Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Bioorganic & Medicinal Chemistry 12 (2004) 1517–1526
Bioorganic & Medicinal Chemistry

Synthesis and duplex DNA recognition studies of oligonucleotides containing a ureido isoindolin-1-one homo-*N*-nucleoside. A comparison of host–guest and DNA recognition studies

Eric Mertz, Sebastiano Mattei and Steven C. Zimmerman*

Department of Chemistry, University of Illinois at Urbana-Champaign, 600 South Matthews Avenue, Urbana, IL 61801, USA

Received 5 August 2003; accepted 18 December 2003

Abstract—In an effort to construct non-natural bases to be used in triplex-based antigene DNA recognition strategies, a uriedo-isoindolin-1-one homo-*N*-nucleoside base was designed to bind the cytosine-guanine (CG) base pair. An organic soluble analogue of this base was shown to effectively complex CG ($K_{assoc} = 740\ M^{-1}$) in chloroform through formation of three hydrogen bonds (Mertz, E.; Mattei, S.; Zimmerman, S. C. *Org. Lett.* **2000**, *2,* 2931–2934). The novel nucleoside base was synthesized and successfully incorporated into oligonucleotides which were used in triple helix melting temperature studies. Low melting temperatures were observed when the isoindolin-1-one base was paired opposite CG as well as GC, TA, and AT, thus indicating that despite favorable recognition in model studies, the artificial base did not effectively recognize duplex DNA to form pyrimidine-purine-pyrimidine type triple helices.

1. Introduction

Formation of DNA triple helices is a potentially powerful approach to sequence-specific DNA binding.[1] Indeed, binding of a certain single stranded oligonucleotide (triplex-forming oligonucleotide or TFO) within the major groove of the target duplex DNA can achieve regulation of gene expression,[2] site-directed DNA modification,[3] and controlled DNA cleavage.[4,5] Despite these remarkable accomplishments, the wider use of the triplex strategy is hampered by the fact that none of the four natural bases recognize the pyrimidine bases thymine (T) and cytosine (C) in the pyrimidine-purine-pyrimidine (pyr-pur-pyr) motif.[6] Thus, formation of the thymine-adenine-thymine (T-AT) and cytosine-guanine-cytosine (*C*-GC) base triplets allows triplex recognition to occur only over homopurine tracts of duplex DNA. (Fig. 1; *C* denotes 2′-deoxycytidine protonated at the 3-position nitrogen).

To expand the scope of triple helix technology, intense efforts have focused on the design and synthesis of artificial nucleoside bases for the selective recognition of the CG and TA base pairs.[6b,7,8] These non-natural bases have ranged from modified natural bases to entirely synthetic receptors. However, once incorporated into oligonucleotide sequences, these novel bases have typically exhibited at best moderate affinity and selectivity for their intended base pair targets.

Three key considerations in the design of CG or TA selective nucleosides are (1) optimization of hydrogen bonding between the artificial base and the targeted base pair, (2) maximization of π-stacking between the artificial base and the adjacent base triplets, and 3) isomorphism between the non-natural base triplet and the canonical T-AT and *C*-GC base triplets.[1d,e] In developing new artificial nucleoside bases for the recognition of the CG base pair, we addressed the first two of these criteria by designing aromatic, heterocyclic bases capable of forming three hydrogen bonds with the major groove (Hoogsteen) side of the CG base pair.[9] The third design criterion (isomorphism) was addressed by comparing two-dimensional and three-dimensional models of proposed base triplets with the canonical T-AT and *C*-GC triplets. In this analysis, the position of the 2′-deoxyribose group of the artificial base should match closely with the position of the 2′-deoxyribose group in the TFO strand of T-AT or *C*-GC.

Keywords: DNA Recognition; Nucleoside; Triplex DNA.
* Corresponding author. Tel.: +1-121-733-36655; fax: +1-121-724-49919; e-mail: scz@scs.uiuc.edu

0968-0896/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2003.12.022

We have investigated the heterocyclic hydrogen bonding modules **1**–**3** as receptors for the CG base pyr-pur-pyr type triple helices (Fig. 2).[9,10] Prior to incorporation of each base into oligonucleotides, the ability of **1**–**3** to effectively complex the CG base pair in isolation from the constraints of the DNA duplex was measured through ^{1}H NMR studies in an organic solvent. Other groups have subsequently used this method to evaluate the hydrogen bonding of other base pair receptors.[11]

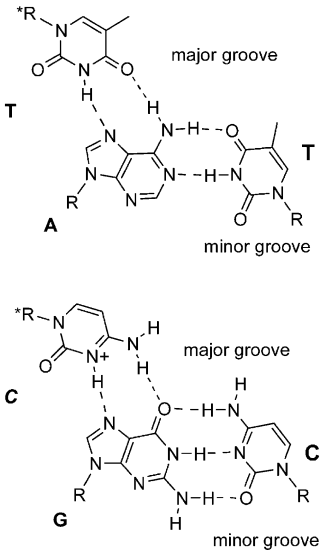


Figure 1. Canonical base triplets in the pyrimidine-purine-pyrimidine (pyr-pur-pyr) triple helix motif. (R = 2′-deoxyribose of Watson and Crick strands and *R = 2′-deoxyribose of the TFO.)

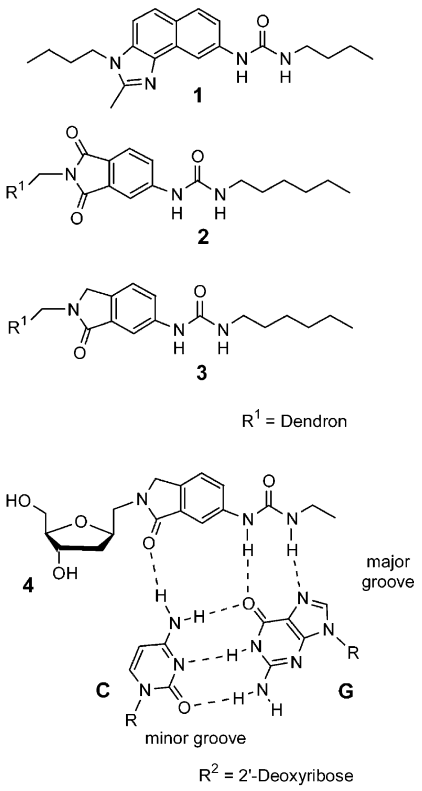


Figure 2. (a) Structures of non-natural receptors designed to form three hydrogen bonds with the major groove side of the CG base pair. (b) Isoindolin-1-one homo-*N*-nucleoside **4** complexed to CG.

The ureido naphthimidazole base **1** was found to complex CG with an association constant (K_{assoc}) of 160 M^{-1} in $CDCl_3$.[9] The complexes of the ureidophthalimide base **2** and the ureidoisoindolin-1-one base **3** with CG were found to be more robust with K_{assoc} values of 1000 M^{-1} and 740 M^{-1}, respectively.[10,12] Unfortunately, the 2′-deoxyribose nucleoside derivative of **2** was found to be unstable to the conditions needed for incorporation of this base into oligonucleotides.[13] Incorporation of a nucleoside analogue of isoindolin-1-one **3** into a TFO was more successful. Herein we report the synthesis of a ureido isoindolin-1-one homo-*N*-nucleoside **4**. This artificial nucleoside was integrated into oligonucleotide sequences, its ability to bind CG within a pyr-pur-pyr triple helix was determined using thermodenaturation studies, and the results compared to the previously reported host–guest, model studies.

