4510–4517 Nucleic Acids Research, 1995, Vol. 23, No. 21

Single strand targeted triplex formation: targeting purine–pyrimidine mixed sequences using abasic linkers

Ekambar R. Kandimalla, Adrienne N. Manning, Ganesh Venkataraman[1], V. Sasisekharan[1] and Sudhir Agrawal*

Hybridon, Inc., One Innovation Drive, Worcester, MA 01605, USA and [1]Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received May 18, 1995; Revised and Accepted September 28, 1995

ABSTRACT

Foldback triplex-forming oligonucleotides (FTFOs) that contain an abasic linker, [2-(4-aminobutyr-1-yl)-1,3-propanediol] (APD linker), in the Hoogsteen domain against pyrimidine bases of a C:G and a T:A base pair were studied for their relative stability and sequence specificity of triplex formation. In general, the APD linker has less destabilizing effect against a C:G base pair than a T:A base pair. Incorporation of three APD linker moieties resulted in decreased binding to the target, which was comparable to results observed with three imperfectly matched natural base triplets. The APD linker incorporation did not result in the loss of sequence specificity of FTFOs, unlike in the case of normal triplex-forming oligonucleotides (TFOs). The introduction of a positively charged abasic linker, however, resulted in decreased stability of the triplex, because of loss of hydrogen bonding and stacking interactions in the major groove. The results of a molecular modeling study show that APD linker can be readily incorporated without any change in the conformation of the natural sugar–phosphate backbone conserving overall triple helix geometry. Further, the modeling study suggests a hydrogen bond formation between the amino group of linker and N4 of cytosine mediated by a solvent molecule (water) in the floor of the base triplet in addition to a contribution from the positive charge on the APD linker amino group. Either a direct or water-mediated hydrogen bond between the amino group of the APD linker and the O4 of thymine is unlikely when the linker is placed against a T:A base pair.

INTRODUCTION

Polypurine and polypyrimidine sequences form Py·Pu:Py or Pu·Pu:Py triple helices (1–5). In a triple helix, an oligonucleotide third strand binds in the major groove of the double helix forming Hoogsteen hydrogen bonds with the purine target strand in either parallel or antiparallel orientation (6,7). Widespread use of the triple helix approach is impeded by the pH-dependence of C^+·G:C triplet formation and the fact that their target sites are restricted to purines. Several alternate designs (8–10) and modified bases (11–16) have been proposed to overcome these problems, but with limited success. Recently, two approaches have been proposed to target all pyrimidine target sites using either parallel-stranded hairpin duplexes in the Py.Pu:Py motif (17) or circular oligonucleotides in the Pu.Pu:Py motif (18).

Another approach involves, incorporation of an abasic (null) residue (1,2-dideoxy-D-ribose) in oligonucleotides opposite Py·Pu base pairs to skip a pyrimidine base in a purine-rich target sequence (19). The introduction of an abasic ribose into oligonucleotides significantly reduced the stability of the triplex compared with the triplets T·A:T, C^+·G:C, and G·T:A, and the decrease in binding produced by an abasic residue was similar to that observed with imperfectly matched natural base triplets (19). A recent report described use of a 2-aminobutyl-1,3-propanediol (APD) linker in triplex-forming oligonucleotides (TFOs) to skip over pyrimidine interruptions in the target sequence and to form triplex in a Pu·Pu:Py motif (20). The APD linker stabilized the triplex but decreased sequence specificity. The increased triplex stabilization by the APD linker containing TFOs did not result in increased inhibition of Sp1 binding to the Ha-ras promoter (20). However, utility of this linker in the Py·Pu:Py motif and its affinity against C:G and T:A base pairs with respect to mismatched bases was not investigated.

In the present study, we have substituted an abasic linker, 2-aminobutyl-1,3-propanediol [2-(4-amino-butyr-1-yl)-1,3-propanediol] (APD), in the Hoogsteen domain of foldback triplex-forming oligonucleotides (FTFOs) (21–26). The APD linker matches up with one or three pyrimidine bases in the target sequence in order to overcome the limitation of target sites to polypurines in triplex formation. The linker maintains the normal internucleotide distance (Fig. 1A).

We selected a 19 base long sequence from the initiation codon region of gag m-RNA of the HIV-1 genome as a target (Fig. 1B). It is a purine-rich sequence with three pyrimidine base interruptions at base sites 8, 12 and 14 from the 5′-end (see Fig. 1B). A phosphorothioate oligonucleotide of 25 bases long that encompasses the current target site has antiviral activity against

* To whom correspondence should be addressed

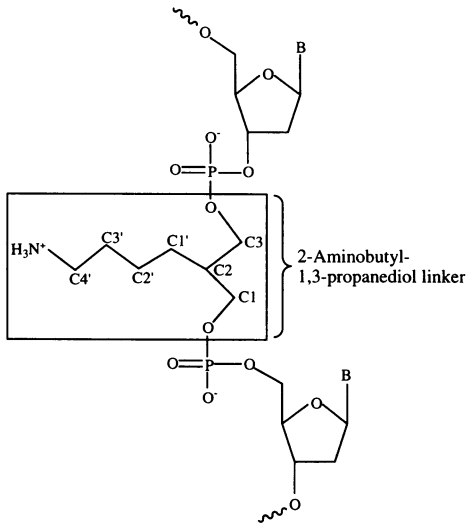

Figure 1. (**A**) Structure of 2-aminobutyl-1,3-propanediol (APD) linker, as incorporated in oligonucleotides at required position, is shown in box. Numbering of the linker carbons used to describe some of the molecular modeling results is shown and B indicates nucleotide base. (**B**) Purine-rich target sequence used in the study. Pyrimidine interruptions are shown in bold.

