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

Sarah A. Cassidy, Peter Slickers[1,+], John O. Trent[1], Daniel C. Capaldi[2], Peter D. Roselt[2], Colin B. Reese[2], Stephen Neidle[1] and Keith R. Fox*

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

ABSTRACT

We have attempted to alleviate the pH dependency of triplex recognition of guanine by using intermolecular triplexes containing 2-amino-5-(2-deoxy-D-ribofuranosyl)pyridine (AP) as an analogue of 2′-deoxycytidine (dC). We find that for the β-anomer of AP, the complex between $(AP)_6T_6$ and the target site $G_6A_6 \cdot T_6C_6$ 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 C_6T_6 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 pK_a 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 pK_a 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 pK_a. A carbocyclic analogue of 5-methylcytosine also has a higher pK_a than cytosine (4.80 instead of 4.35) and has been shown to enhance triplex formation at elevated pHs (16). 6-Oxocytosine (17), pseudoisocytosine (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-D-ribofuranosyl)pyridines (β- and α-AP) (Fig. 1B and C) which have pK_as 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

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 $G_6A_6 \bullet T_6C_6$ using a series of modified oligonucleotides based on C_6T_6, 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

C_6T_6 was purchased from Oswel DNA Service. Modified oligonucleotides containing 2′-deoxy-5-methylcytidine (^{Me}C) 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 $G_6A_6 \bullet T_6C_6$, are shown in Figure 1D. The integrity of the oligonucleotides was confirmed at various times by labelling with [γ-^{32}P]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 $G_6A_6 \bullet T_6C_6$ cloned into the *Bam*H1 site of pUC19. A radiolabelled restriction fragment containing this sequence was prepared as previously described by cutting the plasmid with *Hin*dIII, labelling at the 3′-end with [α-^{32}P]dATP using reverse transcriptase and cutting again with *Eco*R1. 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 with ethanol and redissolved in 10 mM Tris–HCl, pH 7.5 containing 0.1 mM EDTA.

DNase I footprinting

DNase I footprinting was performed as previously described (24,26–28). 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 MgCl$_2$, 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 MgCl$_2$, 2 mM MgCl$_2$). 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 $G_6A_6 \bullet T_6C_6$ 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 $β_6T_6$, $α_6T_6$, $α_3β_3T_6$, $β_3α_3T_6$ and $(αβ)_3T_6$.
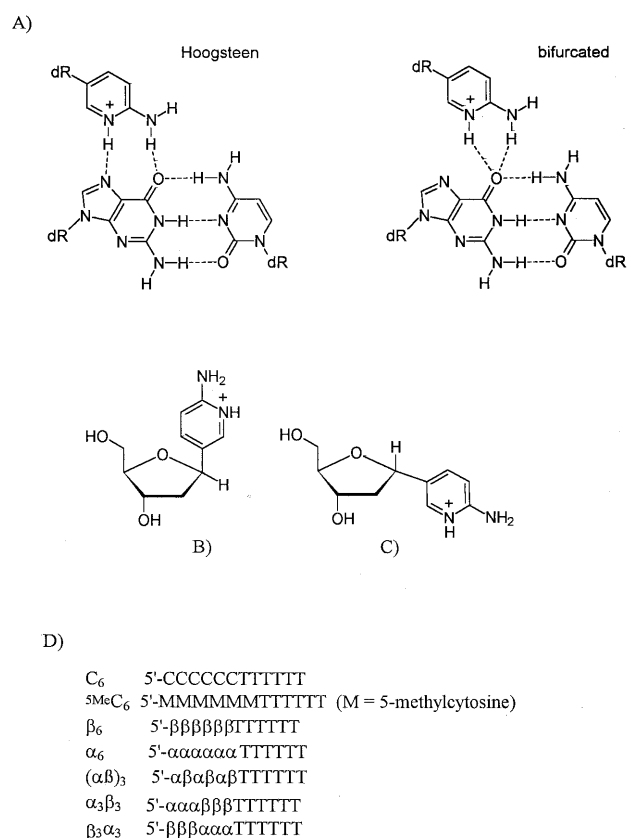

Figure 1. (**A**) Proposed hydrogen bonding schemes between a protonated 2-aminopyridine (AP) nucleoside and a GC base pair (dR, deoxyribose). (**B**) The β-anomer of AP [2-amino-5-(2-deoxy-D-ribofuranosyl)pyridine]. (**C**) The α-anomer of AP. (**D**) Sequences of the oligonucleotides used in this work. 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 α- and β-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 χ about the glycosidic bond of AP nucleotides by the atoms O4', C1', C5 and C6.

