Selectivity and affinity of triplex-forming oligonucleotides containing 2'-aminoethoxy-5-(3-aminoprop-1-ynyl)uridine for recognizing AT base pairs in duplex DNA

Sadie D. Osborne, Vicki E. C. Powers, David A. Rusling[1], Oliver Lack, Keith R. Fox[1] and Tom Brown*

School of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ, U.K.
[1]School of Biological Sciences, University of Southampton, Bassett Crescent East, Southampton SO16 7PX, U.K.

SUPPLEMENTARY MATERIAL

Chemical synthesis

General reagents for chemical synthesis

Reagents for chemical synthesis were purchased from Aldrich, Avocado, Cruachem, Fluka, Lancaster or Link Technologies and used without purification with the exception of the following solvents, which were purified by distillation: methanol (over iodine and magnesium), THF (over sodium wire and benzophenone), DCM, DIPEA, pyridine and TEA (over calcium hydride). All chemical reactions were carried out under argon using oven dried glassware. Column chromatography was carried out under pressure using Fisher scientific DAVISIL 60A (35-70 micron) silica. Compounds were visualised by irradiation at 254 nm or by staining with anisaldehyde. Thin layer chromatography was performed using Merck Kieselgel 60 F24 (0.22mm thickness, aluminium backed). ^{1}H NMR spectra were measured at 300MHz on a Bruker AC300 spectrometer or 400 MHz on a Bruker DPX400 spectrometer. ^{13}C NMR spectra were measured at 75 MHz and 100 MHz on the same spectrometers respectively. Chemical shifts are given in ppm relative to tetramethylsilane, and *J* values are given in Hz. Low-resolution mass spectra were recorded using electrospray technique on a Fisons VG platform instrument in acetonitrile or a Waters ZMD quadrupole mass spectrometer in methanol or water. High-resolution mass spectra were recorded using electrospray technique using a Bruker APEX III FT-ICR mass spectrometer in methanol, acetonitrile or water. Infrared spectra were recorded on a BIORAD FT-IR using a Golden Gate adapter and BIORAD WIN-IR software or a Satellite FT-IR using a Golden Gate adapter and WIN FIRST-lite software. Absorptions are described as strong (s), medium (m),

broad (b) or weak (w). Melting points were measured on a Gallenkamp electrothermal melting point apparatus and were uncorrected.

1'-O-methyl-3',5'-O-(1,1,3,3,-tetraisopropyldisiloxane-1,3-diyl)-D-ribofuranose, 3

D-Ribose (23.8 g, 159 mmol) was dissolved in distilled methanol (200 ml). A solution of acetyl chloride (0.72 ml) in methanol (50 ml) was added and the mixture was left to stir at rt for 20 hours. Sodium bicarbonate (8.0 g) was added and the reaction mixture was then filtered. The solvent was removed *in vacuo* to give a yellow oil which was coevaporated with anhydrous pyridine and dried under high vacuum before redissolving in anhydrous pyridine (250 ml). The solution was cooled to 0 °C before adding 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane (60.6 ml, 190.4 mmol) dropwise. The reaction mixture was then left to stir at rt for 16 hours, the solvent was removed *in vacuo* and the residue partitioned between ethyl acetate and water. The aqueous layer was separated and the organic layer washed with water, 2M HCl, water again and finally brine then dried (sodium sulfate), filtered and the solvent removed *in vacuo*. The crude product was purified by column chromatography (4:1 hexane:ethyl acetate) to afford **3** as an anomeric mixture (α:β *ca* 2:1) as a colourless oil (46.3 g, 72 %).

