Stable Recognition of TA Interruptions by Triplex Forming Oligonucleotides Containing a Novel Nucleoside†

Yang Wang,‡ David A. Rusling,§ Vicki E. C. Powers,‡ Oliver Lack,‡,|| Sadie D. Osborne,‡,⊥ Keith R. Fox,*,§ and Tom Brown‡

School of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ, United Kingdom, and School of Biological Sciences, University of Southampton, Bassett Crescent East, Southampton SO16 7PX, United Kingdom

Received January 4, 2005; Revised Manuscript Received February 15, 2005

ABSTRACT: We have prepared the 2′-aminoethoxy derivative of the S nucleoside (^{2AE}S) and incorporated it into triplex-forming oligonucleotides for recognition of TA interruptions within a target oligopurine tract. Fluorescence melting, UV melting, and DNase I footprinting experiments show that ^{2AE}S has greater affinity than G or S for a single TA interruption. Stable triplexes are formed at pH 6.0 at an 18-mer target site containing two TA interruptions, even though this contains eight C^+.GC triplets. Although ^{2AE}S and S produce stable triplexes at TA interruptions, they also interact with other base pairs, in particular, CG, although the selectivity for TA improves with increased pH. ^{2AE}S is the best nucleoside described so far for recognition of TA within a triple-helix target.

Triplex-forming oligonucleotides (TFOs) are gene-targeting agents with many potential applications in biotechnology and molecular biology (*1*−*4*). They bind in the major groove of double-stranded DNA, forming specific hydrogen bonds with exposed groups on the base pairs (*5*). TFOs containing pyrimidine bases bind in a parallel orientation to the purine strand of the target DNA duplex, recognizing AT and GC base pairs to form T.AT and C^+.GC triplets. Because there are no simple methods for targeting pyrimidine bases, this restricts triplex formation to polypurine.polypyrimidine stretches of the duplex and imposes a severe limit on the number of potential target sites (*6*). Consequently, the search for nucleoside analogues that recognize TA and CG base pairs (pyrimidine inversions) in DNA and enable triplex formation at mixed-sequence target sites is the focus of intense activity.

The first report of TA base-pair recognition in the parallel-binding motif involved the use of guanine to form a G.TA triplet (*7*). However, this triplet has lower stability than T.AT or C^+.GC, probably because the third strand only forms one hydrogen bond with the TA base pair. The stability of the G.TA triplet is influenced by the surrounding DNA sequences and is more stable when flanked by T.AT than C^+.GC (*8*−*12*). Despite the subsequent development of novel nucleosides with optimized hydrogen-bonding contacts and improved π-stacking interactions, only moderate progress has been made in the recognition of pyrimidine inversion (*13*−*16*). Some success at increasing the affinity of weaker triplets has come from the use of chemically modified backbones and sugars in the third strand (*17*−*21*), and one such sugar modification, 2′-aminoethoxy ribose, greatly increases the affinity (*22*−*28*), association rates, and nuclease resistance (*26*) of TFOs. Fully aminoethoxy-modified TFOs form triplexes of particularly high stability, and this is attributed to preorganization into an RNA-like conformation (*26*−*28*) and specific interactions between the protonated amines (ethylammonium ions) of the TFO and the negatively charged phosphodiester groups of the duplex (*22*). This study highlights the importance of extending the repertoire of available 2′-aminoethoxy-modified nucleosides to produce TFOs that recognize all four base pairs.

It has recently been reported that the unnatural thiazolylaniline deoxyribonucleoside (S) (Figure 1a) recognizes TA base pairs with an affinity comparable to that of the T.AT triplet (*29*−*31*). This synthetic nucleoside has some structural similarity to D_3, which was originally proposed for shape-specific recognition of TA and CG (*13*) but was subsequently shown to bind by intercalation at pyrimidine-purine base steps (*14*). However, a series of studies with S and its derivatives suggested that it forms a specific complex with TA base pairs (*29*−*31*). This is an encouraging advance, and we now describe the synthesis of TFOs containing the 2′-aminoethoxyribo-derivative of S (^{2AE}S). In this paper, we use DNase I footprinting and thermal melting experiments to examine the stability of triplexes that contain this novel nucleoside. The stability and selectivity of triplexes containing ^{2AE}S is compared with each of the natural DNA bases and with S.

MATERIALS AND METHODS

Chemical Synthesis of the ^{2AE}S Monomer. Synthesis of the ^{2AE}S phosphoramidite **11** is shown in Scheme 1. Full

† This work was supported by grants from Cancer Research, U.K., and the European Union. D.A.R. is supported by a research studentship from EPSRC.

* To whom correspondence should be addressed: School of Biological Sciences, University of Southampton, Bassett Crescent East, Southampton SO16 7PX, U.K. Telephone: +44 23 8059 4374. Fax: +44 23 8059 4459. E-mail: K.R.Fox@soton.ac.uk.

‡ School of Chemistry.

§ School of Biological Sciences.

|| Present address: Pharmaceuticals Division, F. Hoffmann−La Roche Ltd., CH-4070 Basel, Switzerland.

⊥ Present address: Department of Chemistry, University of Sheffield, Sheffield S3 7HF, U.K.