HIV-1 and is currently being tested in human clinical trials (27). We have synthesized several FTFOs with one or three APD linkers in the Hoogsteen domain (third strand) against T:A or C:G base pairs (Fig. 2). For comparison, we also synthesized and studied several control oligonucleotides without the linker, but with perfectly matched, or mismatched bases, as shown in Figure 2. In this report we describe the ability of the APD linker to bind to mixed sequences by forming foldback triplexes, which we determined by absorbance thermal melting and gel mobility shift experiments. We also describe the influence of the APD linker in the third strand against C:G and T:A base pairs on binding to the target. In addition, we report the results of a molecular modeling study of the effect of an APD linker on the resulting foldback triplex structure.

Purine rich target sequences

1 AGAGAGAAGGGAGAGAGAG-3'
2 AGAGAGAAGGGAGTGAGAG-3'
3 AGAGAGAAGGGTGAGAGAG-3'
4 AGAGAGATGGGAGAGAGAG-3'
5 AGAGAGATGGGTGTGAGAG-3'
6 AGAGAGAAGGGAGCGAGAG-3'
7 AGAGAGATGGGTGCGAGAG-3'

Duplex-forming complementary sequences

1D TCTCTCTTCCCTCTCTCTC-5'
2D TCTCTCTTCCCTCACTCTC-5'
3D TCTCTCTTCCCACTCTCTC-5'
4D TCTCTCTACCCTCTCTCTC-5'
5D TCTCTCTACCCACACTCTC-5'
6D TCTCTCTTCCCTCGCTCTC-5'
7D TCTCTCTACCCACGCTCTC-5'

Foldback triplex-forming sequences (containing natural bases)

1X TCTCTCTTCCCTCTCTCTC-3' / TCTCTCTTCCCTCTCTCTC-5'
2X TCTCTCTTCCCTCACTCTC-3' / TCTCTCTTCCCTCACTCTC-5'
3X TCTCTCTTCCCACTCTCTC-3' / TCTCTCTTCCCACTCTCTC-5'
4X TCTCTCTACCCTCTCTCTC-3' / TCTCTCTACCCTCTCTCTC-5'
5X TCTCTCTACCCACACTCTC-3' / TCTCTCTACCCACACTCTC-5'
6X TCTCTCTTCCCTCGCTCTC-3' / TCTCTCTTCCCTCGCTCTC-5'
7X TCTCTCTACCCACGCTCTC-3' / TCTCTCTACCCACGCTCTC-5'

Foldback triplex-forming sequences (containing APD linker in the Hoogsteen domain)

2L TCTCTCTTCCCTCLCTCTC-3' / TCTCTCTTCCCTCACTCTC-5'
3L TCTCTCTTCCCLCTCTCTC-3' / TCTCTCTTCCCACTCTCTC-5'
4L TCTCTCTLCCCTCTCTCTC-3' / TCTCTCTACCCTCTCTCTC-5'
5L TCTCTCTLCCCLC LCTCTC-3' / TCTCTCTACCCACACTCTC-5'
6L TCTCTCTTCCCTCLCTCTC-3' / TCTCTCTTCCCTCGCTCTC-5'
7L TCTCTCTLCCCLC LCTCTC-3' / TCTCTCTACCCACGCTCTC-5'

Figure 2. Sequences used in the current study. Bases under study are shown in bold. Target sequences are shown from 5′ to 3′ and all other sequences from 3′ to 5′ for easy comparison. Letter extensions D, X and L for the sequence numbers stand for duplex-forming, foldback triplex-forming and linker containing foldback triplex-forming oligonucleotides, respectively.

MATERIALS AND METHODS

Oligonucleotide synthesis and purification

Oligodeoxyribonucleotides were synthesized on a Milligen 8700 DNA synthesizer using β-cyanoethyl phosphoramidite chemistry. All the monomer synthons were obtained from Milligen. The other reagents and solvents required for the synthesis were purchased from either Milligen or Cruachem. The oligonucleotides containing the APD linker were synthesized on a DNA synthesizer using N-Fmoc-O^1-DMT-O^3-cyanoethoxydiisopropyl-aminophosphinyl-2-aminobutyryl-1,3-propanediol that was purchased from Clontech. The oligonucleotides were deprotected by incubating with concentrated ammonium hydroxide at room temperature for 1.5 h and then further incubating the supernatant at 55°C for 6 h. Oligonucleotides were then purified on a C18 reverse-phase HPLC, using 0.1 M aqueous ammonium acetate and acetonitrile containing 20% 0.1 M ammonium acetate solvent system. HPLC-purified oligonucleotides were detritylated with 80% aqueous acetic acid for 1 h at room temperature. The oligonucleotides were desalted on C18 reverse-phase Sep-Pak cartridges (Waters) and the purity was checked by polyacrylamide gel electrophoresis. Oligonucleotides that were less than 95–97% pure were further purified by preparative 15% polyacrylamide gel electrophoresis.

UV thermal melting experiments

UV thermal melting curves at 260 nm were determined at a heating rate of 0.5°C/min with a Perkin-Elmer Lambda 2 spectrophotometer in 100 mM sodium acetate, pH 5.0 buffer containing 10 mM $MgCl_2$. The oligonucleotide concentration was 1.5 μM per each