RESULTS

Comparison of α-AP, β-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 $G_6A_6 \bullet T_6C_6$. 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 $C^+ \bullet GC$ 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 $^{5Me}C_6T_6$ at pH 5.0. It can be seen that, at this third strand
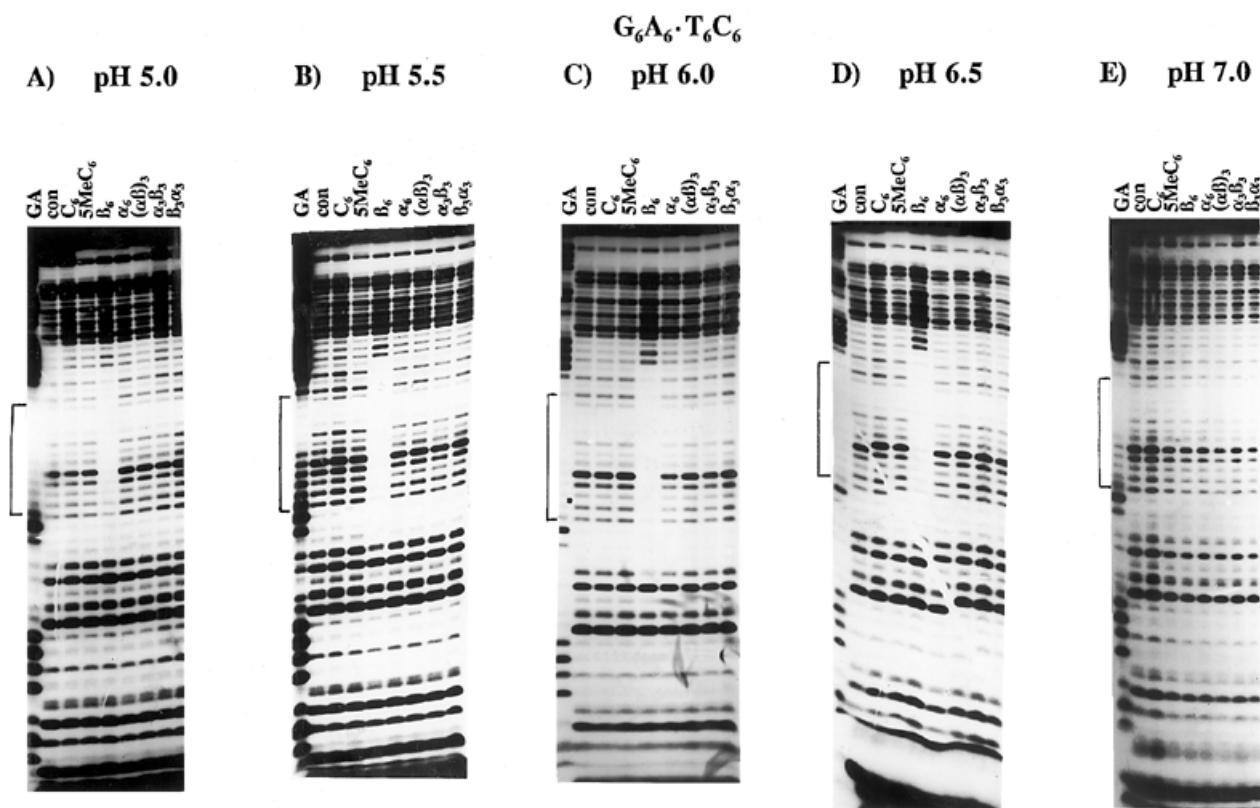
$$G_6A_6 \cdot T_6C_6$$