The homo-*N*-nucleoside **4** is an 'extended nucleoside' containing an extra methylene group between its hydrogen bonding, heterocyclic portion and the 2′-deoxyribose moiety. Such homo-*N*-nucleoside analogues of the natural DNA bases have been investigated by both Beaucage and Herdwijn.[14] To date, base **4** is the first example of the homo-*N*-nucleoside design being employed in a triplex base pair recognition scheme. However, several of the most promising non-natural nucleosides designed for recognition of CG and TA have also possessed modified sugar groups.[15] In one of these examples, Fourrey and Sun's amide-linked thiazole *C*-nucleoside receptor for TA, an extended base design was employed similar to the design of **4**. The intervening methylene group in **4** makes it a longer nucleoside than the corresponding *N*-nucleoside **5**, providing **4**-CG with improved isomorphism with the natural base triplets (Fig. 3). A simple two-dimensional

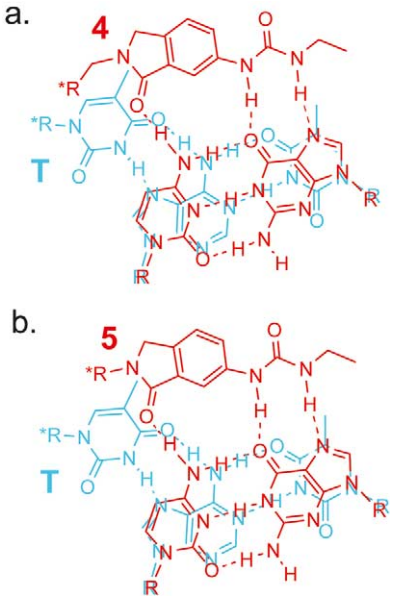
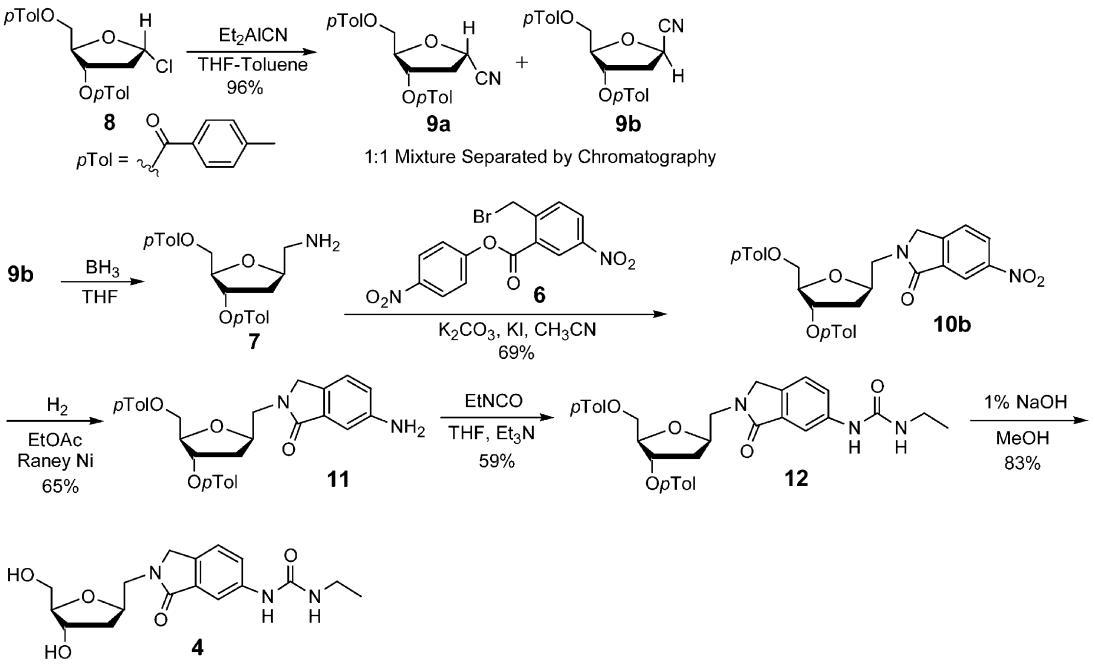


Figure 3. (a) Two-dimensional overlay of the **4**-CG base triplet (red) with the T-AT base triplet (blue). (b) Two-dimensional overlay of the *N*-nucleoside containing **5**-CG base triplet (red) with the T-AT (blue). (R = 2′-deoxyribose of Watson and Crick strands and *R = 2′-deoxyribose of the TFO.)



Scheme 1.

overlap of the **4**-CG triplet with the T-AT base triplet confirms this assessment as the 2′-deoxyribose moiety of **4** extends more closely to the position of the TFO 2′-deoxyribose of a natural base triplet than is possible with base **5**. Simple three-dimensional molecular modeling studies verified the improved isomorphism obtained by employing the homo-*N*-nucleoside design.[16]

2. Results and discussion

The isoindolin-1-one ring system in **4** was constructed using a key lactamization reaction involving reaction of 4-nitrophenyl 6-(bromomethyl)-3-nitrobenzoate (**6**) with aminomethyl-*C*-glycoside **7** (Scheme 1).[17] The base precursor **6** is available in two steps from methyl 6-methyl-3-nitrobenzoate,[18] and its synthesis has been described in our previous communication.[10] The 2′-deoxyribose derivative **7** was synthesized starting from 3′,5′-di-*O*-toluoyl protected 1′-chlorodeoxyribose **8**.[19] Treatment of **8** with diethyl aluminum cyanide according to the method of Beaucage resulted in a 1:1 mixture of nitriles **9a** and **9b**.[14a] To obtain the β-anomer of the desired homo-*N*-nucleoside, this anomeric mixture of was separated by flash chromatography. The α- and β-anomers were each identified by comparison of previously reported melting point and ^{1}H NMR data.[20] Reduction of **9b** to amine **7** was accomplished using borane-THF.[21]

Treatment of **7** with the base precursor **6** and base resulted in the desired lactamization reaction affording the isoindolin-1-one derivative **10b**. Thorough characterization of **10b** using ^{1}H NMR, ^{13}C NMR, COSY, and HMQC supported formation of the desired lactam. Difference NOE studies were used to unambiguously confirm the anomeric assignment of **10b** (Fig. 4). Thus, irradiation of H-2′ in **10b** produced a significant NOE in only one proton (H-3′), while irradiation of H-2″ produced a significant NOE in H-1′ only. These results are consistent with the β-anomer. For comparison, NOE studies were also performed on the α-anomer **10a**, which was synthesized from **9a** using the chemistry described above. Irradiation of H-2′ in **10a** gave a network of significant NOE values between H-2′, H-3′, and H-1′. This observation is consistent with the α-anomer as H-3′, H-2′, and H-1′ all exist on the same face of the furanose ring.[22]

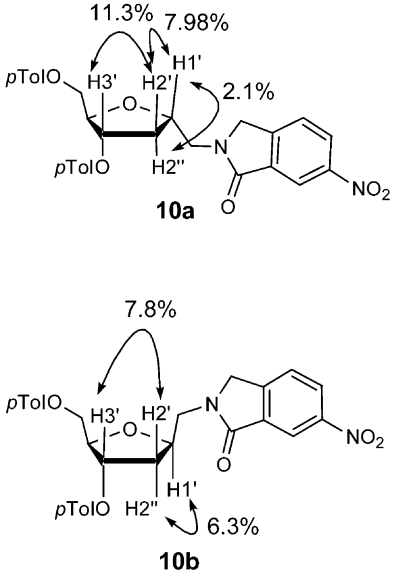
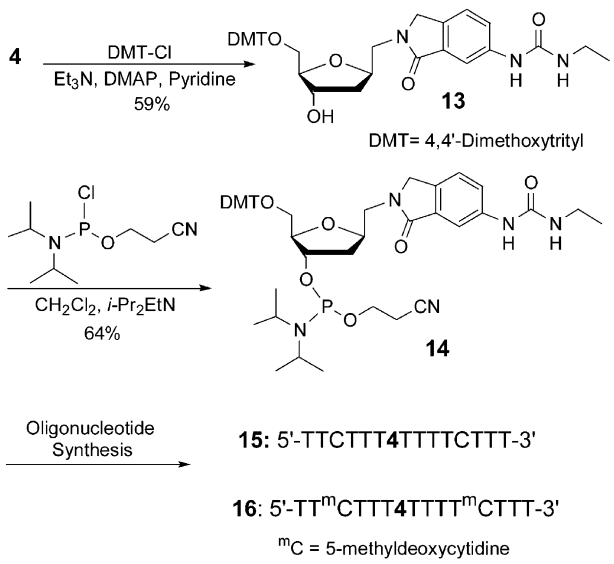


Figure 4. NOE values for the α- and β-anomers of an isoindolin-1-one homo-*N*-nucleoside.



Scheme 2.

The β-homo-*N*-nucleoside **10b** was subjected to Raney nickel catalyzed hydrogenation to produce amine **11** which was subsequently transformed into the ethyl urea **12** using ethylisocyanate and triethylamine. Removal of the toluoyl protecting groups with 1% w/v NaOH-methanol furnished base **4** in good yield.