10.1021/bi050013v CCC: $30.25
Published on Web 03/19/2005

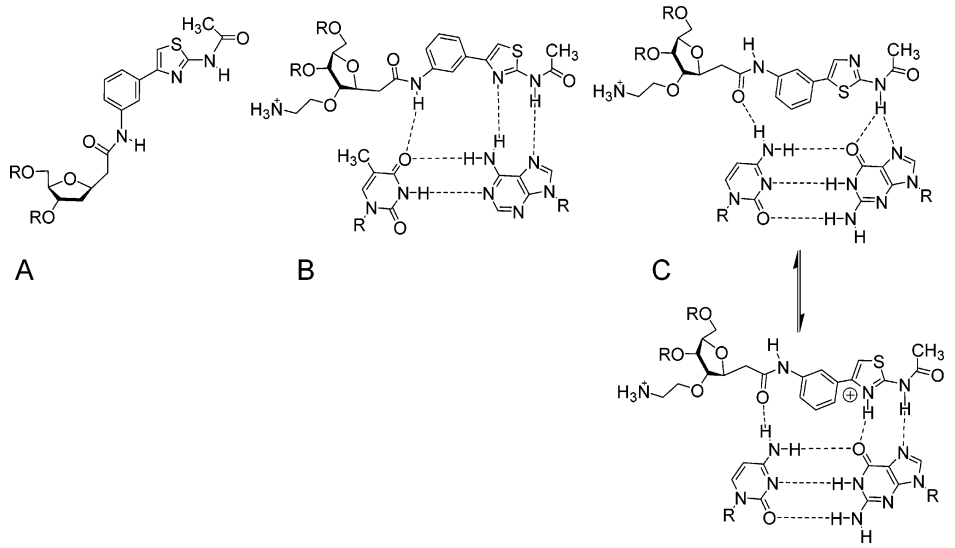
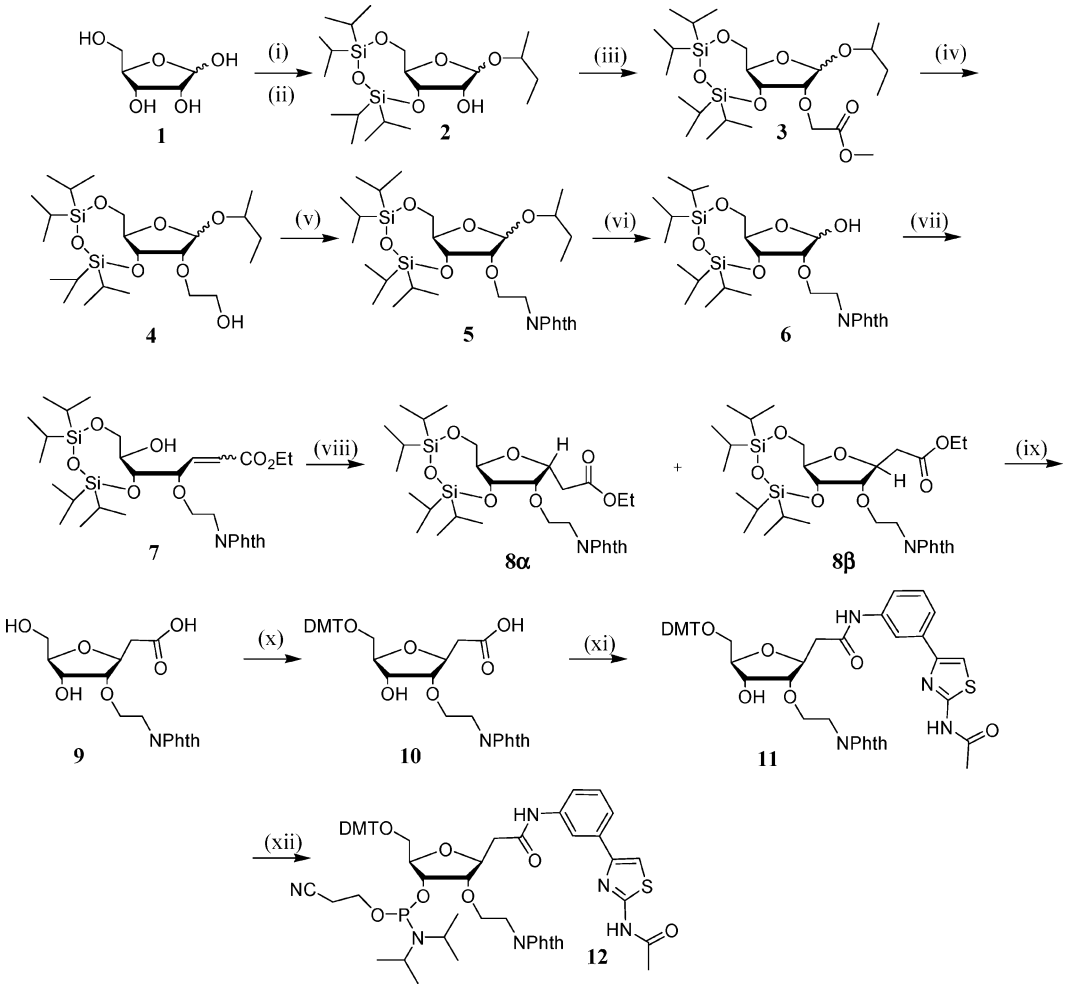


FIGURE 1: Chemical structures of (A) S nucleoside, (B) putative ^{2AE}S.TA triplet, (C) putative ^{2AE}S.CG triplets.

Scheme 1: Synthesis of Phosphoramidite for ^{2AE}S (**12**)[a]



[a] Reagents and conditions: (i) acetyl chloride, butan-2-ol, room temperature, 20 h, quant.; (ii) TIPDSCl (1 equiv), py, rt, 2 h, 74%; (iii) methyl bromoacetate (2.5 equiv), NaH (2.5 equiv), DMF, −10 °C to room temperature, 3.5 h, 80%; (iv) $LiBH_4$ (2.0 equiv), THF, room temperature, 80%; (v) phthalimide (1.1 equiv), PPh_3 (1.4 equiv), DEAD (1.4 equiv), THF, 0 °C−room temperature, 2 h, 73%; (vi) DCM/TFA/H_2O, −15 °C, 3 h, 83%; (vii) ethoxycarbonyl(methylene)triphenyl phosphorane (1.2 equiv), THF, 50 °C, 2 h; then (viii) sodium ethoxide, ethanol (α/β, 1:1) 50%; (ix) concentrated HCl, water, THF, 50 °C, 0.5 h; (x) DMTr-Cl (0.9 equiv), py, room temperature, 5 h, 32%; (xi) *N*-(2-(3-aminophenyl)thiazol-4-yl)acetamide (1.5 equiv), dicyclohexylcarbodiimide (1.5 equiv), DCM, py, butanol, room temperature, 4 h, 80%; (xii) 2-cyanoethoxy-(*N*,*N*-diisopropylamino)chlorophosphine (1.2 equiv), DIPEA (3 equiv), THF, room temperature, 5 h, 84%.

chemical details are presented in the Supporting Information. D-Ribose was first converted into the isobutyl glycoside by stirring at room-temperature overnight in a mixture of acetyl chloride and anhydrous butan-2-ol. The isobutyl group was

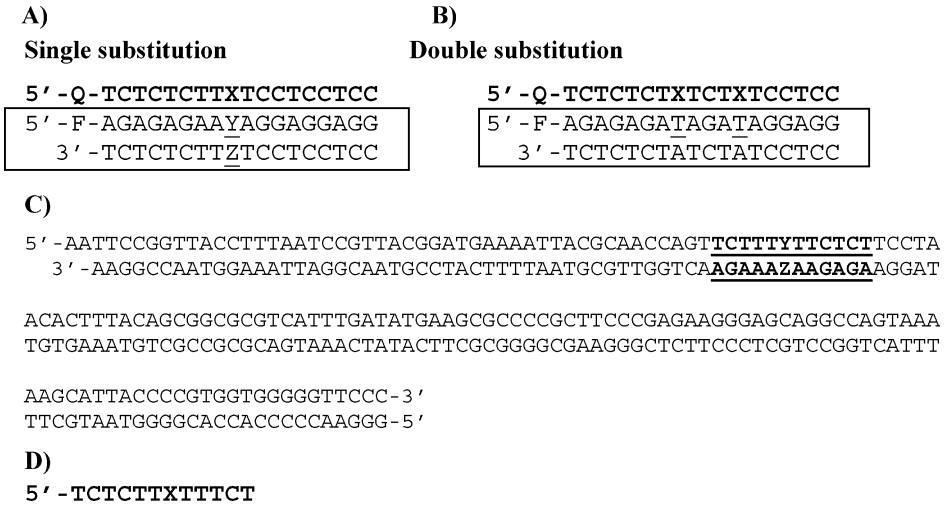


FIGURE 2: (A and B) Sequences of the oligonucleotides used in the fluorescence and UV melting experiments. The duplex target is boxed, while the third strand is shown in bold. For the fluorescence melting experiments, the duplex was labeled with fluorescein at the 5′ end of the purine-rich strand and the third strand was labeled with methyl red L-threoninol at the 5′ end. Identical sequences were used in the UV melting experiments but lacked the fluorophore and quencher. (A) Target sites containing a single pyrimidine inversion. For the duplex, YZ is each base pair in turn (AT, TA, GC, and CG), while for the third strand, X is each base or analogue in turn (A, G, C, T, S, and ^{2AE}S). (B) Target sites containing two TA inversions; in the third strand, X is either G, S, or ^{2AE}S. (C) Sequences of the *tyr*T (43−59) fragments used in the footprinting experiments. The 12 base-pair triplex target site is underlined in bold, and YZ corresponds to each base pair in turn. The DNA was labeled at the 3′ end of the *Eco*RI site (lower strand). (D) Sequence of the triplex-forming oligonucleotide used in the footprinting experiments, where X is each base in turn (G, C, T, S, and ^{2AE}S).