Figure 2. DNase I footprinting patterns for the interaction of various oligonucleotides with the fragment containing the target site $G_6A_6 \bullet T_6C_6$. The DNA was labelled at the 3′-end of the *Hin*dIII site, visualising the pyrimidine-containing strand (T_6C_6) of the target site. In each case the oligonucleotide concentration was 10 µM. Reactions were performed at 20°C. The track labelled con shows digestion of the DNA in the absence of added oligonucleotide. GA corresponds to a Maxam–Gilbert marker lane specific for guanine and adenine. The square bracket shows the position and length of the target site. All experiments contained 5 mM $MgCl_2$. The buffers were: (**A**) 50 mM sodium acetate pH 5.0; (**B**) 50 mM sodium acetate pH 5.5; (**C**) 20 mM PIPES containing 10 mM NaCl pH 6.0; (**D**) 20 mM PIPES pH 6.5 containing 10 mM NaCl and (**E**) 10 mM Tris–HCl pH 7.0 containing 10 mM NaCl. All buffers contained 5 mM $MgCl_2$. The oligonucleotide is indicated at the top of each gel lane.
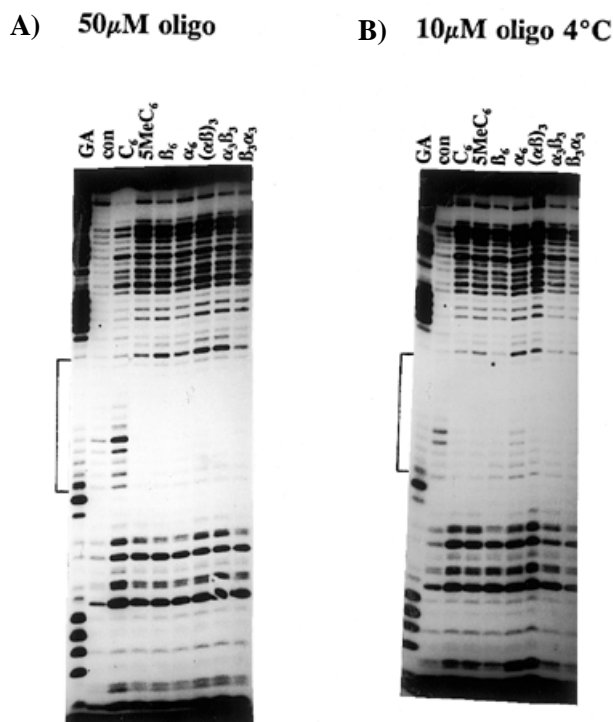
$G_6A_6 \cdot T_6C_6$

A) 50 μM oligo

B) 10 μM oligo 4°C

Figure 3. (**A**) DNase I digestion of a fragment containing the insert $G_6A_6 \bullet T_6C_6$ in the presence and absence of 50 μM AP-containing oligonucleotides. (**B**) DNaseI footprints showing the interaction between 10 μM AP-containing oligonucleotides and the target site $G_6A_6 \bullet T_6C_6$, measured at 4°C. Experiments were performed in 50 mM sodium acetate pH 5.0. The square brackets show the position and length of the target site (T_6C_6). Con indicates digestion of the DNA in the absence of added oligonucleotide. In both cases this control is slightly underdigested relative to the other lanes. The track labelled GA is a Maxam–Gilbert marker specific for G+A.

concentration (10 μM), β_6 is the only oligonucleotide which shows any interaction, producing a clear footprint covering the entire triplex target site. Neither the C_6T_6 or $^{5Me}C_6T_6$ show any interaction with the duplex target. In order to examine whether the interaction with β_6 extends to higher pH this experiment was repeated at pH 5.5 (Fig. 2B), pH 6.0 (Fig. 2C), pH 6.5 (Fig. 2D) and pH 7.0 (Fig. 2E). Clear footprints with 10 μM β_6 are evident up to pH 6.5, though no protection is evident at pH 7.0. Further studies with higher concentrations of β_6 (50 μM, not shown) reveal clear footprints at pH 7.0, but not pH 7.5. These results clearly demonstrate that replacement of dC by β-AP residues dramatically enhances the oligonucleotide binding affinity at pH 5.0 and stabilises the complex at elevated pH. Furthermore this stabilisation is much greater than that achieved with 5-methylcytosine. It should be noted that the experiments performed at pH 6.0 and above used lower ionic strength buffers. Since triplex stability decreases at lower ionic strengths these experiments will underestimate the affinities at higher pHs.