To incorporate **4** into oligonucleotides, the non-natural base underwent further manipulations to its 5′-*O*-dimethoxytrityl (DMT) protected 3′-phosphoramidite (Scheme 2). Thus, reaction of **4** with DMT-Cl installed the 5′-blocking group in nucleoside **13**, and subsequent treatment of **13** with 2-cyanoethyl diisopropylchlorophosphoramidite produced phosphoramidite **14**, the activated starting material for automated oligonucleotide synthesis.[23]

Phosphoramidite **14** was employed in the automated synthesis of 15-mer oligonucleotides **15** and **16**, each of which contained the non-natural base in one position.

The crude product of automated DNA synthesis was analyzed using MALDI-TOF mass spectrometry (MALDI). The MALDI spectra of **15** and **16** using a 2′,4′,6′-trihydroxyacetophenone (THAP) matrix gave mass peaks at $m/z = 4577$ and 4609, respectively. These values were within experimental error of the calculated molecular weights of **15** (4578) and **16** (4606). Lower intensity peaks around 2400 were also observed. These represent the sequence 5′-TTTTCTTT-3′, which results from synthetic oligonucleotides which failed to incorporate phosphoramidite **14**. Crude **15** and **16** were purified using semi-preparative reverse-phase HPLC or denaturing polyacrylamide gel electrophoresis (PAGE).[23]

Coupling MALDI analysis with enzymatic digests of **15** provided complete characterization of the oligonucleotide sequence.[24] In one assay, **15** was treated with snake-venom phosphodiesterase without addition of alkaline phosphatase. Aliquots of the reaction mixture were obtained at several time intervals, and each aliquot was quenched by directly mixing with a 10-fold excess of THAP matrix. MALDI spectra were obtained of each aliquot (Fig. 5). Over time, peaks appeared in the MALDI spectrum corresponding to oligonucleotide fragments resulting from cleavage of bases from the 3′-end of **15**. MALDI analysis of the reaction mixture after 290 min indicated cleavage up to the point of insertion of the non-natural base **4**. No further cleavage was observed,

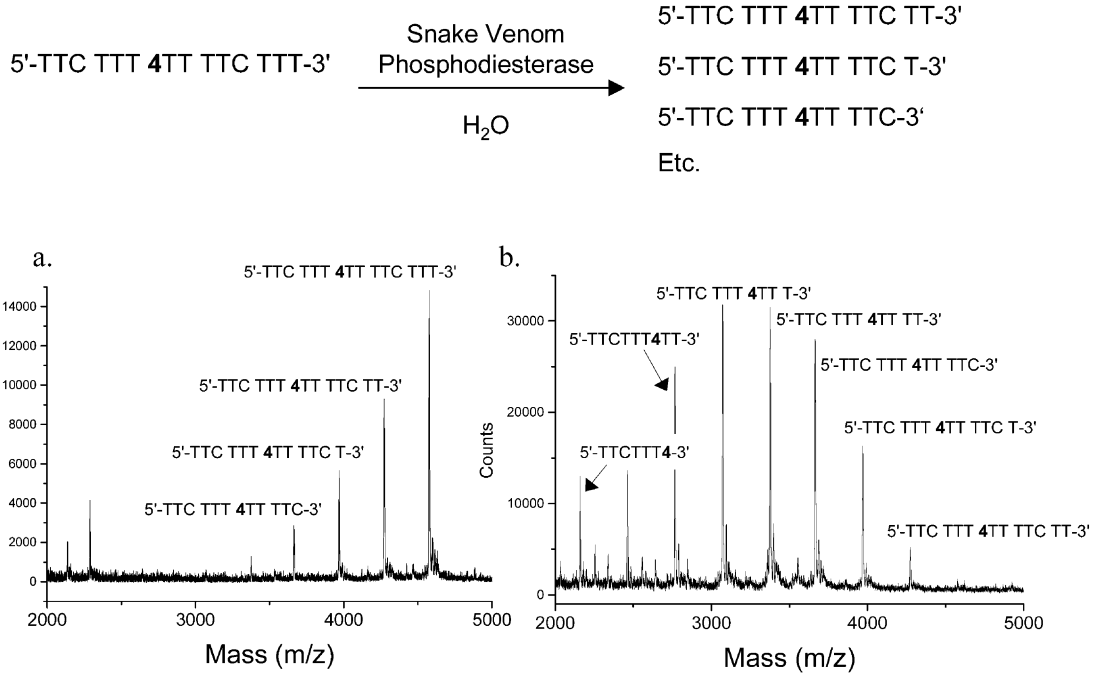


Figure 5. MALDI analysis of digestion of oligonucleotide **15** by snake-venom phosphodiesterase. (a) MALDI spectrum of a digest mixture aliquot after 24 min. (b) MALDI spectrum of a digest mixture aliquot after 290 min.

suggesting that snake-venom phosphodiesterase has trouble cleaving past the artificial base.

A similar MALDI-coupled digest experiment was performed using bovine spleen phosphodiesterase, a 5′-exonuclease which sequentially cleaves bases from the 5′-end of the oligonucleotide.[25] In this case, MALDI analysis of aliquots from the enzymatic digest mixture indicated cleavage of bases from the 5′-end of **15** up to and beyond the point of insertion of the artificial nucleoside **4** (Fig. 6). Taken together, the results of both the 3′-and 5′-exonuclease studies strongly indicate that base **4** is in tact in oligonucleotide **15**. Each MALDI peak observed in the two enzymatic digests is consistent with this finding. Thus, the possibility that modification or decomposition of **4** occurred during oligonucleotide synthesis was ruled out.

Oligonucleotide **15** was employed in triple helix thermodenaturation (T_m) studies in which the TFO targeted a homopurine tract within the 25-mer double helix shown below. Unfortunately, no clear triplex to duplex transitions were observed in studies employing **15** under a variety of thermodenaturation conditions.

15: 5′-TT CTTT**4**TTTTCTTT-3′
5′-CGGGCAAGAAA**Y**AAAAGAAACGCGC-3′
3′-GCGCGTTCTTT**Z**TTTTCTTTGCGCG-5′

One possible method for obtaining sharper triplex to duplex transitions is enhancing the overall TFO binding affinity for targeted duplex sequences. Thus, TFO **16** was next examined. This sequence was designed with 5-methyldeoxycytidine (^{m}C) residues in place of the 2′-deoxycytidine residues. Because the ^{m}C-GC base triplet is more stable than C-GC triplets at pH conditions near neutrality, it was believed that triple helices containing **16** should be more stable under the thermodenaturation experiment conditions.[26]

Triplex thermodenaturation studies were performed using TFO **16** and complementary 25-mer duplex targets of the sequence shown below. Four different 25-mer targets were studied, each had a different base pair (**YZ**) opposite the non-natural base **4**.

16: 5′-TT^{m}CTTT**4**TTTT^{m}CTTT-3′
5′-CGGGCAAGAAA**Y**AAAAGAAACGCGC-3′
3′-GCGCGTTCTTT**Z**TTTTCTTTGCGCG-5′

A triplex dissociation temperature of 18.3 °C was observed when **4** was opposite CG (Fig. 7). This value demonstrates that the **4**-CG complex within a triple helix is not robust. By comparison, T_m values of 27.5 °C and 44.0 °C were recorded when T-AT and C-GC base triplets replaced the **4**-CG base triplet in the identical triple helix sequence. These canonical base triplets are thus approximately 1.8–5.0 kcal/mol more stable than the **4**-CG base triplet.[27] Moreover, a T_m value of 20.1 °C was recorded when an abasic site (ϕ) replaced base **4** in the TFO. This suggests that **4** provided no stabilization to the triple helix over the abasic site.

Within the triple helix motifs examined, base **4** exhibited at best only marginal selectivity in recognizing CG over the other three base pairs (Table 1). Triple helices identical to those described above but with GC, AT, and TA base pairs opposite **4** exhibited T_m values only 0.4 to 1.4 °C lower than the triplex containing the **4**-CG base. This small ΔT_m value likely falls within the experimental error of the thermodenaturation studies. The lack of selectivity for CG is likely a consequence of the low affinity of **4** for CG within a triple helix.