chosen because it can be removed in the presence of the TIPDS group, which was necessary later in the synthesis (*32*). The 3′ and 5′ hydroxyl groups were then protected with TIPDS using the Markiewicz reagent (*33*) in pyridine to afford compound **2** as an anomeric mixture (α/β, approximately 2:1) in 74% yield. Introduction of the protected 2′-aminoethoxy substituent was achieved in three steps. First, alkylation of the 2′-hydroxyl group with an excess of methyl bromoacetate and sodium hydride in DMF at −5 °C gave **3**. This reaction failed to go to completion even with the addition of excess alkylating agent, and unreacted starting material was difficult to remove because of the close polarity of the two compounds. Therefore, compound **3** was used in the next step without further purification. The ester moiety of **3** was reduced with lithium borohydride in THF to give alcohol **4** in good yield, and the hydroxyl group of **4** was then substituted by phthalimide under Mitsunobu conditions (*34*). The phthalimide-protected 2′aminoethoxy moiety was stable for the remaining steps of the synthesis and is compatible with oligonucleotide synthesis and deprotection (*35*). After the procedure of Reese et al. (*32*), the isobutyl group was cleaved to give compound **6** (83%). The aglycon was then attached to the 2′-modified sugar moiety following the protocol of Guianvarc'h et al. (*29*). Thus, treatment of **6** with ethoxycarbonylmethylene triphenylphosphorane (*36*) in THF at reflux yielded **7**, which was then cyclized without isolation to give furanoses $\mathbf{8}_\alpha$ and $\mathbf{8}_\beta$ (α/β, 1:1) (50%). The anomers were separated, and their configurations were assigned by NMR spectroscopy. The β anomer ($\mathbf{8}_\beta$) was treated with concentrated HCl to hydrolyze the ethyl ester moiety and cleave the TIPDS moiety. Acid hydrolysis was chosen because of the presence of the base labile 2′-phthalmide group (it was observed that treatment of $\mathbf{8}_\beta$ with dilute sodium hydroxide results in partial ring open of the phthalimide before the ester group is cleaved). Unpurified compound **9** was protected at the 5′ position using 4,4′-dimethoxytrityl chloride in anhydrous pyridine to afford **10** (32%), which was coupled with *N*-[2-(3-aminophenyl)-thiazol-4-yl]acetamide (*29*) to give **11** (80%). Phosphitylation of **11** using conventional conditions proceeded smoothly to give the ^{2AE}S monomer **12** in high yield.

Preparation of Synthetic Oligonucleotides. All oligonucleotides were synthesized on an Applied Biosystems 394 automated DNA/RNA synthesizer using the standard 0.2 μmol phosphoramidite cycle of acid-catalyzed detritylation, coupling, capping, and iodine oxidation and were purified by reversed-phase HPLC. Further details are provided in the Supporting Information. Purified oligonucleotides were analyzed by MALDI−TOF MS using a ThermoBioAnalysis Dynamo MALDI−TOF mass spectrometer in positive-ion mode (*37*) (Table S1 in the Supporting Information).

Fluorescence Melting Studies. The thermal stability of the intermolecular triplexes was determined with fluorescently labeled oligonucleotides using a Roche LightCycler as previously reported (*38*, *39*). The purine-containing strand of the duplex was labeled at the 5′ end with amidohexyl-fluorescein (6-FAM phosphoramidite, Link Technologies), and the third strand was labeled at the 5′ end with methyl red L-threoninol. These are in close proximity on triplex formation, and the fluorescence is quenched. When the complex is heated, the third strand dissociates and a large increase in fluorescence is observed. In this way, the dissociation of the third strand is observed directly, without interference from dissociation of the duplex. In each case, the quencher was placed on the third strand, so that it could be added in molar excess without increasing the fluorescence signal. All oligonucleotides were prepared in 50 mM sodium acetate buffer at the required pH (5.0, 5.5, or 6.0) containing 200 mM NaCl. The sequences of the oligonucleotides used in these experiments are shown in parts A and B of Figure 2. Melting experiments were performed in a total volume of 20 μL and contained 0.25 μM duplex and 3 μM third strand. The complexes were first denatured by heating to 95 °C and left to equilibrate for 10 min, then cooled to 30 °C (or lower

for experiments conducted at pH 6.0) at a rate 0.2 °C/min, allowed to equilibrate for 10 min, and heated again to 95 °C at 0.2 °C/min. Recordings were taken during both the heating and cooling steps to check for hysteresis. No significant hysteresis was observed with a temperature change of 0.2 °C/min, although the melting and annealing profiles differed by several degrees at faster rates of heating and cooling (0.1 °C/sec) as previous noted for other triplexes (*23*). Although the slowest rate of continuous temperature change in the LightCycler is 0.1 °C/sec, slower melting profiles were obtained by increasing the temperature in 1 °C steps and then leaving the samples to equilibrate for 5 min before recording the fluorescence. The LightCycler has one excitation source (488 nm), and the changes in fluorescence emission were measured at 520 nm. T_m values were determined from the first derivatives of the melting profiles using the Roche LightCycler software. Each reaction was performed in triplicate, and the T_m values usually differed by less than 0.5 °C.

UV Melting Studies. UV melting experiments were performed on a Varian Cary 400 Scan UV−visible spectrophotometer. All oligonucleotides were prepared in 10 mM sodium phosphate buffer (pH 5.8) containing 1 mM EDTA and 200 mM NaCl. The third strand and duplex were mixed in a 5:1 ratio and filtered into Hellma SUPRASIL synthetic quartz cuvettes using Kinesis regenerated cellulose 13 mm filters. The sample chamber was flushed with nitrogen gas to minimize condensation on the cells. For the experiments at pH 5.8, the complexes were initially denatured by heating to 80 °C at 0.5 °C/min, then cooled to 10 °C at a slow rate of 0.1 °C/min, and then heated again to 80 °C at a rate of 0.1 °C/min. Similar experiments were performed at pH 6.5, except the heating and cooling rates were 0.25 °C/min. The change in absorbance at 260 nm was measured during both these annealing and denaturing steps, and T_m values were calculated from the first derivatives. At these rates of heating and cooling, no significant hysteresis was observed between the melting and annealing curves.

Footprinting: DNA Fragments. The *tyr*T(43−59) fragment contains a 17-base oligopurine tract between positions 43−59 (*40*). This was modified to produce four fragments, each containing a different base pair at position ZY (*23*) (see Figure 2C). Radiolabeled fragments were produced by digesting each plasmid with *Eco*RI and *Ava*I and labeling at the 3′ end of the *Eco*RI site using reverse transcriptase and [α-^{32}P]dATP. Each fragment was separated from the remainder of the plasmid DNA on an 8% (w/v) nondenaturing polyacrylamide gel. After elution, the fragment was dissolved in 10 mM Tris-HCl at pH 7.5 containing 0.1 mM EDTA to give about 10 cps/μL as determined on a handheld Geiger counter (<10 nM).

DNase I Footprinting. Radiolabeled DNA (1.5 μL) was mixed with TFOs (3 μL) dissolved in 50 mM sodium acetate at pH 5.0 containing 7.5 mM $MgCl_2$ to produce final oligonucleotide concentrations between 0.3 and 20 μM. The complexes were left to equilibrate at 20 °C overnight. DNase I digestion was carried out by adding 2 μL of DNase I (typically 0.01 units/mL) dissolved in 20 mM NaCl containing 2 mM $MgCl_2$ and 2 mM $MnCl_2$. The reaction was stopped after 1 min by the addition of 4 μL of 80% formamide containing 10 mM EDTA, 10 mM NaOH, and 0.1% (w/v) bromophenol blue. The products of digestion were separated on 9% polyacrylamide gels containing 8 M urea. Samples were heated to 100 °C for 3 min, before rapidly cooling on ice and loading onto the gel. Polyacrylamide gels (40 cm long and 0.3 mm thick) were run at 1500 V for about 2 h and then fixed in 10% (v/v) acetic acid. These were transferred to Whatman 3MM paper and dried under vacuum at 86 °C for 1 h. The dried gels were subjected to phosphorimaging using a Molecular Dynamics Storm phosphorimager.