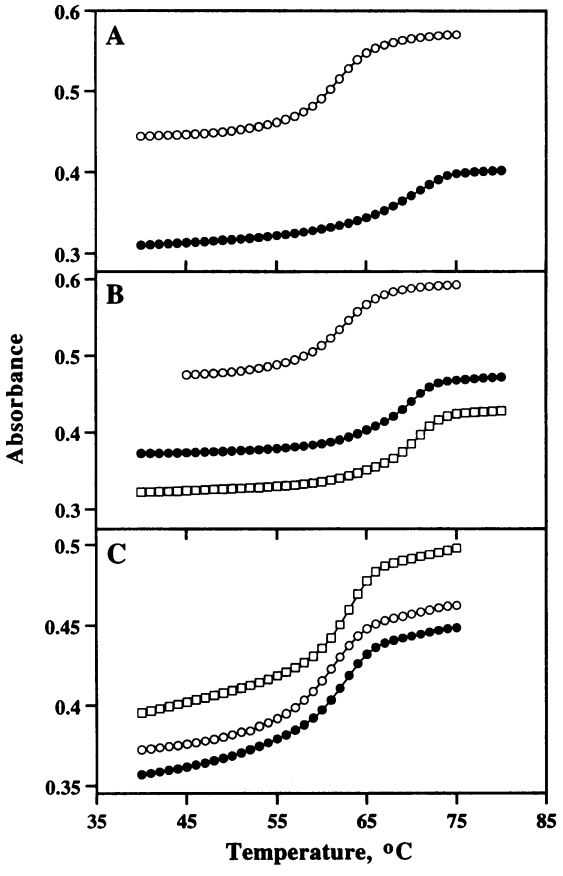

Figure 3. Thermal melting curves of representative foldback triplexes and control duplexes. (**A**) **1+1D** (○) and **1+1X** (●); (**B**) **3+3D** (○), **3+3X** (●) and 3+3L (□); and (**C**) **5+5D** (○), **5+5X** (●) and 5+5L (□).

Table 1. Thermal melting data of the complexes of oligonucleotides[a]

Oligo number	Complex with specific bases under study	$T_m(\Delta T_m)^b$, °C
1+1D	5'—A—A—A— / —T—T—T—	61.8
1+1X	—T—T—T— / 5'—A—A—A— / —T—T—T—	70.9 (9.1)
2+2D	5'—A—A—T— / —T—T—A—	62.3
2+2X	—T—T—A— / 5'—A—A—T— / —T—T—A—	68.8 (6.5)
2+2L	—T—T—L— / 5'—A—A—T— / —T—T—A—	69.4 (7.1)
3+3D	5'—A—T—A— / —T—A—T—	62.4
3+3X	—T—A—T— / 5'—A—T—A— / —T—A—T—	69.1 (6.7)
3+3L	—T—L—T— / 5'—A—T—A— / —T—A—T—	70.5 (8.1)
4+4D	5'—T—A—A— / —A—T—T—	61.9
4+4X	—A—T—T— / 5'—T—A—A— / —A—T—T—	69.0 (7.1)
4+4L	—L—T—T— / 5'—T—A—A— / —A—T—T—	70.5 (8.6)
5+5D	5'—T—T—T— / —A—A—A—	61.3
5+5X	—A—A—A— / 5'—T—T—T— / —A—A—A—	61.5 (0.2)
5+5L	—L—L—L— / 5'—T—T—T— / —A—A—A—	63.0 (1.7)

[a]See Figure 2 for sequences and Materials and Methods for experimental conditions.
[b]ΔT_m (shown in parentheses) is the difference between T_ms of the foldback triplex and the corresponding duplex.

strand. The mid point of the thermal melting curves (T_m) was determined from first derivative curves obtained by plotting d(absorbance at 260 nm)/d(temperature) versus temperature. Each value reported is an average of two individual experiments. The uncertainties in the T_m values were estimated to be within ± 0.5 °C.

Gel mobility shift experiments

The target oligonucleotides (**1–7**) were labeled with ^{32}P at the 5′-end using [γ-^{32}P]ATP and T4 polynucleotide kinase. A small amount of labeled oligonucleotide was mixed with different ratios of appropriate duplex- or triplex-forming oligonucleotide in 100 mM sodium acetate (pH 5.0)/10 mM $MgCl_2$, heated to 95 °C for 10 min and cooled to room temperature before analysis on gels. Each sample contained 0.1 A_{260} U of sonicated salmon sperm DNA in a final volume of 20 µl. All the samples were loaded on gel with glycerol-dye mix. Electrophoresis was carried out on 15% native polyacrylamide gels at room temperature using 50 mM Tris–glycine, pH 5.0 as running buffer. After electrophoresis, gels were dried and autoradiographed using Kodak X-Omat AR film at –70 °C. Both gel and running buffer contained 100 mM sodium acetate and 10 mM $MgCl_2$.

Molecular modeling

Molecular modeling was performed on a Silicon Graphics Iris Indigo workstation using Insight II program (version 2.3.1 Biosym Technologies, San Diego, CA). The coordinates of the target purine-rich strand were generated using the LALS program with the starting conformation as published (28), using the linked atom least squares program (29) with helical parameters $n = 12$ and $h = 3.26$. The Hoogsteen strand was generated from the purine strand coordinates by a rotation of 69.5° about the helix axis. The coordinates of the Watson–Crick base paired pyrimidine chain for the triple helix was generated from the coordinates of the purine strand by applying a 2-fold symmetry operation. The Watson–Crick strand was extended at the 3′-end by five bases (-CTCTC-) in order to create a loop joining to the Hoogsteen strand. The backbone torsions of the loop were then adjusted to form a phosphodiester linkage with the Hoogsteen pyrimidine strand. The loop structure was minimized, keeping the Watson–Crick and the Hoogsteen pyrimidine strands fixed. Energy

Table 2. Thermal melting data of the complexes of oligonucleotides[a]

Oligo number	Complex with specific bases under study	$T_m(\Delta T_m)^b$, °C
6+6D	5'—A—A—C— / —T—T—G—	65.1
6+6X	—T—T—G— / 5'—A—A—C— / —T—T—G—	66.4 (1.3)
6+6L	—T—T—L— / 5'—A—A—C— / —T—T—G—	71.8 (6.7)
7+7D	5'—T—T—C— / —A—A—G—	65.7
7+7X	—A—A—G— / 5'—T—T—C— / —A—A—G—	65.4 (-0.3)
7+7L	—L—L—L— / 5'—T—T—C— / —A—A—G—	67.0 (1.3)
5+1D	5'—T—T—T— / —T—T—T—	48.2
5+1X	—T—T—T— / 5'—T—T—T— / —T—T—T—	49.1
1+5D	5'—A—A—A— / —A—A—A—	49.6
1+5L	—L—L—L— / 5'—A—A—A— / —A—A—A—	54.9

[a]See Figure 2 for sequences and Materials and Methods for experimental conditions.