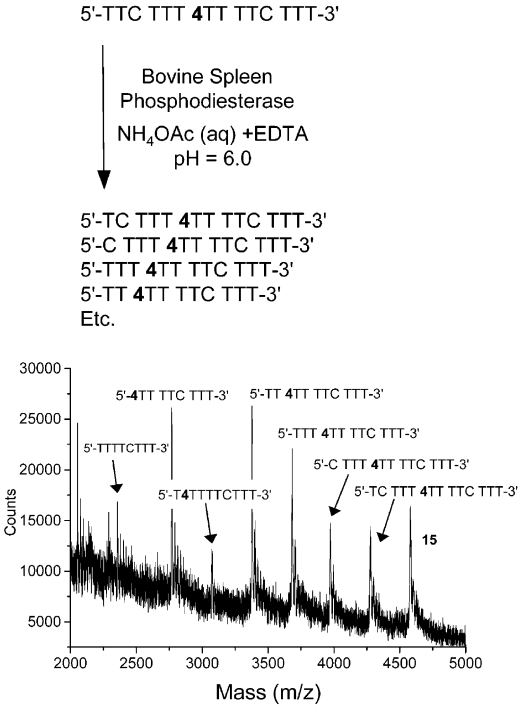


Figure 6. MALDI analysis of an aliquot taken from a bovine spleen phosphodiesterase digestion of oligonucleotide **15** obtained after 24 h.

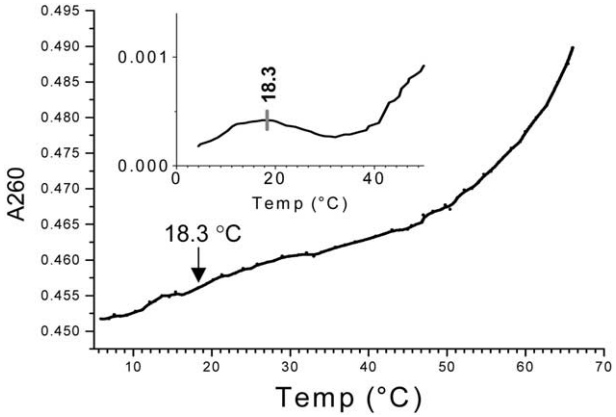


Figure 7. Melting temperature curve for a triple helix containing the **4**-CG base triplet. A T_m value of 18.3 °C was determined from the first derivative of this plot (inset).

Table 1. Summary of melting temperature (T_m) values for triple helices formed from a 15-mer TFO containing base **X** and a 25-mer target duplex containing base pairs **YZ**

5′-TT^{m}CTTT**X**TTTT^{m}CTTT-3′
5′-CGGGCAAGAAA**Y**AAAAGAAACGCGC-3′
3′-GCGCGTTCTTT**Z**TTTTCTTTGCGCG-5′

Base triplet (**X-YZ**)	Triple helix T_m (°C)
4-CG	18.3
4-TA	17.2
4-AT	17.9
4-GC	16.9
ϕ-CG	20.1
T-AT	27.5
C-GC	44.0

(ϕ = abasic site, see ref. 28 for the *C*-GC triplet T_m value.)

3. Conclusions

The isoindolin-1-one homo-*N*-nucleoside **4** was synthesized and successfully incorporated into oligonucleotide sequences. In the sequence motifs investigated, the **4**-CG base triplet was not robust, and little stabilization of triplex formation was observed. This result is in contrast to our previous finding that the organic-soluble isoindolin-1-one analogue **3** strongly and effectively complexed the CG base pair in chloroform.[10] Strong hydrogen bonding in organic solvent thus does not necessarily translate into effective binding within a DNA triple helix. Other subtle factors such as base triplet isomorphism, π-stacking, and conformational preorganization must also be considered in the design of nucleosides targeting TA and CG in a triplex-based recognition scheme. It is possible that the single methylene group in **4** between the 2′-deoxyribose unit and the heterocycle is not sufficient to make the **4**-CG base triplet isomorphous with T-AT and *C*-GC. Additionally, the ability of base **4** to stack with neighboring bases may be relatively low, thus attenuating any stability that the **4**-CG triplet may provide to a triple helix. A productive avenue may involve retaining the effective hydrogen bonding motif of base **4** while increasing the tether between the heterocyclic base and 2′-deoxyribose and improving the ability of the base to stack with neighboring residues. Modifications such as these may ultimately allow the strong hydrogen bonding potential of the ureido isoindolin-1-one design to result in effective recognition of CG in a DNA triple helix.[29]

4. Experimental

4.1. General methods

All reactions were performed under N_2 in glassware that was oven dried at 105 °C prior to use. Reagents and solvents used in reactions were obtained from commercial sources and were used without further purification except as follows: Tetrahydrofuran (THF) was distilled from sodium-benzophenone. Pyridine was dried over 4 Å molecular sieves. Acetonitrile, benzene, methylene chloride (CH_2Cl_2), and triethylamine were distilled from CaH_2. All reactions were monitored by TLC using silica gel 60 F_{254} plastic, aluminum, or glass plates (Merck). Flash chromatography was performed with 32–63 micron silica gel (Merck).[30] Radial chromatography was performed using a Harrison Research Chromatotron on 2 mm plates of silica gel 60 PF_{254} with gypsum. ^{1}H NMR, ^{13}C NMR, ^{31}P NMR data were obtained on either 400 or 500 MHz Varian U400, U500, and 500NB instruments. ^{1}H NMR spectra obtained in $CDCl_3$ were referenced to 7.26 ppm, and those obtained in CD_3OD were referenced to 3.31 ppm. ^{13}C NMR spectra obtained in $CDCl_3$ were referenced to 77.23 ppm, and those obtained in CD_3OD were referenced to 49.15 ppm. ^{31}P NMR spectra were referenced to an external 85% aqueous H_3PO_4 standard. Chemical shifts are reported in parts per million (ppm) and coupling constants are reported in Hertz (Hz). IR spectra were collected on a Mattson FTIR 5000 with major bands reported in cm^{-1}. UV/Vis spectra were obtained using a Shimadzu UV-2501PC recording spectrophotometer. FDMS and FABMS data were collected by the mass spectrometry service at the University of Illinois, and MALDI-TOF spectra were obtained using an Applied Biosystems Voyager-DE STR mass spectrometer. MALDI spectra of oligonucleotides were obtained in a 2,4,6-trihydroxyacetophenone (THAP) matrix. Elemental analysis was performed by the microanalytical laboratory at the University of Illinois. Melting points were measured with a Thomas Hoover melting point apparatus and are uncorrected.

4.2. Synthesis and characterization of new compounds

4.2.1. 4′-Nitrophenyl 6-(bromomethyl)-3-nitrobenzoate (6). A mixture of 1.00 g (5.60 mmol) *N*-bromosuccinimide, 6.0 mg (0.040 mmol) 2, 2′-azobisisobutyronitrile, 0.81 g (2.7 mmol) 4-nitrophenyl-6-methyl-3-nitrobenzoate, and 20 mL benzene was illuminated from below for 5 h with a 100 watt light bulb. During the illumination process, the reaction mixture heated to reflux. Upon cooling to 0 °C, a yellow solid precipitated. This solid was filtered off and discarded. The filtrate was evaporated at reduced pressure to yield an oily orange solid. This residue was purified using flash chromatography with a stepped solvent gradient ranging from 1:19 ethyl acetate–petroleum ether (PE) to 1:4 ethyl acetate-PE to yield **6** as 0.42 g (41%) of an oily yellow solid containing a small amount (5% by ^{1}H NMR integration) of unreacted starting material as an impurity. The product was further purified to a white solid by precipitation from 10:1 hexane-ethyl acetate: mp 109–111 °C. ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.02 (d, $J = 2.5$, 1H, H-2), 8.43 (dd, $J = 8.6$, 2.5, 1H, H-4), 8.35 (A, A′ of AA′XX′, $J_{AA'} = 2.5$, $J_{XX'} = 2.5$, $J_{AX} = 8.0$, $J_{AX'} = 0.5$, 2H, ArH′), 7.77 (d, $J = 8.5$, 1H, H-5), 7.49 (X,X′ of AA′XX′, $J_{AA'} = 2.5$, $J_{XX'} = 2.5$, $J_{AX} = 8.0$, $J_{AX'} = 0.5$, 2H, ArH), 5.01 (s, 2H, $ArCH_2Br$). ^{13}C NMR (125 MHz, $CDCl_3$) δ 162.0, 154.6, 147.3, 146.8, 145.7, 133.2, 128.5, 127.9, 126.8, 126.7, 125.3, 122.49, 28.8. FDMS *m*/*z* 382.1 $(M+1)^+$.