We repeated these experiments using a higher oligonucleotide concentration (50 μM) at the same pH (5.0) in attempt to compare the binding of the α-anomers and the unmodified oligonucleotides. The results of this experiment are shown in Figure 3A . It can be seen that, although C_6T_6 does not affect the digestion pattern, all the oligonucleotides containing ^{5Me}C, α-AP or β-AP generate clear footprints at the target site. From these results it appears that AP-modified oligonucleotides containing a mixture of α- and β-nucleosides [i.e. α_6, $(\alpha\beta)_3$, $\alpha_3\beta_3$ and $\beta_3\alpha_3$] can also bind to this target site and that their binding is stronger than the unmodified oligonucleotide. Previous studies with TC-containing oligo-nucleotides containing all α-anomeric residues have shown that these bind in the opposite orientation to the natural β-anomers (32). Since T•AT triplets can form in either parallel or anti-parallel configuration it is possible that the correct target site for α_6 should be $A_6G_6 \bullet C_6T_6$ (generating an anti-parallel structure) rather than $G_6A_6 \bullet T_6C_6$. We tested this possibility by examining the interaction of 50 μM of the modified oligonucleotides with a fragment containing the target site $A_6G_6 \bullet C_6T_6$. 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 μM of the unmodified C_6T_6 produces a DNase I footprint at the target site. Footprints are also generated with $^{5Me}C_6T_6$, β_6, $\alpha_3\beta_3$ and $\beta_3\alpha_3$. However the binding of α_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 β-anomer of AP is more effective than the α-anomer.

Interaction with β_6

The results presented above clearly show that β_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 β_6 to $G_6A_6 \bullet T_6C_6$ 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, β_6 produces a footprint which persists to concentrations as low as 0.25 μM (Fig. 4A). Since 50 μM of the cytosine-containing oligonucleotide shows no interaction under the same conditions (Fig. 3A), it appears that replacement of cytosine with β-AP enhances the oligonucleotide binding affinity by >200-fold. A similar experiment at pH 6.5 is presented in Figure 4B. In this case, β_6 generates a DNase I footprints down to a concentration of 1 μM. 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•AT 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 α- and β-AP. Figure 5A examines the effect of 10 μM of this compound on the interaction of 10 μM AP-modified oligonucleotides with $G_6A_6 \bullet T_6C_6$, at pH 5.0. In the absence of the ligand, the only
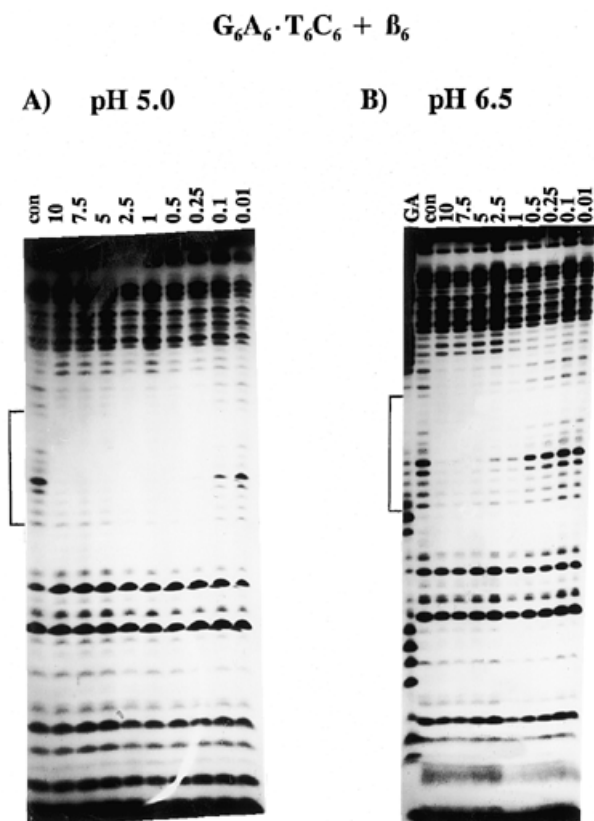
Figure 4. Interaction of oligonucleotide β_6 with a DNA fragment containing the triplex target site $G_6A_6 \cdot T_6C_6$. Oligonucleotide concentrations (μM) are shown at the top of each gel lane. Reactions were performed at 20°C in 50 mM sodium acetate pH 5.0 containing 5 mM $MgCl_2$ (**A**) or 20 mM PIPES pH 6.5 containing 10 mM NaCl and 5 mM $MgCl_2$ (**B**). The brackets shows 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 $G_6A_6 \cdot T_6C_6$. The experiments were performed in (**A**) 50 mM sodium acetate pH 5.0 containing 5 mM $MgCl_2$ or (**B**) 20 mM PIPES pH 6.5 containing 10 mM NaCl and 5 mM $MgCl_2$. 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.