[b]ΔT_m (shown in parentheses) is the difference between T_ms of the foldback triplex and the corresponding duplex.

minimizations were performed using the Discover force field with 100 steps of steepest descents followed by 500 steps of conjugate gradient method in vacuum (dielectric of 1.0). As no solvent or counter ions were modeled, electrostatic interactions were not included during the minimization.

The bases in positions 8, 12, and 14 (opposite pyrimidine interruptions in the target strand) of the Hoogsteen strand were modified to have a sugar phosphate group alone (no base) or the APD linker. The conformation of the APD linker was determined by minimizing the triple helix structure keeping all the other residues fixed and allowing only the conformation of the linker to change. Additional minimizations, allowing all the residues to move without constraints did not alter the structure significantly. The ability of water molecules to bridge hydrogen bonds between the amine of the APD linker and bases was explored by manually docking a water molecule at stereochemically feasible sites close to the base.

RESULTS AND DISCUSSION

Effect of basic (mismatch) and abasic (APD) linker substitution in FTFOs over pyrimidine interruptions in the target

The thermal melting data are shown in Tables 1 and 2. Oligonucleotide **1X** is an ideal FTFO that contains all perfectly matched pyrimidine bases in the Hoogsteen domain. The foldback triple helix complex of oligonucleotide **1X** and its polypurine target strand **1** showed a T_m of 70.9°C with a single, sharp, cooperative melting transition (Fig. 3A). The duplex of oligonucleotide **1D** with the same target strand **1** showed a T_m of

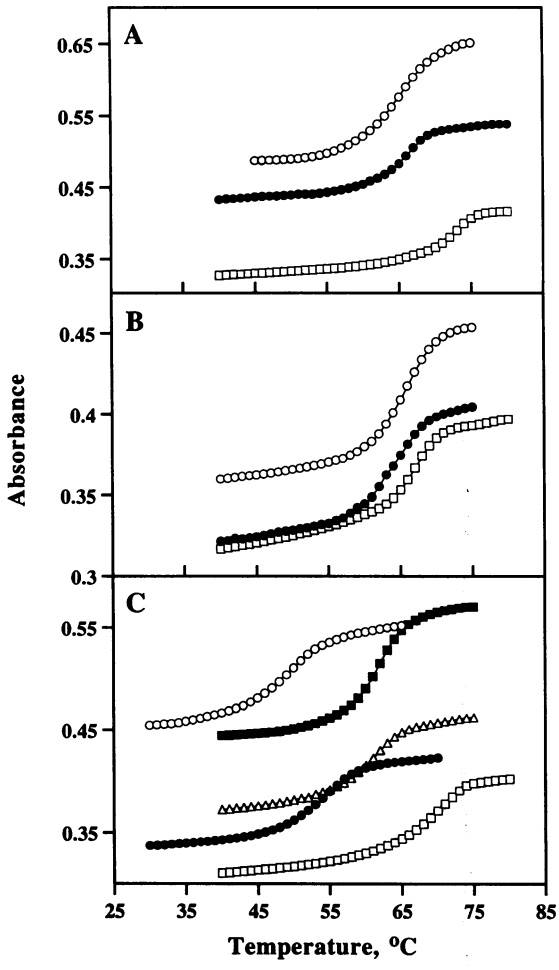

Figure 4. Thermal melting curves of representative foldback triplexes and control duplexes. (**A**) **6+6D** (○), **6+6X** (●) and **6+6L** (□); (**B**) **7+7D** (○), **7+7X** (●) and **7+7L** (□); and (**C**) **1+5D** (○), **1+5L** (●), **1+1X** (□), **1+1D** (■) and **5+5D** (△).

61.8°C (Fig. 3A). The difference (ΔT_m) of 9.1°C in the thermal stability of the foldback triplex (**1–1X**) and the corresponding Watson–Crick double helix (**1–1D**) suggests tight binding of the Hoogsteen domain of the FTFO in the major groove (21).

The foldback triplex of oligonucleotides **2X** (with an adenine against T:A base pair) and the target **2** had a T_m of 68.8°C. The corresponding duplex (**2–2D**) showed a T_m of 62.3°C. The difference (ΔT_m) in the T_ms of the foldback triplex and the duplex is only 6.5°C (Table 1). The lower stability of the foldback triplex of **2–2X** compared with **1–1X** indicates loss of hydrogen bonding as well as stacking interactions that are the result of a mismatched adenine base present in the Hoogsteen domain of oligonucleotide **2X**. The foldback triplex of oligonucleotide **2L** with an APD linker against a T:A base pair and its target sequence **2** showed a T_m of 69.4°C. When the mismatched adenine was replaced with the APD linker (oligonucleotide **2L**), the resulting foldback triplex (**2–2L**) had 0.6°C higher thermal stability. The ΔT_m (Table 1) of the complex **2–2L** compared with **2–2X**, although minimal, suggests that the abasic linker destabilizes the complex to a lesser extent than a mismatched adenine base (see molecular modeling results). The higher thermal stability of the foldback