The intensity of bands within each footprint was estimated using ImageQuant software. These intensities were then normalized relative to a band in the digest, which is not part of the triplex target site and which was not affected by addition of the oligonucleotides. Footprinting plots (*41*) were constructed from these data and fitted using simple binding curves using Sigmaplot for Windows. C_{50} values, indicating the TFO concentration that reduces the band intensity by 50%, were then calculated from these.

RESULTS

The S nucleoside (Figure 1A) was originally designed for triplex recognition of TA interruptions within an oligopurine tract (*29*). Triplex-forming oligonucleotides containing this monomer exhibit good affinity for TA base pairs within parallel DNA triplexes, and the S.TA triplet has been reported to have similar stability to T.AT (*29*−*31*). In the present study, we have examined the affinity and sequence specificity of triplex-forming oligonucleotides that contain its novel 2′-aminoethoxy derivative (^{2AE}S) by fluorescence melting, UV melting, and quantitative DNase I footprinting. These results are compared with those for TFOs that contain S or each of the natural DNA bases.

Fluorescence and UV Melting Studies: Single Incorporation of ^{2AE}S. We examined the stability of the triplexes that contain a single ^{2AE}S residue in the center of the third-strand oligonucleotide by fluorescence melting. Representative fluorescence melting profiles for these complexes are presented in Figure 3. In these studies, ^{2AE}S was investigated against all four base pairs at the central position in the target duplex and the results were compared with S and all of the natural bases at this position, a total of 24 different triplet combinations (X.YZ) at the center of the sequence shown in Figure 2A. Because these target duplexes contain several GC base pairs, each of which is targeted by forming the pH-dependent C^+.GC triplet, these experiments were conducted at low pH values (5.0, 5.5, and 6.0). Triplex formation is kinetically much slower than duplex formation, and significant hysteresis was observed between the melting and annealing profiles at fast rates of heating and cooling (0.1 °C/sec), indicating, as previously described, that the system was not in thermodynamic equilibrium (*23*). No hysteresis was detected at the slower rate of temperature change (0.2 °C/min), and the results of all of these experiments are summarized in Table 1.

Examination of Table 1 reveals that, as expected (*42*−*44*), the most stable triplets with natural bases in the third strand are formed with T against AT, C against GC, G against TA, and C or T against CG. The T_m values of these complexes vary between 67.1 and 57.4 °C at pH 5.0, confirming that C^+.GC is more stable that T.AT and that the triplets involving pyrimidine inversions (T.CG, C.CG,

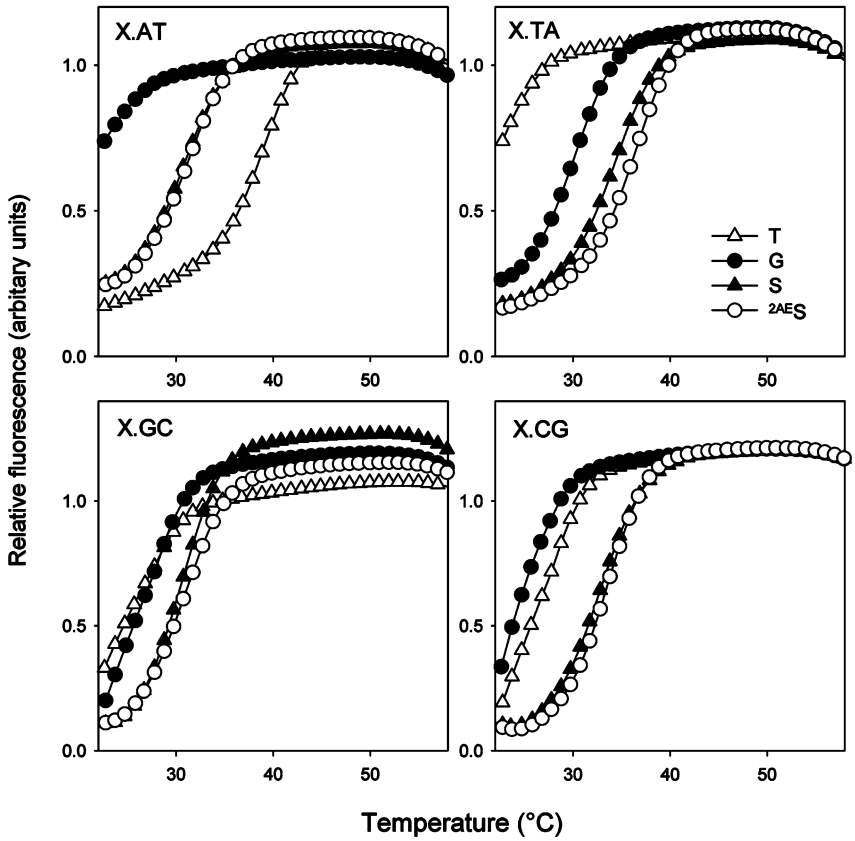


FIGURE 3: Representative fluorescence melting curves showing the interaction of triplex-forming oligonucleotides with target sites containing a variable central base pair (AT, TA, GC, or CG) with a third strand containing T, G, S, and ^{2AE}S in turn (T, △; G, ●; S, ▲; ^{2AE}S, ○). The experiments were performed in 50 mM sodium acetate at pH 6.0 containing 200 mM NaCl. The *y* axis show the normalized fluorescence (arbitrary units), while the *x* axis shows the temperature (°C). The samples were heated at a rate of 0.2 °C/min.

Table 1: T_m Values (°C) Determined for the Fluorescence Melting of the Singly Substituted Intermolecular Triplexes Shown in Figure 2A[a]

	pH	A	G	C	T	S	^{2AE}S
X.AT	5.0	57.1	54.2	54.9	63.8	59.2	59.5
	5.5	43.7	38.2	39.1	54.0	47.2	47.5
	6.0	28.2	<25.0	<25.0	39.4	30.6	31.3
X.TA	5.0	54.4	58.7	53.2	52.7	61.0	62.3
	5.5	39.6	46.7	38.0	38.1	50.5	51.7
	6.0	<25.0	30.6	<25.0	<25.0	34.6	36.5
X.GC	5.0	61.1	57.3	67.1	56.8	60.1	60.3
	5.5	47.3	43.1	56.2	42.9	46.8	47.5
	6.0	30.0	28.0	40.5	27.7	30.4	31.2
X.CG	5.0	54.9	56.3	58.2	57.4	61.1	61.7
	5.5	39.5	41.2	42.5	43.3	49.1	49.3
	6.0	<25.0	<25.0	28.4	28.3	33.0	33.6

[a] The T_m values were determined at pH 5.0, 5.5, and 6.0 as indicated. The duplex concentration was 0.25 μM, while the third strand was 3 μM. Each T_m value is the average of three separate determinations.

and G.TA) are less stable. The difference between C^+.GC and T.AT is less pronounced at higher pH values, but these triplets are still much more stable than those formed at TA and CG. It can also be seen that the four natural bases have very different selectivities for the four possible Watson−Crick base pairs. C and T bind to GC and AT respectively much better than to any other base pairs, while A and G show little discrimination between the different base pairs. This suggests that several of these combinations are effectively "null bases" in which the third-strand base makes little or no contribution to specificity.