4.2.2. *N*-[(3′,5′,-*O*-Ditoluoyl-2′-deoxy-α-D-*ribo*-pentofuranosyl)methyl]-4-nitroisoindolin-1-one (10a). To a solution of 60 mg K_2CO_3, 30 mg KI, and 8.0 mL acetonitrile was added 78 mg (0.20 mmol) (3′,5′,-*O*-ditoluoyl-2′-

deoxy-α-D-*ribo*-pentofuranosyl)aminomethane. The resultant solution was stirred in an ice-water bath for 15 min. A solution of 76 mg (0.20 mmol) **6** in 2 mL acetonitrile was added, and stirring in an ice-water bath continued for 5 h. The reaction mixture was filtered through Celite, and the Celite layer was washed with ethyl acetate. Solvent was evaporated at reduced pressure, and the remaining brown residue was redissolved in ethyl acetate. This solution was washed twice with 20 mL portions of H_2O, and four times with 20 mL of saturated aqueous K_2CO_3. The organic layer was dried over $MgSO_4$, filtered, and evaporated at reduced pressure to yield an amber oil. Flash chromatography using a stepped solvent gradient ranging from 3:7 ethyl acetate–PE to 1:1 ethyl acetate–PE afforded **10a** as 33 mg (30%) of a white solid. Compound **10a** was isolated at 91% purity, as determined by ^{1}H NMR integration: mp 74–80 °C. ^{1}H NMR ($CDCl_3$, 500 MHz) δ 8.60 (d, J = 2.1, 1H, H-3), 8.36 (dd, J = 8.3, J = 2.1, 1H, H-5), 7.88 (d, J = 7.9, 2H, ArH′), 7.74 (d, J = 7.9, 2H, ArH′), 7.53 (d, J = 8.3, 1H, H-6), 7.22 (d, J = 7.9, 2H, ArH′), 7.11 (d, J = 7.9, 2H, ArH′), 5.45 (dt, J = 6.7, J = 2.8, 1H, H-3′), 4.72 (A of AB_q, J_{AB} = 12.1, 1H, H-8), 4.69 (B of AB_q, J_{AB} = 12.1, 1H, H-8′), 4.60 (X of ABX, J_{AX} = 2.7, J_{BX} = 7.0, 1H, H-4′), 4.57 (X of ABX, J_{AX} = 2.7, J_{BX} = 7.0, 1H, H-1′), 4.52 (A of ABX, J_{AB} = 11.6, J_{AX} = 2.7, 1H, H-5′), 4.43 (B of ABX, J_{AB} = 11.6, J_{BX} = 4.1, 1H, H-5″), 3.93 (A of ABX, J_{AB} = 14.4, J_{AX} = 2.7, 1H, H-6′), 3.83 (B of ABX, J_{AB} = 14.4, J_{BX} = 7.0, 1H, H-6″), 2.73 (dt, J = 14.3, 7.6, 1H, H-2′), 2.40 (s, 3H, $ArCH_3$), 2.36 (s, 3H, $ArCH_3$), 2.15 (ddd J = 14.2, J = 5.6, J = 3.0, 1H, H-2″). ^{13}C NMR ($CDCl_3$, 125 MHz) δ 166.9, 166.5, 166.4, 148.6, 147.9, 144.4, 144.3, 134.2, 129.8, 129.7, 129.4, 129.3, 127.0, 126.8, 126.4, 123.9, 119.5, 82.5, 79.0, 76.8, 64.4, 52.1, 46.8, 35.5, 31.1, 21.9, 21.8. FDMS m/z 544.2 (M^+).

4.2.3. *N*-[(3′,5′,-*O*-Ditoluoyl-2′-deoxy-β-D-*ribo*-pentofuranosyl)methyl]-4-nitroisoindolin-1-one (10b). To a solution of 434 mg (3.14 mmol) K_2CO_3, 199 mg (1.20 mmol) KI, and 30 mL acetonitrile was added 547 mg (1.43 mmol) **7**. The resultant solution was stirred in an ice-water bath for 15 min. A solution of 800 mg (2.1 mmol) **6** in 30 mL acetonitrile was added slowly dropwise over 15 min. The reaction mixture was allowed to warm to room temperature, and stirring continued for 4 h. The reaction mixture was filtered through Celite, and the Celite layer was washed with ethyl acetate. Solvent was evaporated at reduced pressure and the remaining brown residue was redissolved in 50 mL ethyl acetate. This solution was washed three times with 30 mL portions of H_2O and five times with 30 mL portions of saturated aqueous K_2CO_3. The organic layer was dried over $MgSO_4$, filtered, and evaporated at reduced pressure to yield an amber oil. Flash chromatography using a stepped solvent gradient ranging from 3:7 ethyl acetate–PE to 3:2 ethyl acetate–PE afforded **10b** as 535 mg (69%) of a white solid at >98% purity as determined by ^{1}H NMR integration. A portion of the product was further purified by recrystallization from 1:1 ethyl acetate–hexane: mp 158–160 °C. ^{1}H NMR ($CDCl_3$, 400 MHz) δ 8.46 (d, J = 1.8, 1H, H-3), 8.20 (dd, J = 8.4, J = 2.2, 1H, H-5), 7.92 (d, J = 8.4, 2H, ArH′), 7.62 (d, J = 8.3, 2H, ArH′), 7.32 (d, J = 8.0, 1H, H-6), 7.25 (d, J = 6.0, 2H, ArH), 7.03 (d, J = 7.9, 2H, ArH), 5.55 (dt, J = 5.9, 1H, H-3′), 4.75 (dd, J = 11.9, 2.7, 1H, H-5′), 4.67 (B of AB_q, J_{AB} = 18.9, 1H, H-8′), 4.54 (A of AB_q, J_{AB} = 18.9, 1H, H-8), 4.54 (m, 1H, H-1′), 4.42–4.34 (m, 2H, H-4′, H-5″), 3.87 (d, J = 3.86, 2H, H-6′, H-6″), 2.42 (s, 3H, $ArCH_3$), 2.35 (dd, J = 13.7, 4.6. 1H, H-2″), 2.32 (s, 3H, $ArCH_3$), 2.10 (ddd J = 13.6, J = 11.3, J = 6.3, 1H, H-2′). ^{13}C NMR ($CDCl_3$, 125 MHz) δ 167.0, 166.5, 166.3, 148.3, 147.9, 144.5, 144.4, 133.8, 129.9, 129.5, 129.3, 127.0, 126.5, 126.3, 123.8, 119.2, 83.6, 79.3, 77.0, 64.8, 52.9, 45.0, 35.9, 21.9, 21.7. IR (film) ν 1706.7, 1685.5, 1608.4, 1527. FDMS m/z 544.2 (M^+).

4.2.4. *N*-[(3′,-5′,-*O*-Ditoluoyl-2′-deoxy-β-D-*ribo*-pentofuranosyl)methyl]-4-amino-isoindolin-1-one (11). To a solution of 650 mg (1.19 mmol) **10b** in 70 mL ethyl acetate was added 3 mL of a thick suspension of Raney nickel in ethyl acetate. The resultant mixture was stirred under 1 atmosphere of hydrogen gas for 23 h. The reaction mixture was filtered through Celite, and the Celite layer was washed with 30 mL ethyl acetate. Combined organic portions were evaporated at reduced pressure to yield 400 mg (65%) of an oily white solid which was recrystallized from 1:1 hexane–ethyl acetate to yield **11** as white needles: mp 96–99 °C. ^{1}H NMR ($CDCl_3$, 500 MHz) δ 7.90 (d, J = 8.1, 2H, ArH′), 7.77 (d, J = 7.9, 2H, ArH′), 7.23 (d, J = 7.9, 2H, ArH), 7.09 (d, J = 8.0, 2H, ArH), 7.06 (d, J = 1.5, 1H, H-3), 6.96 (d, J = 8.2, 1H, H-6), 6.72 (dd, J = 8.2, 2.1, 1H, H-5), 5.48 (d, J = 5.6, 1H, H-3′), 4.59 (B of ABX, J_{AB} = 11.7, J_{BX} = 3.5, 1H, H-5″), 4.50 (X of ABX, 1H, J_{AX} = 3.3, J_{BX} = 3.5, H-4′), 4.48 (A of ABX, J_{AB} = 11.7, J_{AX} = 3.3, 1H, H-5′), 4.39 (s, 2H, H-8, H-8′), 4.33 (X of ABX, J_{AX} = 2.2, J_{BX} = 5.9, H-1′), 3.93 (A of ABX, J_{AB} = 14.4, J_{AX} = 2.2, 1H, H-6″), 3.81 (br s, 2H, NH_2), 3.65 (B of ABX, J_{AB} = 14.4, J_{BX} = 5.9, 1H, H-6″), 2.40 (s, 3H, $ArCH_3$), 2.37 (s, 3H, $ArCH_3$), 2.29 (dd, J = 13.8, 4.8, 1H, H-2″), 2.05 (ddd, J = 13.6, 7.5, 4.7, 1H, H-2′). ^{13}C NMR (125 MHz, $CDCl_3$) δ 169.4, 166.5, 166.4, 146.7, 144.3, 143.9, 133.5, 131.9, 129.9, 129.6, 129.4, 129.3, 127.1, 126.9, 123.4, 118.9, 108.9, 83.3, 79.3, 76.9, 64.7, 52.2, 45.6, 35.9, 21.8. FDMS m/z 514.3 (M^+). Anal. calcd for $C_{30}H_{30}N_2O_6$: C, 70.02; H, 5.88; N, 5.44. Found: C, 70.14; H, 5.94; N, 5.08.