oligonucleotide which shows any interaction at 10 μM is β_6 (Fig. 2). In the presence of the ligand clear footprints are evident with the oligonucleotides containing cytosine and 5-methylcytosine residues and for β_6, $(\alpha\beta)_3$ and $\alpha_3\beta_3$. No interaction is detected with α_6 or $\beta_3\alpha_3$, 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 $G_6A_6 \cdot T_6C_6$ at pH 6.5. It can be seen that only β_6 generates a clear footprint. The binding of this oligonucleotide is not surprising since 10 μM β_6 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 $\alpha_3\beta_3$ may be indicated by a reduction in cleavage intensity around the triplex target site.

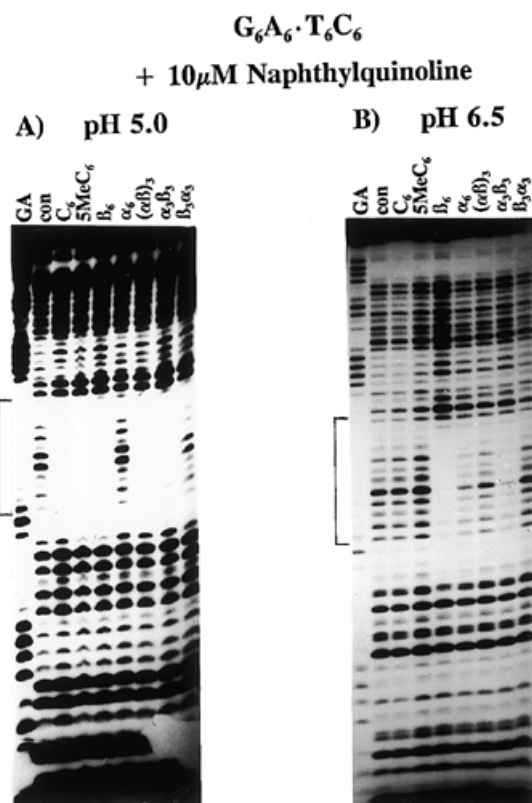The results of these footprinting experiments are summarised in Table 1.

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 β_6T_6 and α_6T_6 are presented in Figures 6 and 7, respectively.
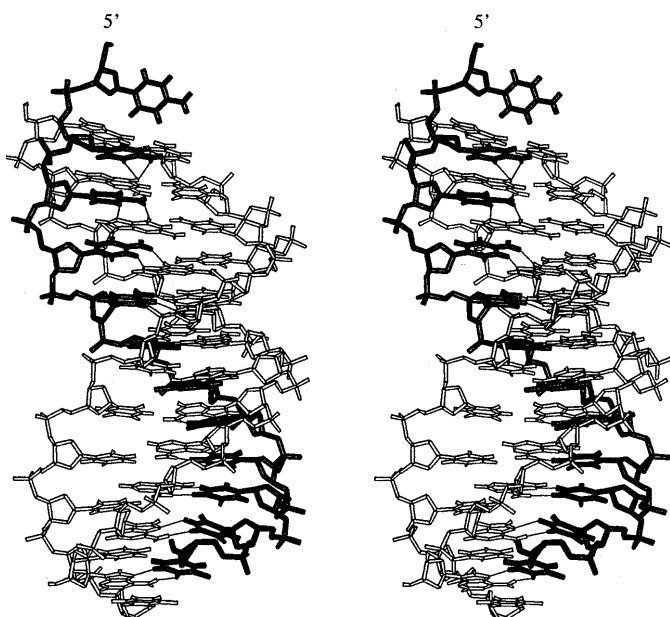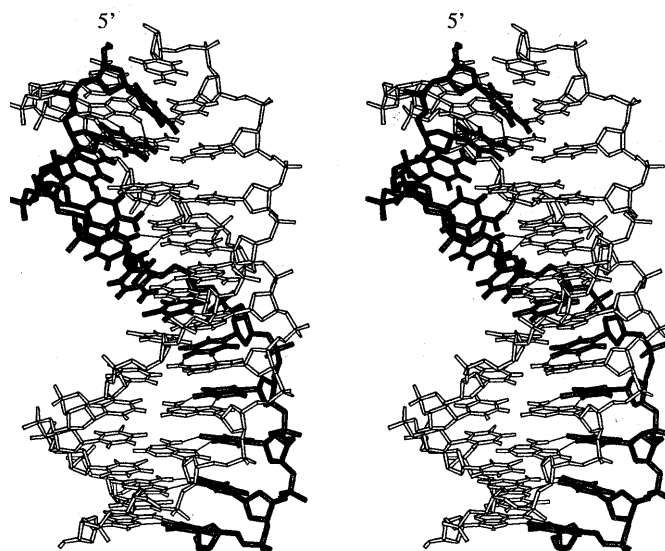
Two types of hydrogen bonding schemes are found between AP bases and guanine, canonical Hoogsteen hydrogen bonds and a bifurcated arrangement, where both H1 and H21 of AP are hydrogen-bonded to guanine O6. These two types are illustrated in Figure 1A. The 5′-terminal triplets are only loosely hydrogen bonded and do open during the simulation in all but the $\beta_3\alpha_3T_6$ complex. The α_6T_6 triplex has the least number of triplets in a full Hoogsteen arrangement. It is notable that the α/β interface in the $\alpha_3\beta_3T_6$ and $\beta_3\alpha_3T_6$ 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 $G_6A_6 \bullet T_6C_6$ target site under a range of experimental conditions