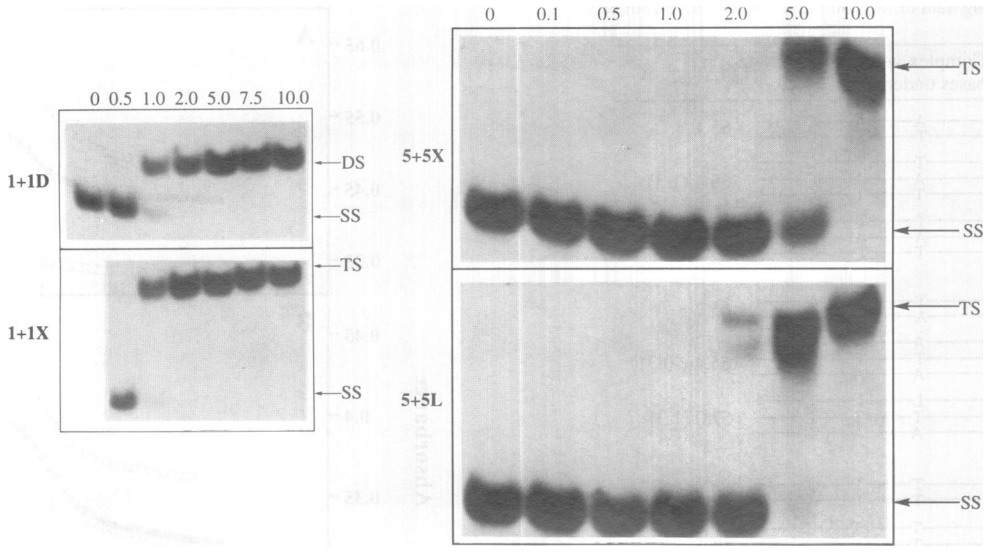

Figure 5. Relative binding affinity of representative duplex and foldback triplex-forming oligonucleotides as a function of gel mobility shift. Target sequences **1** and **5** were ^{32}P-labeled at the 5′-end. The ratio of the added oligonucleotide in each lane is shown on top of the gel. Gels contain oligonucleotide combinations as labeled. Bands labeled SS, DS and TS correspond to single, double and triple strands, respectively.

triplex of oligonucleotide **2–2L** could result from charge interactions between protonated amino group of the APD linker and bases in the target strand and/or phosphodiester backbone.

The foldback triplexes of oligonucleotides **3X** and **4X**, which have an adenine in the Hoogsteen domain, with their corresponding target strands (**3** and **4**) showed T_ms of 69.1 (Fig. 3B) and 69°C, respectively. The corresponding duplexes **3–3D** and **4–4D** showed T_ms of 62.4 and 61.9°C, respectively. The lower T_ms of these foldback triplexes compared with the perfectly matched triplex **1–1X** could be the result of the mismatch in the Hoogsteen domain. The foldback triplexes of oligonucleotides **3L** and **4L**, which have an APD linker substituted for an adenine in the Hoogsteen domain, with their target sequences **3** (Fig. 3B) and **4**, respectively, showed T_ms of 70.5°C. These T_m values are very similar to the T_m of the perfectly matched foldback triplex of oligonucleotides **1–1X** (Table 1). These results suggest that the APD linker may be useful as a null site against a pyrimidine base in the formation of a stable foldback triplex. The differences in the T_ms of different duplexes and triplexes could result from local sequence-dependent conformational effects.

The foldback triplex of oligonucleotide **5X**, which has three mismatched bases in the Hoogsteen domain, and its target sequence **5** showed a T_m of 61.5°C (Fig. 3C). Comparison of T_m of this mismatched foldback triplex with that of corresponding duplex, **5–5D** (61.3°C) suggests that the third strand contributes insignificantly to the stability of the foldback complex. The foldback triplex of oligonucleotide (**5L**), which contains three APD linkers in place of the three mismatched adenine bases (**5X**) in the Hoogsteen domain, had a T_m of 63.0°C (Fig. 3C). The higher stability of ~1.7°C for the complex of **5–5L** over **5–5X** could result from the absence of destabilizing (steric) interactions or the occurrence of charge interactions between protonated amino group of the APD linker and the phosphate backbone. The T_m of the complex **5–5L** is lower than that of **1–1X**, suggesting a loss of hydrogen bonding and stacking interactions in the **5–5L** complex (19). These results suggest that multiple pyrimidine sites might not be the best choice for targeting through triplex or foldback triplex formation with APD linker substitution. A pyrimidine site in a purine-rich environment can be targeted using the APD linker approach without losing significant binding affinity in Py·Pu:Py motif.

Effect of APD linker over C:G base pair compared with A:T base pair

The foldback triplex of oligonucleotides **6–6X** and the corresponding duplex (**6–6D**) showed T_ms of 66.4 and 65.1°C, respectively (Fig. 4A). The T_m of this foldback triplex indicates that the presence of a G in the third strand against a C in the target strand destabilizes the complex considerably (30). When G in the Hoogsteen domain is substituted with the APD linker, the resulting foldback triplex of oligonucleotides **6–6L** showed a T_m of 71.8°C (Fig. 4A). The T_m for the complex of **6–6L** is 5.4°C (Table 2) higher than that for **6–6X**, suggesting that the APD linker against a C:G base pair stabilizes the triplex.

The foldback triplex of oligonucleotides **7–7X**, which has three mismatched purine bases in the Hoogsteen domain against the three pyrimidine interruptions in the target, showed a T_m of 65.4°C (Fig. 4B). Comparison of the T_m (Table 2) of complex **7–7X** with that of the duplex **7–7D** (65.7°C) suggests insignificant binding of the third strand in the major groove as a result of three purine mismatches in the Hoogsteen domain of oligonucleotide **7X**. The foldback triplex of oligonucleotides **7–7L** had a T_m of 67°C (Fig. 4B), which is about 1.6°C higher than the mismatched foldback triplex of oligonucleotides **7–7X** (Table 2). Oligonucleotide **7L** has three APD linkers; one against a C:G base pair and two against two T:A base pairs. These results further confirm that use of an APD linker over pyrimidine interruptions in the target stabilizes the triplex complex compared with a mismatched purine base. It is also clear that a mismatched purine base in the third strand destabilizes the triplex considerably, at least in the case of A against T and G against C.