4.2.5. *N*-[(3′,-5′,-*O*-Ditoluoyl-2′-deoxy-β-D-*ribo*-pentofuranosyl)methyl]-4-[(ethylamino)-carbonyl]amino-isoindolin-1-one (12). To a solution of 200 mg (0.39 mmol) **11** in 12 mL THF was added 0.20 mL triethylamine and 0.10 mL (1.26 mmol) ethyl isocyanate. The resultant solution was stirred at 60 °C for 48 h. Solvent was evaporated at reduced pressure to yield a white foam. Purification using flash chromatography with 1:1 ethyl acetate–CH_2Cl_2 produced 135 mg (59%) of **12** as a white solid: mp 196–198 °C. ^{1}H NMR ($CDCl_3$, 400 MHz) δ 8.06 (dd, J = 8.3, 2.1, 1H, H-5), 7.91 (dt, J = 8.2, 1.7, 2H, ArH″), 7.75 (dt, J = 8.2, 1.5, 2H, ArH″), 7.58 (br s, 1H, NH-2), 7.45 (d, J = 1.8, 1H, H-3), 7.24 (dd, J = 8.6, 0.5, 2H, ArH′), 7.16 (d, J = 8.4, 1H, H-6), 7.08 (dd, J = 8.6, 0.6, 2H, ArH′), 5.52 (br t, J = 6.2, 1H, NH-1), 5.51 (d, J = 4.5, 1H, H-3′), 4.66 (dd, J = 11.9, 3.34, 1H, H-5″), 4.53 (m, 1H, H-1′), 4.50 (s, 2H, H-8, H-8′), 4.48 (dd, J = 11.9, 4.0, 1H, H-5″), 4.36 (m, 1H, H-4′), 3.96 (dd,

J = 14.5, 2.5, 1H, H-6″), 3.67 (dd, J = 14.5, 5.9, 1H, H-6′), 3.34 (qd, J = 7.3, 6.9, 2H, CH_2), 2.42 (s, 3H, $ArCH_3$), 2.36 (s, 3H, $ArCH_3$), 2.29 (dd, J = 14.0, 5.1, 1H, H-2″), 2.09 (ddd, J = 13.9, 11.3, 6.4, 1H, H-2′), 1.19 (t, J = 7.1, 3H, CH_3). ^{13}C NMR ($CDCl_3$, 125 MHz) δ 169.6, 166.3, 166.2, 156.1, 144.4, 144.6, 134.9, 132.1, 129.8, 129.5, 129.3, 129.2, 126.8, 126.6, 123.4, 122.7, 112.4, 83.3, 78.8, 76.7, 64.5, 52.7, 45.7, 36.1, 34.9, 21.8, 21.7, 15.7. FABMS m/z 586.3 (MH^+).

4.2.6. *N*-[(2′-Deoxy-β-D-*ribo*-pentofuranosyl)methyl]-4-[(ethylamino)carbonyl]-amino-isoindolin-1-one (4). A solution of 177 mg (0.30 mmol) **12** in 5 mL 1% w/v NaOH-methanol was stirred for 1 h. Solvent was evaporated at reduced pressure with minimal heating to yield a yellow solid. Purification by flash chromatography with 1:19 methanol-acetone afforded **4** as 87 mg (83%) of an oily white solid which contained 5–10% residual silica gel as an impurity: mp 86–93 °C. ^{1}H NMR (CD_3OD, 400 MHz) δ 7.76 (d, J = 2.1, 1H, H-3), 7.60 (dd, J = 8.2, 2,1, 1H, H-5), 7.41 (d, J = 8.2, 1H, H-6), 4.58 (A of ABq, J_{AB} = 17.8, 1H, H-8′), 4.55 (B of ABq, J_{AB} = 17.8, 1H, H-8), 4.40 (X of ABX, 1H, J_{AX} = 3.2, J_{BX} = 7.3, 1H, H-1′), 4.20 (X of ABX, J_{AX} = 6.1, J_{BX} = 7.0, 1H, H-4′), 3.78 (dd, J = 8.2, 5.1, 1H, H-3′), 3.75 (A of ABX, J_{AB} = 11.5, J_{AX} = 3.2, 1H, H-6‴), 3.67 (B of ABX, J_{AB} = 11.5, J_{BX} = 7.3, 1H, H-6′), 3.55 (A of ABX, J_{AB} = 11.7, J_{AX} = 6.1, 1H, H-5″), 3.52 (B of ABX, J_{AB} = 11.7, J_{BX} = 7.0, 1H, H5′), 3.23 (q, J = 7.2, 2H, CH_2), 1.95 (ddd, J = 13.0, 5.8, 2.6, 1H, H-2″), 1.84 (ddd, J = 13.2, 9.3, 6.2, 1H, H-2′), 1.15 (J = 7.2, 3H, CH_3). ^{13}C NMR (CD_3OD, 125 MHz) δ 171.2, 158.2, 141.6, 137.3, 133.9, 124.4, 124.1, 114.2, 89.1, 78.8, 73.7, 63.8, 52.9, 48.0, 39.5, 35.8, 15.8. HRMS (FAB) calcd for $C_{17}H_{23}N_3O_5$ $(M+H)^+$ 350.1716, found 350.1714.

4.2.7. *N*-[5′-*O*-[Bis(4-methoxyphenyl)phenylmethyl]-(2′-deoxy-β-D-*ribo*-pentofuranosyl)methyl]-4-[(ethylamino)-carbonyl]amino-isoindolin-1-one (13). To a suspension of 70 mg (0.20 mmol) **4**, 4.0 mg 4-(dimethylamino)pyridine, 0.050 mL triethylamine, and 1 mL pyridine was added 81 mg of 4, 4′-dimethoxytrityl chloride. The resultant clear solution was stirred at room temperature for 1.5 h. Additional portions of 81 mg 4, 4′-dimethoxytrityl chloride and 0.050 mL triethylamine were added, and the reaction mixture was stirred at 45 °C for 13 h. The reaction mixture was combined with 20 mL H_2O and extracted with 30 mL ethyl acetate. The organic layer was evaporated at reduced pressure to afford a yellow oil. Purification using radial chromatography with a stepped solvent gradient of 1:9 ethyl acetate–CH_2Cl_2 to neat ethyl acetate on silica deactivated with triethylamine afforded 95 mg (72%) of **13** as an oily off-white solid. The product was isolated at >98% purity as determined by ^{1}H NMR integration: mp 105–110 °C. ^{1}H NMR ($CDCl_3$, 400 MHz) δ 8.49 (s, 1H, NH-2), 8.30 (dd, J = 8.31, 1.8, 1H, H-5), 7.44 (dd, J = 8.1, 1.5, 2H, ArH), 7.31 (dt, J = 8.9, 2.2, 4H, ArH″), 7.29 (d, J = 1.9, 1H, H-3), 7.25 (t, J = 7.76, 2H, ArH), 7.19–7.15 (m, 1H, ArH), 7.12 (d, J = 8.4, 1H, H6), 6.79 (d, J = 9.0, 4H, ArH′), 6.19 (br t, J = 5.0, 1H, NH-1), 4.56–4.26 (m, 2H, H-1′, H-3′), 4.49 (A of ABq, J_{AB} = 17.3, 1H, H-8′), 4.33 (B of ABq, J_{AB} = 17.3, 1H, H-8″), 4.10–3.94 (m, 2H, H-5′, H-5″), 3.75 (s, 7H, OCH_3), 3.42–3.14 (m, 5H, CH_2, H-6′, H-6″ H-4′), 2.00–1.81 (m, 2H, H-2′, H-2″), 1.15 (t, J = 7.2, 3H, CH_3). ^{13}C NMR ($CDCl_3$, 100 MHz) δ 168.8, 158.3, 155.9, 144.7, 140.3, 135.9, 135.8, 134.1, 131.9, 129.9, 130.0, 129.9, 128.0, 127.7, 122.8, 121.5, 112.9, 111.5, 86.5, 85.9, 77.4, 73.8, 64.4, 55.1, 52.2, 46.8, 38.8, 34.6, 15.4. HRMS (FAB) calcd for $C_{38}H_{41}N_3O_7$ (M^+) 651.2944, found 651.2943.