S has a greater affinity than G for the TA base pair, with an increase in T_m of 2 °C at pH 5.0 and 4 °C at pH 6.0, confirming previous findings (*29*−*31*). Addition of the 2′-aminoethoxy group to S increases triplex stability at TA by a further 2 °C at pH 6.0. Therefore, of the nucleosides examined in this study, ^{2AE}S produced the most stable triplets with TA base pairs, and these are much more stable than G.TA triplets. The S and ^{2AE}S nucleosides also produced more stable complexes than any of the natural bases with CG base pairs. At pH 5.0, the melting temperatures of ^{2AE}S.TA and ^{2AE}S.CG triplexes are essentially the same and are only marginally higher than ^{2AE}S.GC and ^{2AE}S.AT. Pleasingly, the discrimination between ^{2AE}S.TA and ^{2AE}S.CG improves significantly at pH 6.0, and ^{2AE}S is superior in selectivity to S (ΔT_m = 2.9 °C compared to 1.6 °C). Importantly, at the higher pH values, the ^{2AE}S.TA triplet is only 3 °C less stable than T.AT, compared with G.TA, which is less stable than T.AT by 9 °C.

Because these triplexes contain several C.GC triplets, they become less stable at higher pH values and no complex formation was detected at pH 7.0 because this would have occurred at temperatures below the useful range of the LightCycler (<28 °C). At pH 6.5, complexes containing central T.AT and C.GC triplets could be stabilized by the addition of 10 mM $MgCl_2$ and 1 mM spermine, showing T_m values of 37 and 36 °C, respectively. Under these conditions, no stable triplexes were detected with any other combinations of natural bases although the complexes with S.TA, ^{2AE}S.TA,

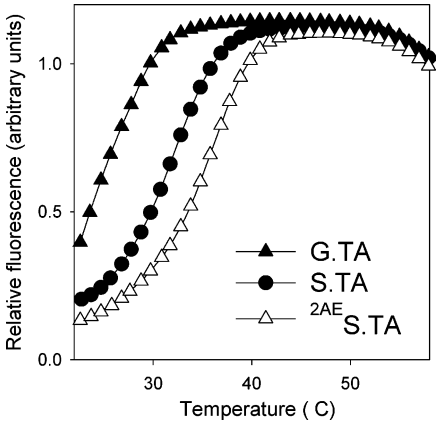


FIGURE 4: Representative fluorescence melting curves showing the interaction of triplex-forming oligonucleotides with the target site containing two TA interruptions. The third strands contained X = G, S, or ^{2AE}S in turn (G, ▲; S, ●; ^{2AE}S, △). The experiments were performed in 50 mM sodium acetate at pH 6.0 containing 200 mM NaCl. The *y* axis show the normalized fluorescence (arbitrary units), while the *x* axis shows the temperature (°C). The samples were heated at a rate of 0.2 °C/min.

Table 2: T_m Values (°C) Determined for the Melting of the Doubly Substituted Intermolecular Triplexes Shown in Figure 2B[a]

	pH	G	S	^{2AE}S
X.TA	5.0	52.3	56.7	58.9
	5.5	40.3	46.3	49.2
	6.0	<25	32.1	36.3

[a] The T_m values were determined by the fluorescence melting technique at pH 5.0, 5.5, and 6.0 as indicated. The duplex concentration was 0.25 μM, while the third strand was 3 μM. Each value is the average of three separate determinations.

S.CG, and ^{2AE}S.CG show T_m values of 29.5, 30.0, 27.8, and 28.1 °C, respectively.

Next, we carried out UV melting experiments to confirm that the quencher employed in the fluorescence melting studies does not influence triplex selectivity. The experiments, which require much larger quantities of unlabeled oligonucleotides, were performed at pH 5.8. The TFO with a central ^{2AE}S residue shows the same relative selectivity as in the fluorescence melting, TA > CG > AT ~ GC, giving melting temperatures of 44, 41, 39, and 38 °C, respectively. In comparison, the oligonucleotide with a G in this position produced a stable complex only when placed opposite TA (forming the G.TA triplet), with a T_m of just 36 °C. As expected, no melting transitions were observed at pH 7.0, although at pH 6.5, T_m values (melting) of 18.7, 15.8, 16.1, and 14.7 °C were determined for ^{2AE}S.TA, ^{2EA}S.CG, S.TA, and S.CG, respectively. These were all more stable than the complex containing a central G.TA triplet (T_m = 11.4 °C).

Triplex Formation at Targets Containing Two TA Interruptions. Similar fluorescence melting studies were also performed with a duplex target containing two TA interruptions (Figure 2B), targeting these positions with third-strand oligonucleotides containing S, ^{2AE}S, and G. Representative fluorescence melting curves for the data obtained at pH 6.0 are shown in Figure 4, and the T_m values are summarized in Table 2. At pH 6.0, the presence of two G.TA triplets is sufficiently destabilizing to prevent triplex formation above 28 °C (note that a single G.TA triplet gave a T_m of only 30.6 °C). In contrast, the complex with two S.TA triplets

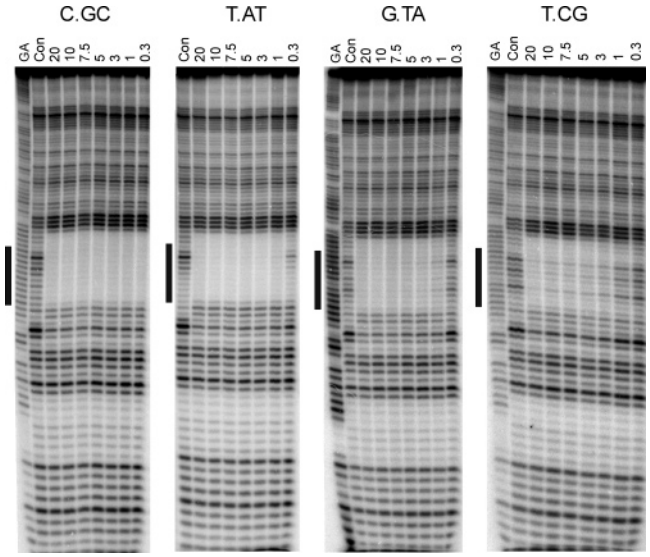


FIGURE 5: Representative DNase I footprinting experiments showing the interaction of the best unmodified triplex-forming oligonucleotides with DNA fragments derived from *tyr*T(43−59). The fragments contain different base pairs at the center (Figure 2C) (GC, AT, TA, and CG in each panel in turn). The interaction is shown for oligonucleotides 5′-TCTCTTXTTTCT-3′, where X is C, T, G, or T with the different duplex targets generating C.GC, T.AT, G.TA, and T.CG triplets, respectively. The lanes labeled "GA" are Maxam−Gilbert markers specific for purines, while "con" indicates DNase I cleavage in the absence of an added oligonucleotide. The oligonucleotide concentration (μM) is shown at the top of each gel lane. The experiments were performed in 50 mM sodium acetate at pH 5.0 containing 5 mM magnesium chloride, and the complexes were left to equilibrate overnight at 20 °C. The filled boxes show the position of the triplex target site.

displayed a T_m of 32.1 °C, and the stability was further enhanced by incorporation of two ^{2AE}S residues (T_m = 36.3 °C). In this general sequence context, there is little difference in stability between the complexes with one and two ^{2AE}S.TA triplets. The differences between ^{2AE}S and S were again greater at pH 6.0, showing an increase in T_m of about 2 °C for each addition of the ^{2AE}S monomer relative to S.