Experimental conditions	Oligonucleotide						
	C_6	$^{5Me}C_6$	β_6	α_6	$(\alpha\beta)_3$	$\alpha_3\beta_3$	$\beta_3\alpha_3$
10µM oligonucleotide pH 5.0-6.5	×	×	✓	×	×	×	×
10µM oligonucleotide pH 7.0-7.5	×	×	×	×	×	×	×
10µM oligonucleotide pH 5.0, 4°C	✓	✓	✓	W	W	✓	✓
50µM oligonucleotide pH 5.0	×	✓	✓	✓	✓	✓	✓
10µM oligonucleotide pH 5.0, 10µM 1	✓	✓	✓	×	✓	✓	W
10µM oligonucleotide pH 6.5, 10µM 1	×	×	✓	×	×	W	×

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 β_6T_6 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.

Figure 7. Stereo view of the averaged structure (150–200 ps) of the α_6T_6 triplex.

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 $(\alpha\beta)_3T_6$ 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.

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

Triplet	α_6	β_6	$\alpha_3\beta_3$	$\beta_3\alpha_3$	$(\alpha\beta)_3$
1	open	open	open	Hoogsteen	open
2	bifurcated	bifurcated	bifurcated	Hoogsteen	open
3	open	Hoogsteen	Hoogsteen	Hoogsteen	Hoogsteen
4	Hoogsteen	Hoogsteen	Hoogsteen	Hoogsteen	Hoogsteen
5	bifurcated	Hoogsteen	Hoogsteen	Hoogsteen	Hoogsteen
6	bifurcated	bifurcated	bifurcated	Hoogsteen	open

DISCUSSION

Cytosine versus β-AP

The results presented in this paper demonstrate that substitution of dC by β-AP dramatically enhances the stability of parallel triplexes. Unmodified C_6T_6 shows no interaction with $G_6A_6 \cdot T_6C_6$ at pH 5.0, even at a concentration of 50 μM. However, replacement of the dC with β-AP residues allows triplex formation at oligonucleotide concentrations as low as 0.25 μM, 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 β-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 μM. Triplex formation is extended to neutral pH with 50 μM $β_6$. The position of the $β_6$ footprint, terminating exactly at the edge of the triplex target site, provides evidence that the β-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•AT triplets.

The β-AP nucleoside differs from a number of other nucleoside analogues which have been designed to recognise GC base pairs such as 8-oxoadenine and derivatives (19,20) and a pyrazolo-pyrimidinone derivative (12), in that it is an isostructural analogue of cytosine. This ensures minimal structural distortion of the third strand backbone, allowing the formation of a continuous YYR triplex.

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_a, the enhancement in triplex stability probably results from favourable hydrophobic effects generated by the formation of a spine of methyl groups in the major groove (15). Our results are consistent with these reports, showing that the binding of C_6T_6 is increased when the cytosine residues are replaced with ^{5Me}C. Substitution with β-aminopyridine imparts even greater stability than ^{5Me}C: the complex with 10 μM β-AP is stable up to pH 6.5, while 10 μM ^{5Me}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_a), rather than attaching a hydrophobic methyl group to the C-5 position.

