₆ triplexes did not involve any destabilisation of hydrogen bonding, with full Hoogsteen arrangements for triplet 3 as well as triplet 4. On the other hand, triplet 6 at the interface between
Table 1. Summary of the binding of AP-modified oligonucleotides to the G6A6•T6C6 target site under a range of experimental conditions

<table>
<thead>
<tr>
<th>Experimental conditions</th>
<th>Oligonucleotide</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C6</td>
</tr>
<tr>
<td>10μM oligonucleotide</td>
<td>×</td>
</tr>
<tr>
<td>pH 5.0-6.5</td>
<td></td>
</tr>
<tr>
<td>10μM oligonucleotide</td>
<td>×</td>
</tr>
<tr>
<td>pH 7.0-7.5</td>
<td></td>
</tr>
<tr>
<td>10μM oligonucleotide</td>
<td>✓</td>
</tr>
<tr>
<td>pH 5.0, 4°C</td>
<td></td>
</tr>
<tr>
<td>50μM oligonucleotide</td>
<td>×</td>
</tr>
<tr>
<td>pH 5.0</td>
<td></td>
</tr>
<tr>
<td>10μM oligonucleotide</td>
<td>✓</td>
</tr>
<tr>
<td>pH 5.0, 10μM 1</td>
<td></td>
</tr>
<tr>
<td>10μM oligonucleotide</td>
<td>×</td>
</tr>
<tr>
<td>pH 6.5, 10μM 1</td>
<td></td>
</tr>
</tbody>
</table>

Successful binding of the oligonucleotide is indicated by a tick (✓), while a cross (×) represents no evidence of binding. W, indicates a possible weak interaction.

AP and T adopts a bifurcated or even open arrangement in all complexes.

Figure 6. Stereo view of the averaged structure (150–200 ps) of the β6T6 triplex. The 5′-end of the third strand (shown in dark) is located at the top of the plot. H-bonds between third strand and first strand are indicated by thin lines. The second triple has a bifurcated arrangement of H-bonds while the other triplets have a Hoogsteen arrangement. Hydrogen atoms are shown for the bases only but not for riboses and not for methyl groups of thymine.

Table 2. Hydrogen-bonded characteristics of the 5-triplex models

<table>
<thead>
<tr>
<th>Triplet</th>
<th>α</th>
<th>β</th>
<th>αβ</th>
<th>βαβ</th>
<th>(αβ)2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>open</td>
<td>open</td>
<td>open</td>
<td>Hoogsteen</td>
<td>open</td>
</tr>
<tr>
<td>2</td>
<td>bifurcated</td>
<td>bifurcated</td>
<td>bifurcated</td>
<td>Hoogsteen</td>
<td>open</td>
</tr>
<tr>
<td>3</td>
<td>open</td>
<td>Hoogsteen</td>
<td>Hoogsteen</td>
<td>Hoogsteen</td>
<td>Hoogsteen</td>
</tr>
<tr>
<td>4</td>
<td>Hoogsteen</td>
<td>Hoogsteen</td>
<td>Hoogsteen</td>
<td>Hoogsteen</td>
<td>Hoogsteen</td>
</tr>
<tr>
<td>5</td>
<td>bifurcated</td>
<td>Hoogsteen</td>
<td>Hoogsteen</td>
<td>Hoogsteen</td>
<td>Hoogsteen</td>
</tr>
<tr>
<td>6</td>
<td>bifurcated</td>
<td>bifurcated</td>
<td>bifurcated</td>
<td>Hoogsteen</td>
<td>open</td>
</tr>
</tbody>
</table>

The sugar pucker phase angles of the α-AP residues in the triplex structures are found in the C1′-endo range (with pseudorotation phase angles from –72° to –36°), and the torsion angle χ about the glycosidic bond is between –60° and 0° (a syn conformation). The same conformational type was also found for the energy-minimised monomer. Hence this is the preferred conformation of α-nucleotides and not primarily the result of binding forces in the complex. The sugars of the β-AP residues in the triplexes adopt a C2′-endo pucker and an anti conformation about the glycosidic bond, although some repuckering to O4′-endo and C3′-endo was observed. The terminal 5′ β-AP residues show some transitions to a syn conformation around the glycosidic bond.