R_f (4:1 hexane:ethyl acetate) 0.42; ν_{max}(neat)/ cm^{-1} 2940 (m, C-H str), 2850 (m, C-H str), 1464 (m, C-H bend), 1036 (s); δ_H (400 MHz, $CDCl_3$) includes the following signals due to the α anomer 4.84 (1H, s, $\mathbf{H}^{1'}$), 4.50 (1H, t, J = 5.3 Hz, $\mathbf{H}^{3'}$), 4.00-4.08 (3H, m, $\mathbf{H}^{2'}$, $\mathbf{H}^{4'}$ and $\mathbf{H}^{5'}$), 3.76 (1H, dd, J = 10.8, 8.8 Hz, $\mathbf{H}^{5'}$), 3.31 (3H, s, Me), 2.97 (1H, bs, O$\mathbf{H}^{2'}$), 1.03-1.07 (28H, m, TIPDS) and includes the following signals due to the β anomer 4.90 (1H, d, J = 3.5 Hz, $\mathbf{H}^{1'}$), 4.20 (1H, dd, J = 7.0, 4.1 Hz, $\mathbf{H}^{3'}$), 4.00-4.08 (3H, m, $\mathbf{H}^{2'}$, $\mathbf{H}^{4'}$ and $\mathbf{H}^{5'}$), 3.76 (1H, dd, J = 10.8, 8.8 Hz, $\mathbf{H}^{5'}$), 3.31 (3H, s, Me), 2.97 (1H, bs, O$\mathbf{H}^{2'}$), 1.03-1.07 (28H, m, TIPDS); δ_C (100 MHz, $CDCl_3$) includes the following signals due to the α anomer 107.8 (d, $\mathbf{C}^{1'}$), 83.1 (d, $\mathbf{C}^{4'}$), 76.3 (d, $\mathbf{C}^{2'}$), 75.5 (d, $\mathbf{C}^{3'}$), 66.6 (t, $\mathbf{C}^{5'}$), 55.4 (s, Me) and includes the following signals due to the β anomer 103.0 (d, $\mathbf{C}^{1'}$), 83.6 (d, $\mathbf{C}^{4'}$), 71.8 (d, $\mathbf{C}^{2'}$), 71.5 (d, $\mathbf{C}^{3'}$), 64.5 (t, $\mathbf{C}^{5'}$) and the following indistinguishable signals due to both anomers 18.0, 17.9, 17.8, 17.8, 17.7, 17.6, 17.5, 17.5, 17.5, 17.4, 17.4, 17.3 (q, TIPDS–$\mathbf{C}H_3$) 14.0, 13.8, 13.8, 13.5, 13.3, 13.3, 13.1, 13.0 (d, TIPDS-$\mathbf{C}$H); *m/z* LRMS [ES^+, MeOH] 429 (M+Na^+, 100%); *m/z* HRMS [ES^+, MeOH] found 429.2097 $C_{18}H_{38}O_6Si_2Na$ requires 429.2099.

1'-O-methyl-2'-O-methylethanoyl-3',5'-O-(1,1,3,3,-tetraisopropyldisiloxane-1,3-diyl)-D-ribofuranose, 4

Compound **3** (46.3 g, 114 mmol) was dissolved in anhydrous DMF (250 ml) and cooled to –5 °C (ice/methanol) before adding methyl bromoacetate (55.5 ml, 586 mmol) followed by sodium hydride (60% dispersion in mineral oil) (19.95 g, 498 mmol) portionwise. The reaction mixture was left to warm to rt then stirred at rt for 48 hours. Saturated KCl was added cautiously and the reaction mixture extracted with diethyl ether. The organic layers were combined then dried (sodium sulfate), filtered and the solvent removed *in vacuo* to give upon purification by column chromatography (4:1 hexane:ethyl acetate) **4** as an anomeric mixture (α:β *ca* 2:1) as a colourless oil (33.3 g, 61%).