4.2.8. *N*-[5′-*O*-[Bis(4″-methoxyphenyl)phenylmethyl]-(2′-deoxy-β-D-*ribo*-pentofuranosyl)-methyl]-4-[(ethylamino)-carbonyl]amino-isoindolin-1-one-3′-*O*-(2-cyanoethyl-*N*,*N*′-diisopropylaminophosporamidite (14). A solution of 109 mg (0.167 mmol) **13**, 0.056 mL (0.35 mmol) 2-cyanoethyl diisopropylchlorophosphoramidite and 0.87 mL diisopropylethylamine in 1.0 mL CH_2Cl_2 was stirred at room temperature for 1.5 h. Methanol was added to the reaction mixture dropwise and solvent was subsequently evaporated at reduced pressure to yield a white foam. Purification using radial chromatography with 1:1 PE–CH_2Cl_2 on silica deactivated with triethylamine afforded 91 mg (64%) of a white foam consisting of a mixture of two diastereomers of **14**: mp 109–116 °C. ^{1}H NMR ($CDCl_3$, 500 MHz) δ 8.28 (dd, J = 8.3, 2.1, 1H, H-5), 8.13 (br s, 1H, NH-2), 7.56 (d, J = 1.9, 1H, H-3), 7.41 (m, 2H, ArH), 7.35–7.19 (m, 8H, H6, ArH, ArH, ArH′), 6.78 (d, J = 9, 2H, ArH′), 6.76 (d, J = 9, 2H, ArH′), 5.94 (br t, J = 7.5, 1H, NH-1), 4.60 (A of ABq, J_{AB} = 12.4, 1H, H-8′), 4.59 (A of ABq J_{AB} = 17.4, 1H, H-8′), 4.55 (B of ABq, J_{AB} = 12.4., 1H, H-8), 4.53 (B of ABq, J_{AB} = 17.4, 1H, H-8), 4.42 (m, 2H, H3′, H1′), 4.08 (m, 1H, H-5′), 3.95 (dd, J = 14.8, 3.1, 1H, H-5″), 4.0–3.60 (m, 3H, H-4′, OCH_2). 3.78 (s, 6H, OCH_3), 3.60–3.48 (m, 2H, *i-pr*CH), 3.21–3.09 (m, 2H, H-6′, H-6″), 2.57 (t, J = 6.7 Hz, 2H, CH_2CN), 2.39 (t, J = 6.7, 2H, CH_2CN), 2.13 (dd, J = 12.3, 5.3, 1H, H-2″), 2.04 (dd, J = 12.3, 5.3, 1H, H2″), 1.94–1.82 (m, 1H, H-2′), 1.18 (t, J = 7.4, 3H, CH_3), 1.15 (d, J = 6.8, 6H, *i-Pr*CH_3), 1.14 (d, J = 6.7, 6H, *i-Pr*CH_3), 1.12 (d, J = 6.8, 6H, *i-Pr*CH_3), 1.04 (d, J = 6.8, 6H, *i-Pr*CH_3). ^{13}C NMR ($CDCl_3$, 125 MHz) δ 169.6, 158.7, 156.0, 144.9, 140.7, 136.2, 136.1, 135.1, 132.6, 130.3, 128.4, 128.0, 127.0, 123.7, 123.1, 117.7, 116.7, 113.3, 112.8, 86.3, 86.1, 78.2, 75.3, 75.1, 63.9, 58.4, 58.3, 55.4, 52.7, 46.6, 43.4, 43.4, 37.8, 35.0, 24.8, 24.7, 24.6, 20.3, 15.7. ^{31}P NMR ($CDCl_3$, 500 MHz) δ 149.1, 148.7. FDMS m/z 852.6 $(M+H)^+$.

4.3. Synthesis, analysis, and purification of synthetic oligonucleotides

Solutions of **14** in acetonitrile (1.0 M) were submitted to the W.M. Keck Center for Comparative and Functional Genomics (Urbana, IL) for oligonucleotide synthesis on a 1 μmol scale. Crude products were supplied as lyophilized powders. For purification of the crude oligonucleotides, semi-preparative reverse phase HPLC was performed on a 150 mm×4.6 mm Cosmosil column with a solvent gradient of 5% acetonitrile–0.10 M triethylammonium acetate to 25% acetonitrile–0.10 M triethylammonium acetate at 1.0 mL/min. Fractions containing the desired oligonucleotide were lyophilized, and the resulting material was used in T_m studies without further purification. For MALDI-coupled enzymatic

digests, oligonucleotide samples were prepared using semi-preparative polyacrylamide gel electrophoresis (PAGE). PAGE was performed on a 0.10 mm 20% polyacrylamide slab gel prepared with urea. Oligonucleotide samples were loaded as 10 μL of 0.023 mg/μL solutions in 1:1 H_2O–formamide. Electrophoresis proceeded at 311 V with 1× tris-Borate as the buffer. The dominant UV absorbing bands were excised and eluted into 0.50 M ammonium acetate solutions containing 10 mM magnesium acetate. After filtration, oligonucleotide solutions were desalted using SEP-PAK cartridges (Waters). Desalted oligonucleotides were eluted from the cartridges using 6:4 methanol–water, and the resultant solutions were lyophilized.

4.4. MALDI-coupled enzymatic digest assays

To 60 μL of a 40 μM solution of TFO **15** in an Eppendorf tube was added 3 μL of a 0.067 mg/μL solution of snake-venom phosphodiesterase (Sigma). The resulting solution was thoroughly mixed. Aliquots (1.0 μL) of the reaction mixture were obtained at regular time intervals up to 24 h. These aliquots were quenched upon combination with 5.0 μL of a THAP/ammonium citrate matrix. Samples (1.0 μL) of the resulting mixture were loaded directly onto a MALDI plate for analysis.

MALDI-coupled spleen phosphodiesterase analysis proceed in the same manner with the following modifications: 10 μL of an 0.15 mg/90 μL solution of calf-spleen phosphodiesterase (Worthington Enzymes; solution was buffered to pH 6.0 with a 0.15 M ammonium acetate + 0.01 M EDTA buffer) was added to a solution of 6 μg TFO **15** in 20 μL water + 2 μL pH 6.0 ammonium acetate buffer. The resulting mixture was incubated at 37 °C for 24 h. During this time period, two aliquots from the mixture were frozen, and the frozen samples were analyzed using MALDI mass spectrometry.

4.5. UV–Vis triple helix T_m studies

In a typical triple helix thermodenaturation study, a 1.0 mL target duplex solution was prepared by combining a homopurine strand (0.46–0.5 μM) and a homopyrimidine strand (0.48–0.92 μM) in a pH 6.5 tris–HCl melt buffer (0.050 M tris, 0.040 M $MgCl_2$, 0.10 M NaCl, adjusted to pH 6.5 using concd HCl) at 70 °C.[26b] This duplex solution was annealed slowly (cooling rate = 1 °C/min) to 5 °C, and the duplex T_m was determined by increasing the temperature of the duplex solution (heating rate = 1 °C/min) and observing the absorbance at 260 nm. The temperature of the duplex mixture was increased until 70 °C. An aliquot of a freshly prepared aqueous solution of a non-natural base containing TFO was added to the duplex mixture such that the TFO concentration was 1.35 μM. The resulting triplex mixture was annealed slowly to 5 °C. The temperature of the triplex solution was increased at a rate of 1.0 °C/min. During this process, the absorbance at 260 nm was observed and recorded at 1 min intervals.

To determine triple helix T_m values, plots of absorbance at 260 nm vs. temperature were produced in *Origin* by Microcal. The first derivative of each plot was obtained, and the local maximum of the first derivative data was recorded after performance of one smoothing operation. For all T_m studies, solution temperatures were measured directly with a Cole-Parmer digital thermistor temperature gauge. The UV/Vis sample chamber was continuously flushed with dry nitrogen gas to insure a low humidity environment, and care was taken to reduce formation of bubbles in duplex and triplex solutions.