The intrastrand C1′-C1′ distance of two successive nucleotides was found to be ~5 Å, between β nucleotides. This is the case with the first, second and third strand, although it changes at α/β steps. This is summarized in Table 2. In 5′-α/β steps the C1′-C1′ distance is increased to ~6 Å, while it decreases to <4.5 Å at 5′-β/α steps. Accordingly, in the (αβ)3T6 complex the alternating 5′-α/β and 5′-β/α steps compensate each other and the average C1′-C1′ distance remains at 5.3 Å.
DISCUSSION

Cytosine versus $\beta$-AP

The results presented in this paper demonstrate that substitution of dC by $\beta$-AP dramatically enhances the stability of parallel triplexes. Unmodified C\(\beta\)T\(\beta\) shows no interaction with G\(\alpha\)A\(\alpha\)T\(\alpha\)C\(\alpha\) at pH 5.0, even at a concentration of 50 \(\mu\)M. However, replacement of the dC with $\beta$-AP residues allows triplex formation at oligonucleotide concentrations as low as 0.25 \(\mu\)M, representing an increase in binding affinity of >200-fold. Moreover, the AP nucleoside extends the recognition of GC base pairs to higher pH, thereby relieving the pH-dependency of the parallel binding motif. In contrast to the unmodified oligonucleotide, the $\beta$-AP-containing oligonucleotide shows significant interaction at pH 6.5 at which clear DNase I footprints are generated at oligonucleotide concentrations as low as 1 \(\mu\)M. Triplex formation is extended to neutral pH with 50 \(\mu\)M $\beta$e. The position of the $\beta_e$ footprint, terminating exactly at the edge of the triplex target site, provides evidence that the $\beta$-aminopyridines are actively contributing to the integrity of the complex and that the footprint is not merely caused by a strong interaction of the T\(\alpha\)A\(\alpha\)T triplets.

The $\beta$-AP nucleoside differs from a number of other nucleoside analogues which have been designed to recognise GC base pairs via reverse-Hoogsteen hydrogen bonds (32). One of the first modifications of cytosine involved its methylation at the C-5 position (13,14). Although this derivative has a slightly higher pK$_{\alpha}$, the enhancement in triplex stability probably results from favourable hydrophobic effects generated by the formation of a spire of methyl groups in the major groove (15). Our results are consistent with these reports, showing that the binding of C\(\beta\)T\(\beta\) is increased when the cytosine residues are replaced with $^{5\alpha}$C. Substitution with $\beta$-aminopyridine imparts even greater stability than $^{5\alpha}$C: the complex with 10 \(\mu\)M $\beta$-AP is stable up to pH 6.5, while 10 \(\mu\)M $^{5\alpha}$C-oligonucleotide does not generate a DNase I footprint at any pH. These results demonstrate that the binding of cytosine to GC base pairs is best improved by enhancing the degree of N-3 protonation (by increasing the pK$_{\alpha}$), rather than attaching a hydrophobic methyl group to the C-5 position.

5-Methylcytosine versus 2-aminopyridine

One of the first modifications of cytosine involved its methylation at the C-5 position (13,14). Although this derivative has a slightly higher pK$_{\alpha}$, the enhancement in triplex stability probably results from favourable hydrophobic effects generated by the formation of a spire of methyl groups in the major groove (15). Our results are consistent with these reports, showing that the binding of C\(\beta\)T\(\beta\) is increased when the cytosine residues are replaced with $^{5\alpha}$C. Substitution with $\beta$-aminopyridine imparts even greater stability than $^{5\alpha}$C: the complex with 10 \(\mu\)M $\beta$-AP is stable up to pH 6.5, while 10 \(\mu\)M $^{5\alpha}$C-oligonucleotide does not generate a DNase I footprint at any pH. These results demonstrate that the binding of cytosine to GC base pairs is best improved by enhancing the degree of N-3 protonation (by increasing the pK$_{\alpha}$), rather than attaching a hydrophobic methyl group to the C-5 position.

Anomeric configuration

The susceptibility of natural $\beta$-oligonucleotides to nuclease attack may be overcome by using synthetic $\alpha$-anomers (34,35). Many studies have compared the binding of oligonucleotides consisting of entirely $\alpha$- or $\beta$-nucleotides. Due to the change in anomeric configuration, the orientation of an $\alpha$-oligonucleotide is expected to be reversed in comparison to $\beta$-oligomers. However oligonucleotides consisting entirely of $\beta$-T are reported to bind in the same parallel orientation as the $\beta$-T\(\alpha\) oligonucleotide, presumably via reverse-Hoogsteen hydrogen bonds (32).