R_f (4:1 hexane:ethyl acetate) 0.35, 0.14; ν_{max}(neat)/ cm^{-1} 2947 (s, C-H str), 2867 (s, C-H str), 1735(s, C=O str), 1438 (m, C-H bend), 1312 (m, CH_3 str), 1165 (s), 1035 (s); δ_H (400 MHz, $CDCl_3$) includes the following signals due to the α anomer 4.86 (1H, s, $\mathbf{H}^{1'}$), 3.78-4.50 (5H, m, $\mathbf{H}^{2'}$, $\mathbf{H}^{3'}$, $\mathbf{H}^{4'}$ and $\mathbf{H}^{5'}$), 3.71 (3H, s, Me), 3.74 (2H, s, $\mathbf{H}^{1''}$), 3.32 (3H, s, Me), 0.91-1.08 (28H, m, TIPDS) and includes the following signals due to the β anomer 5.01 (1H, d, J = 4.0 Hz, $\mathbf{H}^{1'}$), 3.78-4.50 (5H, m, $\mathbf{H}^{2'}$, $\mathbf{H}^{3'}$, $\mathbf{H}^{4'}$ and $\mathbf{H}^{5'}$), 3.71 (3H, s, Me), 3.73 (2H, s, $\mathbf{H}^{1''}$), 3.47 (3H, s, Me), 0.91-1.08 (28H, m, TIPDS); δ_C (100 MHz, $CDCl_3$) includes the following signals due to the α anomer 168.4 (s, **C**O), 106.3 (d, $\mathbf{C}^{1'}$), 83.6 (d, $\mathbf{C}^{2'}$), 81.3 (d, $\mathbf{C}^{4'}$), 74.5 (d, $\mathbf{C}^{3'}$), 68.7 (t, $\mathbf{C}^{5'}$), 64.2 (t, $\mathbf{C}^{1''}$), 55.1 (q, Me), 51.0 (q, Me) and includes the following signals due to the β anomer 171.3 (s, **C**O), 102.9 (d, $\mathbf{C}^{1'}$), 81.9 (d, $\mathbf{C}^{2'}$), 78.6 (d, $\mathbf{C}^{4'}$), 71.3 (d, $\mathbf{C}^{3'}$), 68.3 (t, $\mathbf{C}^{5'}$), 62.2 (t, $\mathbf{C}^{1''}$), 56.3 (q, Me), 52.2 (q, Me) and the following indistinguishable signals due to both anomers 17.9, 17.8, 17.7, 17.7, 17.6, 17.5, 17.5, 17.5, 17.4, 17.4, 17.3, 17.2 (q, TIPDS–$\mathbf{C}H_3$) 13.6, 13.4, 13.3, 13.3, 13.3, 13.2, 13.1, 12.9 (d, TIPDS-**C**H); *m/z* LRMS [ES^+, MeOH] 501 ($M+Na^+$, 100%); *m/z* HRMS [ES^+, MeOH] found 501.2309 $C_{21}H_{42}O_8Si_2Na$ requires 501.2310.

1'-O-methyl-2'-O-ethoxy-3',5'-O-(1,1,3,3,-tetraisopropyldisiloxane-1,3-diyl)-D-ribofuranose, 5

Compound **4** (33.3 g, 69.7 mmol) was dissolved in anhydrous THF (250 ml). $LiBH_4$ (2.79 g, 128 mmol) was added portionwise and the reaction mixture left to stir at rt for 30 mins. A solution of methanol (20 ml) in THF (150 ml) was added with extreme caution and the reaction mixture was then left to stir at rt for 30 minutes. Methanol (50 ml) was added and the

mixture diluted with diethyl ether. The organic layer was washed with water, then dried (sodium sulfate), filtered and the solvent removed *in vacuo* to give upon purification by column chromatography (4:1 hexane:ethyl acetate) **5** as an anomeric mixture (α:β *ca* 2:1) as a colourless oil (26.8 g, 86 %).