Anomeric configuration

The susceptibility of natural β-oligonucleotides to nuclease attack may be overcome by using synthetic α-anomers (34,35). Many studies have compared the binding of oligonucleotides consisting of entirely α- or β-nucleotides. Due to the change in anomeric configuration, the orientation of an α-oligonucleotide is expected to be reversed in comparison to β-oligomers. However oligo-nucleotides consisting entirely of α-T are reported to bind in the same parallel orientation as the β-d(T)$_n$ 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 α- and β-anomers. They reported that oligonucleotides containing a mixture of β-thymidine and α-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 α- and β-aminopyridines bind to $G_6A_6 \cdot T_6C_6$ but show no interaction with $A_6G_6 \cdot C_6T_6$. Molecular modelling studies (21) suggested that α-anomers of AP can be accommodated within an otherwise β-anomeric triplex, with minimal structural perturbation. The results presented in this paper demonstrate that oligonucleotides containing α-AP form less stable parallel triplexes than β-AP. Comparing the activity of $α_6$ and $β_6$, we find that the β-AP-oligonucleotide is consistently more stable than the corresponding α-AP-oligonucleotide.

The modelling results indicate that the α-AP residues are as readily accommodated within a parallel triplex as the isomorphous β-nucleosides. The constitutional modification at C1′ is compensated by changes of the sugar pucker phase angle and the torsion about the glycosidic bond. (We have to bear in mind that the α-anomer is not a mirror image of the β-anomer, since the ribose ring has several chiral centres.) The preferred sugar conformation of the α-AP residues is C1′-*endo* and the torsion angle about the glycosidic bond is in the negative *syn* range. This compensates for the altered configuration at the C1′ atom and allows one to integrate α-nucleotides into a parallel triple helix. Previous studies have reported that the binding of α-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 stabilities 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 α-AP and β-AP nucleosides [i.e. $(αβ)_3$, $α_3β_3$ and $β_3α_3$], bind more strongly than $α_6$ under some conditions. These differences are most pronounced under triplex-stabilising conditions. At 4°C, $α_3β_3$ and $β_3α_3$ show strong interaction at 10 μM, with a weaker interaction detected for $α_6$ and $(αβ)_3$. Furthermore, at pH 5.0 the naphthylquinoline triplex-binding ligand induces the binding of $(αβ)_3$ and $α_3β_3$ but not $α_6$ and $β_3α_3$. In general it seems that structures containing a run of only three α-AP residues are tolerated more readily than complexes which include a run of six α-nucleosides. This supports the proposal that introduction of several consecutive α-AP residues into an otherwise β-anomeric triplex induces some destabilisation. The higher stability of $α_3β_3$ than $β_3α_3$ can be attributed to the presence of nine contiguous β-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 $(αβ)_3$ and $α_3β_3$ at pH 5.0 it does not potentiate the binding of $(αβ)_3$ at pH 6.5, though some weak interaction is detected with $α_3β_3$. In other words stabilisation of the six T•AT triplets is insufficient to facilitate formation of the adjoining stretch of $(α-AP^+) \cdot GC$ triplets at pH 6.5.