Acknowledgements

We wish to thank the W. M. Keck Center for Comparative and Functional Genomics for assistance with oligonucleotide synthesis. Funding for this work was generously provided by the University of Illinois and the National Institutes of Health.

References and notes

1. (a) Reviews and lead references on triplex DNA: Felsenfeld, G.; Davies, D. R.; Rich, A. *J. Am. Chem. Soc.* **1957**, *79*, 2023. (b) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645. (c) Le Doan, T.; Perrouault, L.; Praseuth, D.; Habhoub, N.; Decout, J.-L.; Thuong, N. T.; Lhomme, J.; Hélène, C. *Nucleic Acids Res.* **1987**, *15*, 7749. (d) Thuong, N. T.; Hélène, C. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 666. (e) Doronina, S. O.; Behr, J.-P. *Chem. Soc. Rev.* **1997**, 63. (f) Praseuth, D.; Guieysse, A. L.; Hélène, C. *Biochim. Biophys. Acta* **1999**, *1489*, 181.
2. Maher, L. J., III; Wold, B.; Dervan, P. B. *Science* **1989**, *245*, 725.
3. Takasugi, M.; Guendouz, A.; Chassignol, M.; Decout, J.-L.; Lhomme, J.; Thuong, N. T.; Hélène, C. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 5602.
4. Strobel, S. A.; Dervan, P. B. *Science* **1990**, *249*, 73.
5. (a) Reviews of other triplex DNA applications: Maher, L. J., III *Cancer Investigation* **1996**, *14*, 66. (b) Chan, P. P.; Glazer, P. M. *J. Molec. Med.* **1997**, *75*, 267. (c) Malvy, C., Harel-Bellan, A., Pritchard, L. L., Eds. *Triple Helix Forming Oligonucleotides*; Kluwer Academic Publishers: Boston, 1999.
6. (a) Mergny, J. L.; Sun, J.-S.; Rougee, M.; Montenay-Garestier, T.; Barcelo, F.; Chomilier, J.; Helene, C. *Biochemistry* **1991**, *30*, 9791. (b) Griffin, L. C.; Kiessling, L. L.; Beal, P. A.; Gillespie, P.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 7976.
7. Robles, J.; Grandas, A.; Pedroso, E.; Luque, F. J.; Eritja, R.; Orozco, M. For reviews of non-natural nucleoside bases for base pair recognition in triple helices see *Current Org. Chem.* **2002**, *6*, 1333. (b) Purwanto, M. G. M.; Weisz, K. *Current Org. Chem.* **2003**, *7*, 427.
8. It should be noted that natural guanosine residues were found to recognize the TA base pair in the pyr-pur-pyr triple helix motif: Griffin, L. C.; Dervan, P. B. *Science* **1989**, *245*, 967.
9. Zimmerman, S. C.; Schmitt, P. *J. Am. Chem. Soc.* **1995**, *117*, 10769.
10. Mertz, E.; Mattei, S.; Zimmerman, S. C. *Org. Lett.* **2000**, *2*, 2931.
11. (a) Sasaki, S.; Nakashima, S.; Nagatsugi, F.; Tanaka, Y.; Hisatome, M.; Maeda, M. *Tetrahedron Lett.* **1995**, *36*, 9521. (b) Cho, Y. L.; Jeong, K.-S. *Tetrahedron Lett.* **1997**, *38*, 8337. (c) Lecubin, F.; Benhida, R.; Fourrey, J.-L.; Sun, J.-S. *Tetrahedron Lett.* **1999**, *40*, 8085. (d) Saito, A.;

Kuroda, R. *Tetrahedron Lett*. **1999**, *40*, 4837. (e) López de la Paz, M.; González, C.; Vicent, C. *Chem. Commun*. **2000**, 411. (f) Purwanto, M. G. M.; Lengeler, D.; Weisz, K. *Tetrahedron Lett*. **2002**, *43*, 61.
12. Complexation of a 4-amino phthalimide base with CG was also studied in: Lengeler, D.; Weisz, K. *Tetrahedron Lett*. **2001**, *42*, 1479.
13. Zimmerman, S. C.; Mattei, S. Unpublished Results.
14. (a) Boal, J. H; Wilk, A.; Scremin, C. L.; Gray, G. N.; Phillips, L. R.; Beaucage, S. L. *J. Org. Chem*. **1996**, *61*, 8617. (b) Hossain, N.; Blaton, N.; Peeters, O.; Rozenski, J.; Herdewijn, P. *Tetrahedron* **1996**, *52*, 5563.
15. (a) Obika, S.; Hari, Y.; Sekiguchi, M.; Imanishi, T. *Angew. Chem., Int. Ed*. **2001**, 40. (b) Guianvarc'h, D.; Benhida, R.; Fourrey, J.-L.; Maurisse, R.; Sun, J.-S. *Chem. Commun*. **2001**, 1814.
16. Molecular models of triplex DNA were generated using (a) *Macromodel*, version 5.0, produced by Columbia University or (b) *Insight II*, produced by Accelrys. (c) For each structure, a natural base was excised from a triplex structure and a minimized ureido isoindolin-1-one homo-*N*-nucleoside structure was placed in the resulting gap. In the case of models generated by *Macromodel*, coordinates were obtained through the Protein Data Bank for an NMR structure reported in Radhakrishnan, I.; Patel, D. J. *Structure* **1994**, *2*, 17.
17. (a) A related lactamization approach was reported in: Almansa, C; Gómez, L. A.; Cavalcanti, F. L.; Rodriguez, R.; Carceller, E.; Bartroli, J.; Garcia-Rafanell, J.; Forn, J. *J. Med. Chem*. **1993**, *36*, 2121. (b) Egbertson, M. S.; Naylor, A. M.; Hartman, G. D.; Cook, J. J.; Gould, R. J.; Holahan, M. A.; Lynch, J. J., Jr.; Lynch, R. J.; Stranieri, M. T.; Vassallo, L. M. *Bioorg. Med. Chem. Lett*. **1994**, *4*, 1835.
18. Phadnis, A. P.; Nanda, B.; Patwardhan, S. A.; Gupta, A. S. *Indian J. Chem., Sect. B*. **1984**, *23*, 1098.
19. Hoffer, M. *Chem. Ber*. **1960**, *93*, 2777.
20. Iyer, R. P.; Phillips, L. R.; Egan, W. *Synth. Commun*. **1991**, *21*, 2053.
21. Scremin, C. L.; Boal, J. H.; Wilk, A.; Phillips, L. R.; Beaucage, S. L. *Bioorg. Med. Chem. Lett*. **1996**, *6*, 207.
22. Ren, R. X.-F.; Chaudhuri, N. C.; Paris, P. L.; Rumney, S., IV; Kool, E. T. *J. Am. Chem. Soc*. **1996**, *118*, 7671.
23. Gait, M. J., Ed. *Oligonucleotide Synthesis*; IRL Press: Oxford, UK, 1984.
24. Pieles, U.; Zürcher, W.; Schär, M.; Moser, H. E. *Nucleic Acids Res*. **1993**, *21*, 3191.
25. Enzymatic reaction conditions modified from Bernardi, A.; Bernardi, G. *The Enzymes IV*, 3rd Ed.; Academic Press: New York, 1971; p. 329.
26. (a) Lee, J. S.; Woodsworth, M. L.; Latimer, L. J. P.; Morgan, A. R. *Nucleic Acids Res*. **1984**, *12*, 6603. (b) Conditions for thermodenaturation studies were adapted from Miller, P. S.; Bhan, P.; Cushman, C. D.; Trapane, T. L. *Biochemistry* **1992**, *31*, 6788.
27. $\Delta\Delta G$ estimates are based on triplex T_m data from Booher, M. A.; Wang, S.; Kool, E. T. *Biochemistry* **1994**, *33*, 4645.
28. Zimmerman, S. C.; Schmitt, P. unpublished results.
29. Alternative strategies for sequence specific DNA recognition have also been developed. Most notable is an approach involving minor groove recognition by polyamides: Dervan, P. B.; Bürli, R. W. *Current Opinion in Chemical Biology* **1999**, *3*, 688.
30. Still, W. C.; Kahn, M.; Mitra, A. *J. Org. Chem*. **1978**, *43*, 2923.