R_f (4:1 hexane:ethyl acetate) 0.19; ν_{max}(neat)/ cm^{-1} 3486 (bw, O-H str), 2944 (s, C-H str), 2867 (s, C-H str), 1465 (s, C-H bend), 1385 (m, CH_3 str), 1031 (s); δ_H (400 MHz, $CDCl_3$) includes the following signals due to the α anomer 4.65 (1H, s, $\mathbf{H}^{1'}$), 4.38 (1H, dd, J = 8.0, 4.5 Hz, $\mathbf{H}^{3'}$), 3.49-4.03 (8H, m, $\mathbf{H}^{2'}$, $\mathbf{H}^{4'}$, $\mathbf{H}^{5'}$, $\mathbf{H}^{1''}$ and $\mathbf{H}^{2''}$), 3.35 (3H, s, Me), 2.31 (1H, bs, O**H**), 0.91-1.08 (28H, m, TIPDS) and includes the following signals due to the β anomer 4.51 (1H, d, J = 8.0 Hz, $\mathbf{H}^{1'}$), 3.20 (3H, s, Me) 3.49-4.03 (9H, m, $\mathbf{H}^{2'}$, $\mathbf{H}^{3'}$, $\mathbf{H}^{4'}$, $\mathbf{H}^{5'}$, $\mathbf{H}^{1''}$ and $\mathbf{H}^{2''}$), 2.31 (1H, bs, O**H**), 0.91-1.08 (28H, m, TIPDS); δ_C (100 MHz, $CDCl_3$) includes the following signals due to the α anomer 106.3 (d, $\mathbf{C}^{1'}$), 84.0 (d, $\mathbf{C}^{2'}$), 79.6 (d, $\mathbf{C}^{4'}$), 73.5 (t, $\mathbf{C}^{5'}$), 72.1 (d, $\mathbf{C}^{3'}$), 66.5 (t, $\mathbf{C}^{1''}$), 63.0 (t, $\mathbf{C}^{2''}$), 56.7 (q, Me) and includes the following signals due to the β anomer 100.7 (d, $\mathbf{C}^{1'}$), 81.5 (d, $\mathbf{C}^{2'}$), 79.6 (d, $\mathbf{C}^{4'}$), 74.0 (t, $\mathbf{C}^{5'}$), 71.5 (d, $\mathbf{C}^{3'}$), 65.3 (t, $\mathbf{C}^{1''}$), 62.1 (t, $\mathbf{C}^{2''}$), 55.0 (q, Me) and the following indistinguishable signals due to both anomers 17.9, 17.8, 17.8, 17.7, 17.7, 17.7, 17.6, 17.4, 17.3 (q, TIPDS–$\mathbf{C}H_3$) 13.8, 13.8, 13.6, 13.1, 13.1, 13.0 (d, TIPDS-**C**H); *m/z* LRMS [ES^+, MeOH] 473 ($M+Na^+$, 100%); *m/z* HRMS [ES^+, MeOH] found 473.2362 $C_{20}H_{42}O_7Si_2Na$ requires 473.2361.

1'-O-methyl-2'-O-(2-phthalimidoethyl)-3',5'-O-(1,1,3,3,-tetraisopropyldisiloxane-1,3-diyl)-D-ribofuranose, 6

Compound **5** (26.8 g, 59.6 mmol) was dissolved in distilled THF (140 ml). Triphenylphosphine (21.2 g, 81 mmol) and phthalimide (11.98 g, 81 mmol) were then added followed by diethylazodicarboxylate (12.9 ml, 81 mmol) as a solution in THF (75 ml). The reaction mixture was left to stir at rt for 2 hours, evaporated and the residue redissolved in the minimum amount of diethyl ether then left at 0 °C for 30 mins. The white solid was filtered off and the filtrate condensed to give upon purification by column chromatography (4:1 hexane:ethyl acetate) **6** as a white solid (31.5 g, 91%).