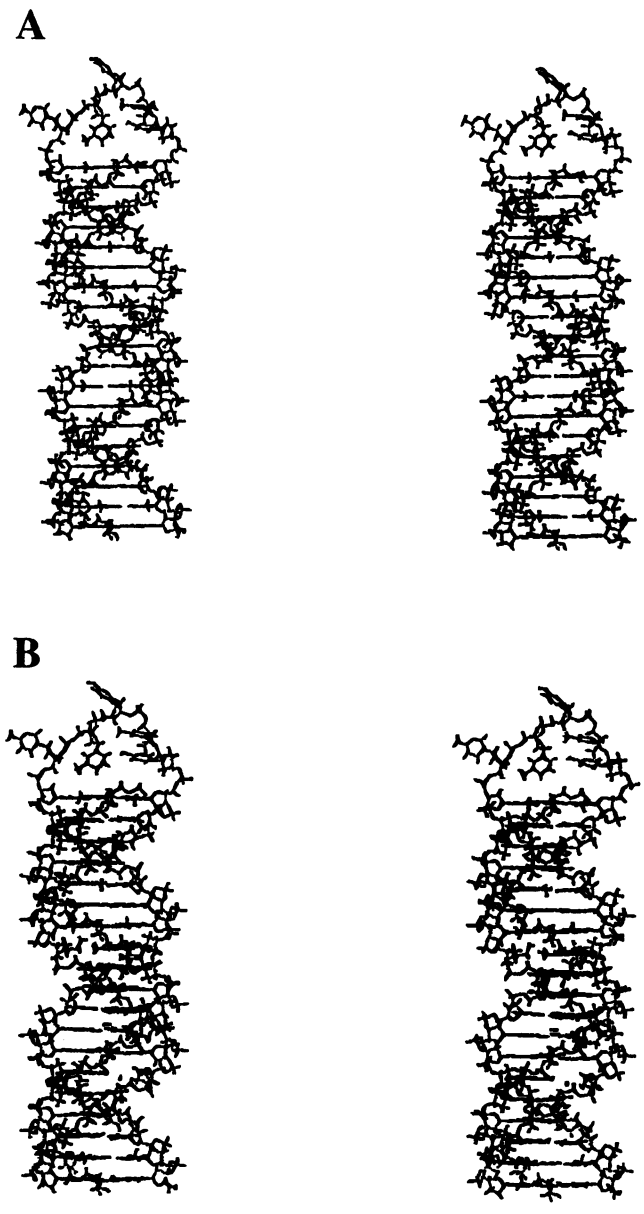

Figure 6. Stereo drawings of the energy minimized foldback triplex structures of (**A**) oligonucleotides **1–1X** with matched bases in the Hoogsteen strand and (**B**) oligonucleotides **7–7L** containing three APD linkers in the Hoogsteen domain against T:A and C:G base pairs.

Comparison of the T_m of the foldback triplex of oligonucleotides **6–6L** (71.8°C) with that of oligonucleotides **2–2L** (69.4°C) further suggests that an APD linker is better tolerated against a C:G base pair than a T:A base pair. These experiments suggest that (i) a guanine in the third strand is sterically incompatible with a C:G base pair, as is the case of an adenine in the third strand against a T:A base pair (31); (ii) an adenine in the third strand against a T:A base pair is better tolerated than a G against a C:G base pair (compare ΔT_ms of triplexes **2–2X** and **6–6X**) (Table 1 and 2); and (iii) the APD linker in the third strand against a C:G base pair has a greater stabilizing effect on the triplex than against a T:A base pair (compare T_ms of foldback triplexes of oligonucleotides **2–2L** with **6–6L**) (see also molecular modeling results).

Sequence specificity of FTFOs containing abasic substitution

Complementary base recognition through Watson–Crick duplex and Hoogsteen triplex formation is highly sequence-specific. The foldback and circular oligonucleotides that bind to the single-stranded target sequence are much more sequence-specific than oligonucleotides that bind to the target either by just Watson–Crick hydrogen bonding (antisense) or Hoogsteen hydrogen bonding (antigene or TFOs). The FTFOs are more sequence-specific because they read the target sequence twice: first when they form a duplex with the target sequence, and second when the Hoogsteen domain binds to form the triplex (21,24). An abasic site introduced into a TFO intended to bind to double-stranded DNA containing pyrimidine interruptions can decrease the sequence specificity by allowing the TFO to bind to non-targeted sequences (19,20). This problem should not arise with FTFOs that contain an APD linker in the Hoogsteen domain, because sequence specificity is mainly determined by the Watson–Crick domain, while the Hoogsteen strand contributes to the stability of the resulting foldback triple helix.

The complex of oligonucleotides **5–1X** contained three mismatches each in both the Watson–Crick and the Hoogsteen domains. The T_m measured for this complex was 49.1°C, which is about 0.9°C higher than the corresponding mismatched duplex of oligonucleotides **5–1D** (Table 2), suggesting that the third strand binds insignificantly because of the three mismatches. The mismatched foldback triplex (**5–1X**) has 12.7 and 12.2°C lower thermal stability compared with the perfectly matched duplexes **1–1D** and **5–5D**, respectively. The mismatched duplex of oligonucleotides **1–5D** had a T_m of 49.6°C, which is about 12.2 and 11.7°C lower than the perfectly matched duplexes of oligonucleotides **1–1D** and **5–5D**, respectively. The foldback triplex of oligonucleotides **1–5L**, which contains three APD linkers in the Hoogsteen domain and three mismatches in the Watson–Crick domain, showed a T_m of 54.9°C (Fig. 4C). This T_m is about 8.1°C lower than the corresponding triplex (**5–5L**) that has no mismatches in the Watson–Crick domain and three APD linkers in the Hoogsteen domain. Similarly, the triplex of **1–5L** has T_ms 6.8 and 6.4°C lower than the corresponding duplexes **1–1D** and **5–5D** without mismatches (Fig. 4C). The higher T_m for the complex of oligonucleotides **1–5L** compared with that of the complexes of oligonucleotides **5–1X** and **1–5D** could result from the presence of three positively charged amino groups in the former complex. This result demonstrates that the sequence specificity of FTFOs is not altered significantly by the presence of APD linkers in the Hoogsteen domain. The stability of triplexes, however, depends on the base composition and experimental conditions as well as flanking sequences, and therefore may be different for each individual sequence.