Bates et al. (21) were the first group to investigate the triplex binding of oligonucleotides containing a mixture of $\alpha$- and $\beta$-anomers. They reported that oligonucleotides containing a mixture of $\beta$-thymidine and $\alpha$-AP residues bind in a parallel orientation with respect to the duplex purine strand. Our results confirm this orientation since oligonucleotides containing various combinations of $\alpha$- and $\beta$-aminopyridines bind to G\(\alpha\)A\(\alpha\)T\(\alpha\)C\(\alpha\) but show no interaction with A\(\alpha\)G\(\alpha\)C\(\alpha\)T\(\alpha\)C\(\alpha\). Molecular modelling studies (21) suggested that $\alpha$-anomers of AP can be accommodated within an otherwise $\beta$-anomeric triplex, with minimal structural perturbation. The results presented in this paper demonstrate that oligonucleotides containing $\alpha$-AP form less stable parallel triplexes than $\beta$-AP. Comparing the activity of $\alpha_e$ and $\beta_e$, we find that the $\beta$-AP-oligonucleotide is consistently more stable than the corresponding $\alpha$-AP-oligonucleotide.

The modelling results indicate that the $\alpha$-AP residues are as readily accommodated within a parallel triplex as the isomorphous $\beta$-nucleosides. The constitutional modification at C 1’ is compensated by changes of the sugar puckering phase angle and the torsion about the glycosidic bond. (We have to bear in mind that the $\alpha$-anomer is not a mirror image of the $\beta$-anomer, since the ribose ring has several chiral centres.) The preferred sugar conformation of the $\alpha$-AP residues is C 1’-endo and the torsion angle about the glycosidic bond is in the negative syn range. This compensates for the altered configuration at the C 1’ atom and allows one to integrate $\alpha$-nucleotides into a parallel triple helix. Previous studies have reported that the binding of $\alpha$-oligonucleotides to duplex DNA is fairly weak, requiring the additional attachment of a stabilising acridine group to the end of the oligonucleotide (32,36). We can infer from our modelling studies that this is primarily caused by steric restrictions. The major differences between the five simulated structures are seen in the behaviour of the hydrogen bonds between first strand guanine and third strand AP bases. The number of triplets which have a canonical Hoogsteen arrangement of hydrogen bonds approximately accords with the relative stability found by footprinting (Table 1). It should, however, be borne in mind that the 5-fold differences in concentrations correspond to only a small difference in binding free energy.

Oligonucleotides including a mixture of $\alpha$-AP and $\beta$-AP nucleosides [i.e. ($\alpha$\(\beta\))$\alpha$ ($\alpha$\(\beta\))$\alpha$ ($\beta$\(\alpha\))$\alpha$] bind more strongly than $\alpha$\(\alpha\) under some conditions. These differences are most pronounced under triplex-stabilising conditions. At 4°C, $\alpha$\(\beta\)$\alpha$ and $\alpha$\(\beta\)$\alpha$ show strong interaction at 10 \(\mu\)M, with a weaker interaction detected for $\alpha$\(\alpha\) and ($\alpha$\(\beta\))$\alpha$. Furthermore, at pH 5.0 the naphthylquinoline triplex-binding ligand induces the binding of ($\alpha$\(\beta\))$\alpha$ and $\alpha$\(\beta\)$\alpha$ but not $\alpha$\(\alpha\) and $\beta$\(\alpha\)$\alpha$. In general it seems that structures containing a run of only three $\alpha$-AP residues are tolerated more readily than complexes which include a run of six $\alpha$-nucleosides. This supports the proposal that introduction of several consecutive $\alpha$-AP residues into an otherwise $\beta$-anomeric triplex induces some destabilisation. The higher stability of $\alpha$\(\beta\)$\alpha$ than $\beta$\(\alpha\)$\alpha$ can be attributed to the presence of nine contiguous $\alpha$-residues in the former.

We have previously shown that the triplex-binding ligand enhances the stability of 10mer TC-containing oligonucleotides by >100-fold. However the compound was unable to shift triplex formation to physiological pH, presumably due to diminished N-3 protonation on the third strand cytosine residues. Although this ligand induces the binding of ($\alpha$\(\beta\))$\alpha$ and $\alpha$\(\beta\)$\alpha$ at pH 5.0 it does not potentiate the binding of ($\alpha$\(\beta\))$\alpha$ at pH 6.5, though some weak interaction is detected with $\alpha$\(\beta\)$\alpha$. In other words stabilisation of the six T\(\alpha\)A\(\alpha\)T triplets is insufficient to facilitate formation of the adjoining stretch of ($\alpha$-AP)$\alpha$GC triplets at pH 6.5.
ACKNOWLEDGEMENTS

This work was supported by grants from the Cancer Research Campaign (to S.N., C.B.R. and K.R.F.) and the Medical Research Council (K.R.F.).

REFERENCES

29 Stewart, J.P.P., Mopac 6.0 (QXPE), available from the Quantum Chemistry Program Exchange, Indiana University, Bloomington, IN 47405, USA.