R_f (4:1 hexane:ethyl acetate) 0.26, 0.16; mpt 222-225 °C (DCM) ν_{max}(neat)/ cm^{-1} 2943 (m, C-H str), 2864 (m, C-H str), 1775 (w, C=O str), 1713 (s, C=O str), 1466 (m, as CH_3 str), 1394 (s, s CH_3 str), 1034 (s, C-O-C str); δ_H (400 MHz, $CDCl_3$) includes the following signals due

to the α anomer 7.84 (2H, dd, *J* = 5.5, 3.0 Hz, C**H**^{Ar}), 7.70 (2H, dd, *J* = 5.5, 3.0 Hz, C**H**^{Ar}), 4.67 (1H, s, **H**$^{1'}$), 4.43 (1H, dd, *J* = 7.5, 4.0 Hz, **H**$^{3'}$), 3.81-3.98 (7H, m, **H**$^{4'}$, **H**$^{5'}$, **H**$^{1''}$ and **H**$^{2''}$), 3.72 (1H, d, *J* = 4.0 Hz, **H**$^{2'}$), 3.29 (3H, s, Me), 0.95-1.08 (28H, m, TIPDS); δ_C (100 MHz, $CDCl_3$) includes the following signals due to the α anomer 168.6 (s, **C**O), 134.3 (d, **C**H^{Ar}), 132.7 (s, **C**H^{Ar}), 123.7 (d, C**H**^{Ar}), 106.5 (d, **C**$^{1'}$), 83.4 (d, **C**$^{2'}$), 81.3 (d, **C**$^{4'}$), 74.1 (d, **C**$^{3'}$) 68.5 (t, **C**$^{1''}$), 64.1 (t, **C**$^{2''}$), 55.1 (q, Me), 38.3 (t, **C**$^{5'}$), 17.9, 17.8, 17.8, 17.8, 17.8, 17.7, 17.5, 17.4 (q, TIPDS–**C**H_3) 14.7, 13.8, 13.7, 13.2 (d, TIPDS-**C**H); *m/z* LRMS [ES$^+$, MeOH] 602 (M+Na$^+$, 100%); *m/z* HRMS [ES$^+$, MeOH] found 602.2585 $C_{28}H_{45}NO_8Si_2Na$ requires 602.2576.

1',3',5'-Tri-O-acetyl-2'-O-(2-phthalimidoethyl)-D-ribofuranose,7

Concentrated sulfuric acid (1.6 ml) was added to a solution of **6** (15.6 g, 26.9 mmol) in glacial acetic acid:acetic anhydride (150 ml, 1:1, v/v). The reaction mixture was stirred at rt for 2 hours then a saturated solution of sodium bicarbonate was slowly added to the green reaction mixture and the organic layers extracted with DCM. The combined organic layers were dried over magnesium sulfate and concentrated *in vacuo* to dryness. The residue was purified by column chromatography (1:1 hexane:ethyl acetate), to yield an anomeric mixture (α:β, 1:2) of **7** as a colourless oil (9.51 g, 79 %).