DNase I Footprinting Studies. The stability and selectivity of triplexes containing S and ^{2AE}S was also examined by DNase I footprinting. The interaction of 12-mer oligonucleotides of sequence TCTCTTXTTTCT (X = S, ^{2AE}S, G, C, and T) with duplex targets containing the oligopurine tract 5′-AGAAAZAAGAGA, where Z is each base in turn opposite the third-strand position X were measured by quantitative DNase I footprinting (Figure 2C). Representative footprinting gels for these interactions are shown in Figures 5−7, for experiments performed in 50 mM sodium acetate (pH 5.0) containing 5 mM $MgCl_2$. Footprinting gels for the most stable complex with each fragment, using third-strand oligonucleotides containing only natural DNA bases (forming T.AT, C^+.GC, G.TA, and T.CG triplets at the central position X.ZY) are shown in Figure 5. Footprints for oligonucleotides containing S or ^{2AE}S at the central position against all four DNA fragments are shown in Figures 6 and 7, respectively. C_{50} values, (the oligonucleotide concentration that reduces the intensity of bands in the footprint by 50%) are shown in Table 3. The oligonucleotides containing the best natural bases in the central position all produced footprints at micromolar concentrations, although the affinity (C_{50}) varied according to the identity of the central triplet in the order

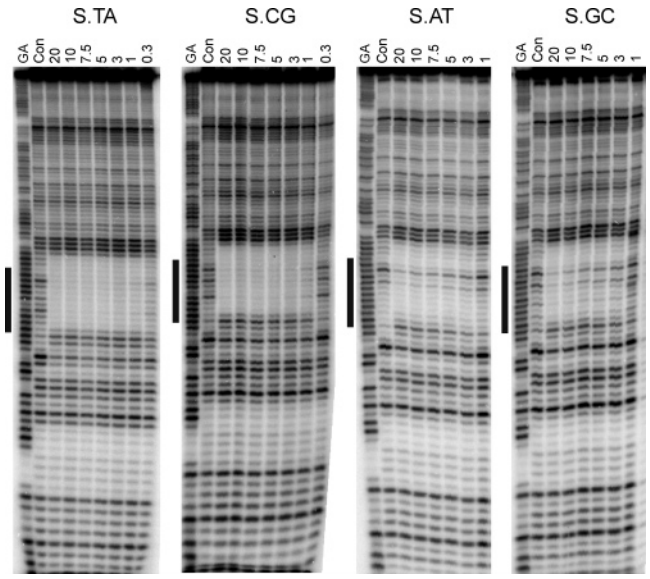


FIGURE 6: DNase I footprinting experiments showing the interaction of oligonucleotide 5′-TCTCTTSTTTCT-3′ with DNA fragments derived from *tyr*T(43−59) containing each base pair in turn at the center of the oligopurine tract, generating the central triplex S.TA, S.CG, S.AT, and S.GC in turn. The lanes labeled "GA" are Maxam−Gilbert markers specific for purines, while "con" indicates DNase I cleavage in the absence of added oligonucleotide. The oligonucleotide concentration (μM) is shown at the top of each gel lane. The experiments were performed in 50 mM sodium acetate at pH 5.0 containing 5 mM magnesium chloride, and the complexes were left to equilibrate overnight at 20 °C. The filled boxes show the position of the triplex target site.

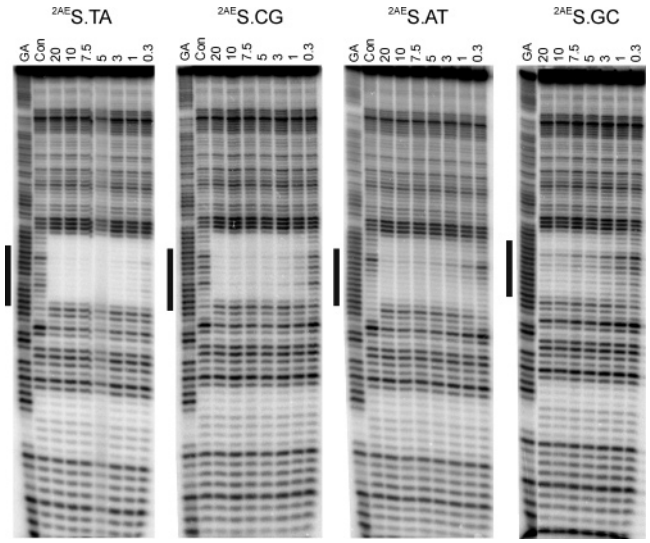


FIGURE 7: DNase I footprinting experiments showing the interaction of oligonucleotide 5′-TCTCTT^{2AE}STTTCT-3′ with DNA fragments derived from *tyr*T(43−59) containing each base pair in turn at the center of the oligopurine tract, generating the central triplex ^{2AE}S.TA, ^{2AE}S.CG, ^{2AE}S.AT, and ^{2AE}S.GC in turn. The lanes labeled "GA" are Maxam−Gilbert markers specific for purines, while "con" indicates DNase I cleavage in the absence of added oligonucleotide. The oligonucleotide concentration (μM) is shown at the top of each gel lane. The experiments were performed in 50 mM sodium acetate at pH 5.0 containing 5 mM magnesium chloride, and the complexes were left to equilibrate overnight at 20 °C. The filled boxes show the position of the triplex target site.

T.AT = C$^+$.GC > G.TA > T.CG. Incorporation of a G in the center of the third strand showed about 10-fold selectivity for TA over the next best triplet (G.GC); similarly, a central C (for which the footprint at central GC was too tight to measure accurately) was approximately 10-fold better at GC than at CG and T and was about 7-fold selective for AT over CG. As expected, several base combinations at this central position failed to produce footprints (G.AT, G.CG, C.AT, C.TA, and T.TA). In contrast, both S and ^{2AE}S produced DNase I footprints at all four target sites, although with different affinities. S showed a 5-fold selectivity for TA over CG, and the complex containing an S.TA triplex was about twice as stable as G.TA. Addition of the 2′-aminoethoxy group further increased the affinity for all of the targets, although the increase in affinity was greatest for the targets with a central TA and AT. Again ^{2AE}S.TA is more stable than G.TA. ^{2AE}S also appears to be able to discriminate between TA and CG with at least a 4-fold difference in C_{50} values.

Table 3: C_{50} Values (μM) Determined from Quantitative DNase I Footprinting Experiments for the Interaction of Triplex-Forming Oligonucleotides 5′-TCTCTTXTTTCT (X = G, C, T, S, or ^{2AE}S) with Target Duplexes Containing the Sequence 5′-AGAAAZAAGAGA/5′-TCTCTTZTTTCT (YZ = AT, TA, GC, or CG)[a]

	YZ			
X	AT	TA	GC	CG
G	nd	0.3 ± 0.1	2.5 ± 0.5	nd
C	nd	nd	<0.1	0.8 ± 0.3
T	<0.1	nd	2.1 ± 0.3	0.6 ± 0.1
S	2.1 ± 0.3	0.11 ± 0.05	2.1 ± 0.7	0.6 ± 0.2
^{2AE}S	0.8 ± 0.1	<0.1	1.2 ± 0.2	0.5 ± 0.1

[a] The experiments were performed in 50 mM sodium acetate at pH 5.0 containing 200 mM NaCl and 5 mM $MgCl_2$. C_{50} values less than 0.1 μM could not be accurately determined. nd indicates that no footprint was detected.