Gel mobility shift

The relative binding strength of oligonuclotides to their target sequences can be determined by the gel mobility shift assay. Figure 5 shows the gel mobility of representative duplex and triplex complexes. The control FTFO (**1X**) without any mismatches formed a stable complex at 1:1 (**1:1X**) ratio (Fig. 5A). For comparison, the mobility of the duplex of oligonucleotides **1–1D** is also shown in Figure 5A.

The FTFOs with an adenine mismatch in the Hoogsteen domain (**2X–4X**) formed up to 75% triplex with their respective target

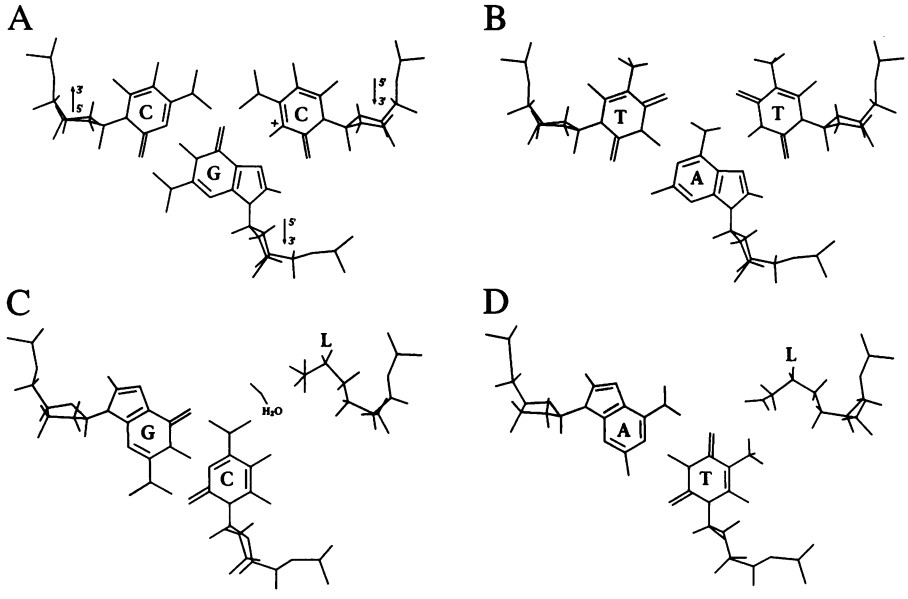

Figure 7. Normal C:G·C⁺ (**A**), and T:A·T (**B**) triplets. The triplets containing linker against C:G and T:A base pairs are shown in (**C**) G:C·L and (**D**) A:T·L as observed in energy minimized structures. Note position of water molecule that could mediate hydrogen bonding between amino group of the APD linker and N4 of cytosine in (C). The arrows in (A) indicate polarity of each strand. The three strands in (B), (C) and (D) have the same polarity as in (A). The letters C, G, A, T and L stand for cytosine, guanosine, adenosine, thymidine and the APD linker, respectively.

strands (**2–4**) at a ratio of 1:2 as determined by triplex band intensity (data not shown). Under the same experimental conditions FTFOs (**2L–4L**) containing an APD linker instead of an adenine mismatch in the Hoogsteen domain showed >90% triplex band formation with their respective target sequences at the same ratio (data not shown). At the next higher ratio (1:5) studied, these oligonucleotides containing an adenine mismatch or an APD linker formed 100% triplex formation. These results support thermal melting data obtained with these oligonucleotides.

Oligonucleotide **5X**, which has three adenine mismatches, showed up to 55% (at 1:5 ratio) and 100% (at 1:10 ratio) triplex formation with its target sequence (**5**) at high concentrations, suggesting weak triplex formation (Fig. 5B). Oligonucleotide **5L** with three APD linkers in place of the three adenine mismatches, however, showed a >90% triplex band at 1:5 ratio with the same target sequence (Fig. 5B), indicating that the APD linkers are better tolerated than the mismatched adenines in the Hoogsteen domain of FTFOs. These results support conclusions drawn from thermal melting experiments. We obtained similar results with oligonucleotides **7X** and **7L** with their target sequence **7** (data not shown).

Molecular modeling

A stereo drawing of an energy-minimized structure of a perfectly matched foldback triplex of oligonucleotides **1–1X** is shown in Figure 6A. Interstrand hydrogen bonding (see Fig. 7A and B for hydrogen bonding pattern), intrastrand stacking, sugar–base and phosphate–base interactions mainly contribute to the stability of the complex. The triplex structure (**7–7L**) in which the Hoogsteen strand contains APD linkers opposite the pyrimidine interruptions in the purine-rich target strand is shown in Figure 6B. Comparison of the two structures reveals that the backbone 1,3-propanediol (APD) structure adopts a conformation very similar to the natural sugar–phosphate backbone, conserving the overall geometry of the triple helical structure. The two methylene bridges that join C1′ of the butyl chain with C2 of propanediol and C1′–C2′ of butyl chain adopt an extended trans conformation while the other two linkages joining C2′–C3′ and C3′–C4′ of butyl chain adopt a gauche conformation (see Fig. 1A for numbering). The flexible APD linker does not interfere with Watson–Crick hydrogen bonding at this site or the neighboring sites.

We further investigated if the amino group of the APD linker could involve itself in hydrogen bonding with either the backbone or the bases to stabilize the complex. When the APD linker is placed against a C:G base pair, the orientation of the amine group of the APD linker allows sufficient room between the bases to accommodate a solvent (water) molecule in the floor of the major groove, which could bridge a hydrogen bond between the amine of the APD linker and the N4 of cytosine (Fig. 7C). When the APD linker is placed against a T:A base pair, however, it adopts a conformation such that no direct hydrogen bond formation between the amino group of the APD linker and the O4 of thymine is possible because of steric hindrance from the C5-methyl group of thymine (Fig. 7D). In addition, the C6-amino group of adenine restricts the available space for a solvent molecule in the floor of the major groove, so that any solvent molecule within this restricted space is not capable of bridging a hydrogen bond between the amine group of the APD linker and the O4 of thymine. The additional hydrogen bond between the amino group of the APD linker and cytosine base might explain the stabilizing effect observed in thermal melting and gel mobility shift studies when the APD linker is placed opposite a C:G base pair rather than a T:A base pair. The positive charge of the APD linker provides additional stabilization energy.