REFERENCES

1 Soyfer, V.N. and Potaman, V.N. (1996) *Triple helical Nucleic Acids* Springer-Verlag, NY.
2 Thuong, N.T. and Hélène, C. (1993) *Angewandte Chemie*, **32**, 666–690.
3 Chubb, J.M. and Hogan, M.E. (1992) *Trends Biotechnol.*, **10**, 132–136.
4 Moffat, A.S. (1991) *Science*, **252**, 1374–1375.
5 Moser, H.E. and Dervan, P.B. (1987) *Science*, **238**, 645–650.
6 Le Doan, T., Perrouault, L., Praseuth, D., Habhoub, N., Decout, J.L., Thuong, N.T., Lhomme, J. and Hélène, C. (1987) *Nucleic Acids Res.*, **15**, 7749–7760.
7 Radhakrishnan, I. and Patel, D.J. (1994) *Structure*, **2**, 17–32.
8 Beal, P.A. and Dervan, P.B. (1991) *Science*, **251**, 1360–1363.
9 Beal, P.A. and Dervan, P.B. (1992) *Nucleic Acids Res.*, **20**, 2773–2776.
10 Radhakrishnan, I., de los Santos, C. and Patel, D.J. (1991) *J. Mol. Biol.*, **221**, 1403–1418.
11 Radhakrishnan, I. and Patel, D.J. (1993) *Structure*, **1**, 135–152.
12 Koh, J.S. and Dervan, P.B. (1992) *J. Am. Chem. Soc.*, **114**, 1470–1478.
13 Lee, J.S., Wordsworth, M.L., Latimer, L.J.P. and Morgan, A.R. (1984) *Nucleic Acids Res.*, **12**, 6603–6614.
14 Povsic, T.J. and Dervan, P.B. (1989) *J. Am. Chem. Soc.*, **111**, 3059–3061.
15 Xodo, L.E., Manzini, G., Quadifoglio, F., van der Marel, G. and van Boom, J. (1991) *Nucleic Acids Res.*, **19**, 5625–5631.
16 Froehler, B.C. and Ricca, D.J. (1992) *J. Am. Chem. Soc.*, **114**, 8320–8322.
17 Huang, C.-Y., Bi, G. and Miller, P.S. (1996) *Nucleic Acids Res.*, **24**, 2606–2613.
18 Ono, A., Ts'o, P.O.P. and Kan, L. (1991) *J. Am. Chem. Soc.*, **113**, 4032–4033.
19 Miller, P.S., Bhan, P. Cushman, C.D. and Trapane, T.L. (1992) *Biochemistry*, **31**, 2999–3004.
20 Jetter, M.C. and Hobbs, F.W. (1993) *Biochemistry*, **32**, 3249–3254.
21 Bates, P.J., Laughton, C.A., Jenkins, T.C., Capaldi, D.C., Roselt, P.D.,Reese, C.B. and Neidle, S. (1996) *Nucleic Acids Res.*, **24**, 4176–4184.
22 Hildbrand, S., Blaser, A., Parel, S.P. and Leumann, C.J. (1997) *J. Am. Chem. Soc.*, **119**, 5499–5511.
23 Hildbrand, S. and Leumann, C. (1996) *Angew. Chem. Int. Ed. Eng.*, **35**, 1968–1970.
24 Cassidy, S.A., Strekowski, L., Wilson, W.D. and Fox, K.R. (1994) *Biochemistry*, **33**, 15338–15347.
25 Wilson, W.D., Tanious, F.A., Mizan, S., Yao, S., Kiselyov, A.S., Zon, G. and Strekowski, L. (1993) *Biochemistry*, **32**, 10614–10621.
26 Chandler, S.P., Strekowski, L., Wilson, W.D. and Fox, K.R. (1995) *Biochemistry*, **34**, 7234–7242.
27 Fox, K.R. (1994) *Nucleic Acids Res.*, **22**, 2016–2021.
28 Stonehouse T.J. and Fox, K.R. (1994) *Biochim. Biophys. Acta*, **1218**, 322–330.
29 Stewart, J.P.P. Mopac 6.0 (QXPE), available from the Quantum Chemistry Program Exchange, Indiana University, Bloomington, IN 47405, USA.
30 Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, D.C., Fox, T., Caldwell, J.W., Kollman, P.A. (1995) *J. Am. Chem. Soc.*, **117**, 5179–5197.
31 Darden, T.A., York, D.M. and Pedersen, L.G. (1993) *J. Chem. Phys.*, **98**, 10089–10092.
32 Sun, C., Giovannangeli, C., Francois, J.C., Kurfurst, K., Monteny-Garestier, T., Asseline, U., Saison-Behmoaras, T., Thuong, N.T. and Hélène, C. (1991) *Proc. Natl. Acad. Sci. USA.*, **88**, 6023–6027.
33 Singleton, S.F. and Dervan, P.B. (1992) *Biochemistry*, **31**, 10995–11003.
34 Morvan, F., Rayner, B., Imbach, J.-L., Thenet, S., Bertrand, J.-R., Paoletti, J., Malvy, C. and Paoletti, C. (1987) *Nucleic Acids Res.*, **15**, 3421–3437.
35 Cazenave C., Chevrier, M., Thuong, N.T., and Hélène, C. (1987) *Nucleic Acids Res.*, **15**, 10507–10521.
36 Thuong, N.T., Asseline, U., Roig, U., Takasuga, M. and Hélène, C. (1987) *Proc. Natl. Acad. Sci. USA.*, **84**, 5129–5133.