DISCUSSION

The studies presented in this paper demonstrate that addition of a 2′-aminoethoxy group to the nucleoside S, generating ^{2AE}S, enhances triplex stability at oligopurine tracts that contain TA inversions. Several studies have now shown the importance of the 2′-aminoethoxy group for stabilizing triplexes, and it has previously been added to T (*24*−*26*) and 5-propargylamino-dU (*22*, *23*) for recognition of AT and to 5-methyl-pyrimidin-2-one (*27*, *28*) for recognition of CG. By analogy with ^{2AE}T, it seems likely that the additional triplex stabilization conferred by the 2′-aminoethoxy group is due to interactions between the protonated amine and the phosphodiester backbone of the duplex (*24*, *25*). Addition of the 2′-aminoexthoxy group will also give the nucleotide an RNA-like configuration, which has been shown to enhance triplex stability (*45*).

In the hydrogen-bonding scheme proposed by Guinvarc'h et al. (*29*−*31*), S recognizes a TA base pair by forming three hydrogen bonds (Figure 1B), one to O4 of thymine and two to adenine (N7 and N6). The relatively high T_m values of the S.TA and ^{2AE}S.TA triplets support this suggestion. The ^{2AE}S.TA triplet provides the most stable means of targeting TA base pairs reported to date; one ^{2AE}S.TA triplet increases the T_m by 4−6 °C relative to G.TA, and it is 1−2 °C more stable than S.TA. This improved triplet has a similar stability to T.AT but produces less stable triplexes at low pH than a C$^+$.GC triplet located in the equivalent position.

Although ^{2AE}S is very effective in stabilizing TA interruptions, this unnatural nucleoside has limited selectivity at

low pH. At pH 5.0, in the sequence context examined in this study, there is only a difference of 3 °C between the most stable (^{2AE}S.TA) and least stable (^{2AE}S.AT) triplets, although this increases to a respectable 5 °C at pH 6.0. This compares with a 14 °C difference between the recognition of GC and TA by C at pH 5.0. It is clearly difficult to achieve simultaneous affinity and selectivity by triplex formation at pyrimidine interruptions, and there remains a scope for developing further base analogues for this purpose. This result is similar to that observed with the bisamino analogue of T (BAU) (*23*), for which addition of the 2′-amino group changed both the affinity and selectivity.

The fluorescence melting and footprinting experiments demonstrate that both ^{2AE}S and S can form triplexes with other base pairs as well as TA, especially CG. It should be noted that S is structurally related to D_3 (*13*, *14*), which was also designed for specific recognition of TA base pairs but was subsequently found to intercalate at YpR steps instead of binding within the major groove (*14*, *15*). Although S could potentially bind in the same manner, a study on nearest neighbor effects suggested that this was not the case (*30*, *31*). However, the conformational flexibility of S may permit hydrogen-bond formation with other base pairs in addition to TA, and schemes for putative S.CG triplets are shown in Figure 1C. In one of these (upper diagram), S forms a hydrogen bond to N4 of cytosine and a bifurcated hydrogen bond to O6 and N7 of guanine, although there are no hydrogen bonds to the thiazole ring. An alternative structure for recognition of CG, which involves protonation of the thiazole nitrogen and consequent hydrogen bonding to O6 of guanine, is shown in the lower part of Figure 2C. The latter triplet, which could exist at low pH, is likely to be the more stable, explaining why S and ^{2AE}S show greater selectivity for TA at elevated pH values. However, detailed structural studies will be necessary to properly elucidate the binding modes of S and ^{2AE}S.

Although ^{2AE}S provides the most stable means for recognizing TA interruptions, the ^{2AE}S.TA triplet is still less stable than T.AT or C^+.GC. This may seem surprising because this analogue is able to form three hydrogen bonds to the TA base pair (Figure 1B), as well as possessing the stabilizing positive charge on the sugar. This suggests that the arrangement of hydrogen-bonding groups in S or its stacking within the third strand is still not optimal. There is still therefore a need for further nucleotides capable of recognizing TA with high affinity and specificity.

In conclusion, ^{2AE}S confers greater triplex stability and selectivity than S. We anticipate that its affinity will be further increased if ^{2AE}S is incorporated in TFOs in which the other nucleosides are also modified with 2′-aminoethoxy groups (*26*, *27*).

SUPPORTING INFORMATION AVAILABLE

Full details of the chemical synthesis and characterization of the phophoramidite for ^{2AE}S and oligonucleotide preparation and characterization. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Hélène, C., and Toulmé, J. (1990) Specific regulation of gene expression by antisense, sense, and antigene nucleic acids, *Biochim. Biophys. Acta 1049*, 99−125.
2. Fox, K. R. (2000) Targeting DNA with triplexes, *Curr. Med. Chem. 7*, 7−17.
3. Buchini, S., and Leuman, C. J. (2003) Recent improvements in antigene technology, *Curr. Opin. Chem. Biol. 7*, 717−726.
4. Seidman, M. M., and Glazer, P. (2003) The potential for gene repair via triple helix formation, *J. Clin. Invest. 112*, 487−494.
5. Felsenfield, G., Davies, R., and Rich, A. (1957) Formation of a three stranded polynucleotide molecule, *J. Am. Chem. Soc. 79*, 2023−2024.
6. Gowers, D. M., and Fox, K. R (1999) Towards mixed sequence recognition by triple helix formation, *Nucleic Acids Res. 27*, 1569−1577.
7. Griffin, L., and Dervan, P. (1989) Recognition of thymine.adenine base pairs by guanine in a pyrimidine triple helix motif, *Science 245*, 967−971.
8. Fossella, A., Kim, Y. J., Shih, H., Richards, E. G., and Fresco, J. R. (1993) Relative specificities in binding of Watson−Crick base pairs by third strand residues in a DNA pyrimidine triplex motif, *Nucleic Acids Res. 21*, 4511−4515.
9. Radhakrishnan, I., and Patel, D. J. (1994) Solution structure of a pyrimidine.purine.pyrimidine DNA triplex containing T.AT, C.GC, and G.TA triples, *Structure 15*, 17−32.
10. Gowers, D. M., and Fox, K. (1997) DNA triple helix formation at oligopurine sites containing multiple contiguous pyrimidines, *Nucleic Acids Res. 25*, 3787−3794.
11. Soto, M., and Marky, L. A. (2002) Thermodynamic contributions for the incorporation of GTA triplets within canonical TAT/TAT and CGC/CGC base-triplet stacks of DNA triplexes, *Biochemistry 41*, 12475−12482.
12. Jiang, L., and Russu, I. M. (2001) Proton exchange and local stability in a DNA triple helix containing a G.TA triad, *Nucleic Acids Res. 29*, 4231−4237.
13. Griffin, L. C., Kiessling, L. L., Beal, P. A., Gillespie, P., and Dervan, P. (1992) Recognition of all four base pairs of double helical DNA by triple helix formation, *J. Am. Chem. Soc. 114*, 7976−7982.
14. Koshlap, K. M, Gillespie, P., Dervan, P., and Feignon, J. (1993) Nonatural deoxyribonucleoside D_3 incorporated in an intramolecular DNA triplex binds sequence specifically by intercalation, *J. Am. Chem. Soc. 115*, 7908−7909.
15. Wang, E., Koshlap, K. M., Gillespie, P., Dervan, P. B., and Feigon, J. (1996) Solution structure of a pyrimidine.purine.pyrimidine triplex containing the sequence-specific intercalating non-natural base D_3, *J. Mol. Biol. 257*, 1057−1069.
16. Mokhir, A. A., Connors, W. H., and Richert, C. (2001) Synthesis and monitored selection of nucleotide surrogates for binding T:A base pairs in homopurine-homopyrimidine DNA triple helices, *Nucleic Acids Res. 29*, 3674−3684.
17. Eldrup, A. B., and Nielson, P. E. (1997) A novel peptide nucleic acid monomer for recognition of thymine in triple-helix structures, *J. Am. Chem. Soc. 119*, 11116−11117.
18. Olsen, A., Dahl, O., and Nielsen, P. E. (2004) Synthesis and evaluation of a conformationally constrained pyridazinone PNA-monomer for recognition of thymine in triple-helix structures, *Bioorg. Med. Chem. Lett. 22*, 1551−1554.
19. Li, J. S., Fan, Y., Zhang, Y., Marky, L. A., and Gold, B. (2003) Design of triple helix forming C-glycoside molecules, *J. Am. Chem. Soc. 125*, 2084−2093.
20. Li, J. S., Shikiya, R., Marky, L. A., and Gold, B. (2004) Triple helix forming TRIPside molecules that target mixed purine/pyrimidine DNA sequences, *Biochemistry 43*, 1440−1448.
21. Sasaki, S., Taniguchi, Y., Takahashi, R., Senko, Y., Kodama, K., Nagatsugi, F., and Maeda, M. (2004) Selective formation of stable triplexes including a TA or a CG interrupting site with new bicyclic nucleoside analogues (WNA), *J. Am. Chem. Soc. 126*, 516−528.
22. Sollogoub, M., Darby, R. A. J., Cuenoud, B., Brown, T., and Fox, K. (2002) Stable DNA triple helix formation using oligonucleotides containing 2′-aminoethoxy,5-propargylamino-U, *Biochemistry 41*, 7224−7231.
23. Osborne, S. D., Powers, V. E. C., Rusling, D. A., Lack, O., Fox, K. R., and Brown, T. (2004) Selectivity and affinity of triplex-forming oligonucleotides containing 2′-aminoethoxy-5-(3-amino-prop-1-ynyl)uridine for recognising AT base pairs in duplex DNA, *Nucleic Acids Res. 32*, 4439−4447.
24. Blommers, M. J. J., Natt, F., Jahnke, W., and Cuenoud, B. (1998) Dual recognition of double stranded DNA by 2′-aminoethoxy-modified oligonucleotides: The solution structure of an intramo-