CONCLUSIONS

A pyrimidine base in a purine-rich sequence can be targeted using an abasic linker (APD) in FTFOs without significant loss of binding affinity. The APD linker showed different affinities against C:G and T:A base pairs. We were unable to distinguish, however, whether this difference resulted from local sequence-dependent conformational effects near the site of modification or differences in the thermodynamic properties resulting from changes in nearest neighbor interactions in the foldback triplex. A molecular modeling study showed sufficient room for a solvent (water) molecule to fit in the floor of the major groove between the bases and mediate hydrogen bond formation between the amino group of the linker and N4 of cytosine. It is unlikely that this would occur with a T:A base pair, which could explain the experimentally observed differences in the affinities of the APD linker against a C:G and a T:A base pair. Sequence specificity of the FTFOs was not affected by APD linker substitution, because the sequence specificity is mainly dictated by the Watson–Crick domain of the oligonucleotide. At the present time viable approaches to targeting pyrimidine sites in the target sequence through triplex formation do not exist. Using the present strategy, one may be able to target at least a limited number of pyrimidine sites in a purine-rich sequence through foldback triplex formation without loss of sequence specificity and significant binding affinity.

REFERENCES

1 Cooney, M., Czernuszewicz, G., Postel, E.H., Flint, S.J. and Hogan, M.E. (1988) *Science* **241**, 456–459.
2 Lyamichev, V.I., Mirkin, S.M., Frank-Kamenetskii, M.D. and Burkholder, G.D. (1988) *Nucleic Acids Res.* **16**, 2165–2178.
3 Maher, L.J. III, Wold, B. and Dervan, P.B. (1989) *Science* **245**, 725–730.
4 Moser, H.E. and Dervan, P.B. (1987) *Science* **238**, 645–650.
5 Peuth, D., Perrouault, L., LeDoan, T., Chassignol, M., Thoung, N. and Hélène, C. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 1349–1353.
6 Cheng, Y-K. and Pettitt, B.M. (1992) *Prog. Biophys. Mol. Biol.* **58**, 225–257.
7 Thoung, N.T. and Hélène, C. (1993) *Angew. Chem. Int. Ed. Engl.* **32**, 666–690.
8 Horne, D.A. and Dervan, P.B. (1990) *J. Am. Chem. Soc.* **112**, 2435–2437.
9 Jayasena, S.D. and Johnston, B.H. (1993) *Biochemistry* **32**, 2800–2807.
10 Ono, A., Chen, C-.N. and Kan, L. (1991) *Biochemistry* **30**, 9914–9921.
11 Jetter, M.C. and Hobbs, F.W. (1993) *Biochemistry* **32**, 3249–3254.
12 Koh, J.S. and Dervan, P.B. (1992) *J. Am. Chem. Soc.* **114**, 1470–1478.
13 Miller, P.S., Bhan, P., Cushman, P.D. and Trapane, T.L. (1992) *Biochemistry* **31**, 6788–6793.
14 Ono, A., Ts'o, P.O.P. and Kan, L.S. (1991) *J. Am. Chem. Soc.* **113**, 4032–4033.
15 Xiang, G., Soussou, W. and McLaughlin, L.W. (1994) *J. Am. Chem. Soc.* **116**, 11155–11156.
16 Young, S.L, Krawczyk, S.H., Matteucci, M.D. and Toole, J.J. (1991) *Proc. Natl. Acad. Sci. USA.* **88**, 10 023–10 026.
17 Kandimalla, E.R., Agrawal, S., Venkataraman, G. and Sasisekharan, V. (1995) *J. Am. Chem. Soc.* **117**, 6416–6417.
18 Wang, S. and Kool, E. (1994) *J. Am. Chem. Soc.* **116**, 8857–8858.
19 Horne, D.A. and Dervan, P.B. (1991) *Nucleic Acids Res.* **18**, 4963–4965.
20 Mayfield, C. and Miller, D. (1994) *Nucleic Acids Res.* **22**, 1909–1916.
21 Kandimalla, E.R. and Agrawal, S. (1994) *Gene* **149**, 115–121.
22 Kandimalla, E.R. and Agrawal, S. (1995) *Nucleosides and Nucleotides* **14**, 991–995.
23 Kandimalla, E.R. and Agrawal, S. (1995) *Nucleic Acids Res.* **23**, 1068–1074.
24 Prakash, G. and Kool, E.T. (1992) *J. Am. Chem. Soc.* **114**, 3523–3528.
25 Xodo, L.E., Manzini, G. and Quadrifoglio, F. (1990) *Nucleic Acids Res.* **18**, 3557–3564.
26 Giovannangeli, C., Montenay-Garestier, T., Rougee, M.; Chassignol, M., Thoung, N.T. and Hélène, C. (1991) *J. Am. Chem. Soc.* **113**, 7775–7776.
27 Agrawal, S. and Tang, J.-Y. (1992) *Antisense Res. Dev.* **2**, 261–266.
28 Raghunathan, G., Miles, H.T. and Sasisekharan, V. (1993) *Biochemistry* **32**, 455–462.
29 Smith, P.J.C. and Arnott, S. (1978) *Acta Cryst.* **a34**, 3–11.
30 Griffin, L.C. and Dervan, P.B. (1989) *Science* **245**, 967–971.
31 Mergny, J.-L., Sun, J.-S., Rougee, M., Montenay-Garestier, T., Barcelo, F.,Chomilier, J. and Hélène, C. (1991) *Biochemistry* **30**, 9791–9798.