lecular triplex obtained by NMR spectroscopy, *Biochemistry 37*, 17714−17725.
25. Cuenoud, B., Casset, F., Husken, D., Natt, F., Wolf, R. W., Altmann, K., Martin, P., and Moser, H. E. (1998) Dual recognition of double-stranded DNA by 2′-aminoethoxy-modified oligonucleotides, *Angew. Chem., Int. Ed. 37*, 1288−1291.
26. Puri, N., Majumdar, A., Cuenoud, B., Miller, P. S., and Seidman, M. M. (2004) Importance of clustered 2′-*O*-(2-aminoethyl) residues for gene targeting activity of triple helix formation, *Biochemistry 43*, 1343−1351.
27. Buchini, S., and Leumann, C. J. (2003) Dual recognition of a C−G pyrimidine-purine inversion site: Synthesis and binding properties of triplex-forming oligonucleotides containing 2′-aminoethoxy-5-methyl-H1-pyrimidin-2-one ribonucleosides, *Tetrahedron Lett. 44*, 5065−5068.
28. Buchini, S., and Leumann, C. J. (2004) Stable and selective recognition of three base pairs in the parallel triple-helical DNA binding motif, *Angew. Chem., Int. Ed. 43*, 3925−3928.
29. Guinvarc'h, D., Fourrey, J., Maurisse, R., Sun, J., and Benhida, R. (2001) Incorporation of a novel nucleobase allows stable oligonucleotide-directed triple helix formation at the target sequence containing a purine.pyrimidine interruption, *Chem. Commun.* 1814−1815.
30. Guinvarc'h, D., Fourrey, J., Maurisse, R., Sun, J., and Benhida, R. (2002) Synthesis, incorporation into triplex forming oligonucleotide, and binding properties of a novel 2′-deoxy-C-nucleoside featuring a 6-(thiazolyl-5)-benzimidazole nucleobase, *Org. Lett. 4*, 4209−4212.
31. Guinvarc'h, D., Fourrey, J., Maurisse, R., Sun, J., and Benhida, R. (2003) Design of nucleobase for the recognition of the AT inversion by triplex-forming oligonucleotides; a structure−stability relationship study and neighbour bases effects, *Biorg. Med. Chem. 11*, 2751−2759.
32. Reese, C. B., and Wu, Q. (2003) Conversion of 2-deoxy-D-ribose into 2-amino-5-(2-deoxy-β-D-ribofuanosyl)pyridine, 2′-deoxypseudouridine, and other C-(2-deoxyribonucleosides), *Org. Biomol. Chem. 1*, 3160−3172.
33. Markiewicz, W. T., Padyukova, N. S., Samek, Z., and Smrt, J. (1980) The reaction of 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane with cytosine arabinoside and 1-(6-deoxy-a-L-talofuranosyl)-uracil, *Collect. Czech Chem. Commun. 40*, 1869−1865.
34. Mitsunobo, O. (1981) The use of diethyl azodicarboxylate and triphenylphosphine in synthesis and transformation of natural-products, *Synthesis 1*, 1−14.
35. Kume, A., Sekine, M., and Hata, T. (1982) Phthaloyl group: A new amino protecting group of deoxyadenosine in oligonucleotide synthesis, *Tetrahedron Lett. 23*, 4365−4368.
36. Hovinen, J., and Salo, H. (1997) C-glycoside phosphoramidate building block for versatile functionalization of oligodeoxyribonucleotides, *J. Chem. Soc., Perkin Trans. 1*, 3017−3020.
37. Langley, G. J., Herniman, J. M., Davies, N. L., and Brown T. (1999) Simplified sample preparation for the analysis of oligonucleotide by matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry, *Rapid Commun. Mass Spectrom. 13*, 1717−1723.
38. Darby, R. A. J., Sollogoub, M., McKeen, C., Brown, L., Risitano, A., Brown, N. M., Barton, C., Brown, T., and Fox, K. R. (2002) High throughput measurement of duplex, triplex, and quadruplex melting curves using molecular beacons and a LightCycler, *Nucleic Acids Res. 30*, e39.
39. James, P., Brown, T., and Fox, K. R. (2003) Thermodynamic and kinetic stability of intermolecular triple helices containing different portions of C^+.GC and T.AT triplets, *Nucleic Acids Res. 31*, 5598−5506.
40. Brown, P. M., Madden, C. A., and Fox, K. R. (1998) Triple helix formation at different positions on nucleosomal DNA, *Biochemistry 37*, 16139−16151.
41. Dabrowiak, J. C., and Goodisman, J. (1989) Quantitative footprinting analysis of drug−DNA interactions, in *Chemistry and Physics of DNA−Ligand Interactions* (Kallenbach, N. R., Ed.) pp 143−174, Adenine Press, New York.
42. Yoon, K., Hobbs, C. A., Koch, J., Sardaro, M., Kutny, R., and Weis, A. L. (1992). Elucidation of the sequence-specific third strand recognition of four Watson−Crick base pairs in a pyrimidine triple helix motif: T.AT, C.GC, G.TA, and T.CG, *Proc. Natl. Acad. Sci. U.S.A. 89*, 3840−3844.
43. Chandler, S. P., and Fox, K. R. (1993) Triple helix formation at $T_8XT_8.A_8XA_8$, *FEBS Lett. 332*, 189−192.
44. Fosella, A., Kim, Y. J., Shih, H., Richards, E. G., and Fresco, J. R. (1993) Relative specificities in binding of Watson−Crick base pairs by third strand residues in a DNA pyrimidine triplex motif, *Nucleic Acids Res. 21*, 4511−4515.
45. Bernal-Méndez, C., and Leumann, C. J. (2002) Stability and kinetics of nucleic acid triplexes with chimaeric DNA/RNA third strands, *Biochemistry 41*, 12343−12340.

BI050013V