R_f (1:1 hexane:ethyl acetate) 0.30; ν_{max} (film, cm^{-1}) 2942 (w, CH^{Ar} str), 2865 (w, CH^{Ar} str), 1741 (s, C=O str), 1736 (s, C=O str), 1711 (s, C=O str), 1706 (s, C=O str), 1465 (w, as CH_3 str), 1390 (m, s CH_3 str), 1215 (s, C-C-O str), 1115 (m, C-O-C str), 1064 (m, C-O-C str), 1005 (s, C-O-C str); δ_H (400 MHz, $CDCl_3$) includes the following signals due to the α anomer 7.81-7.83 (2H, m, C**H**^{Ar}), 7.69-7.71 (2H, m, C**H**^{Ar}), 6.08 (1H, s, **H**$^{1'}$), 4.98 (1H, dd, *J* = 7.0, 4.5 Hz, **H**$^{3'}$), 4.20-4.39 (2H, m, **H**$^{4'}$ plus **H**$^{5'}$), 3.94-4.11 (2H, m, **H**$^{5'}$ plus **H**$^{2'}$), 3.73-3.96 (4H, m, **H**$^{1''}$, **H**$^{2''}$), 2.03 (3H, s, OAc), 2.02 (3H, s, OAc), 1.98 (3H, s, OAc) and includes the following signals due to the β anomer 7.81-7.83 (2H, m, C**H**^{Ar}), 7.69-7.71 (2H, m, C**H**^{Ar}), 6.27 (1H, d, *J* = 4.5 Hz, **H**$^{1'}$), 5.12 (1H, dd, *J* = 6.5, 2.5 Hz, **H**$^{3'}$), 4.20-4.39 (2H, m, **H**$^{4'}$ plus **H**$^{5'}$), 3.94-4.11 (2H, m, **H**$^{5'}$ plus **H**$^{2'}$), 3.73-3.96 (4H, m, **H**$^{1''}$, **H**$^{2''}$), 2.05 (3H, s, Me), 2.02 (3H, s, Me), 1.94 (3H, s, Me); δ_C (100 MHz, $CDCl_3$) includes the following signals due to the α anomer 170.6 (s, **C**O), 170.5 (s, **C**O), 169.5 (s, **C**O), 168.2 (s, **C**O), 168.1 (s, **C**O), 134.1 (d, **C**H^{Ar}), 132.3 (s, **C**^{Ar}), 123.4 (d, **C**H^{Ar}), 98.9 (d, **C**$^{1'}$), 81.7 (d, **C**$^{4'}$), 79.4 (d, **C**$^{2'}$), 71.9 (d, **C**$^{3'}$), 68.5 (t, **C**$^{1''}$), 63.8 (t, **C**$^{5'}$), 37.8 (t, **C**$^{2''}$), 21.2 (q, **C**H_3), 20.9 (q, **C**H_3), 20.8 (q, **C**H_3) and

includes the following signals due to the β anomer 170.6 (s, **C**O), 170.4 (s, **C**O), 170.0 (s, **C**O), 169.5 (s, **C**O), 168.1 (s, **C**O), 134.0 (d, **C**H^{Ar}), 132.1 (s, **C**^{Ar}), 123.2 (d, **C**H^{Ar}), 94.7 (d, **C**$^{1'}$), 81.4 (d, **C**$^{4'}$), 77.8 (d, **C**$^{2'}$), 70.1 (d, **C**$^{3'}$), 69.4 (t, **C**$^{1''}$), 63.6 (t, **C**$^{5'}$), 37.7 (t, **C**$^{2''}$), 21.1 (q, **C**H$_3$), 20.8 (q, **C**H$_3$), 20.5 (q, **C**H$_3$); *m/z* LRMS (ES$^+$, MeOH) 472 (M + Na, 100 %); *m/z* HRMS (ES$^+$, MeOH) found 472.1206, $C_{21}H_{23}NO_{10}Na$ requires 472.1214.

2'-O-(2-phthalimidoethyl)-3',5'-O-diacetyl-5-iodouridine, 8

5-Iodouracil (16.80 g, 70.6 mmol) and TMSCl (9.0 ml, 70.9 mmol) were heated at 120 ^oC for 16 hours in anhydrous HMDS (64 ml). The reaction mixture was cooled to rt and then concentrated to dryness *in vacuo* to yield a colourless oil. This residue was redissolved in anhydrous DCE (50 ml) to which was added a solution of **7** (9.51 g, 21.2 mmol) in anhydrous DCE (40 ml). The resulting mixture was cooled to 0 ^oC and trimethylsilyl triflate (12.24 ml, 67.4 mmol) was added slowly dropwise. The reaction mixture was then warmed to rt and stirred for 14 hours. A saturated solution of sodium bicarbonate was added and the organic layers extracted with DCM. The combined organic layers were washed with brine, dried over magnesium sulfate and concentrated *in vacuo* to afford upon purification by column chromotography (1:1 ethyl acetate:toluene) **8** (as a separable anomeric mixture (α:β *ca* 1:3). The desired β anomer was isolated as a pale yellow foam (6.85 g, 52 %).

R_f(1:1 ethyl acetate: toluene) 0.28 (β), 0.18 (α); ν_{max} (film, cm^{-1}) 1771 (w, C=O str), 1704 (s, C=O str), 1613 (m, C-C-O str), 1429 (m, as CH_3), 1394 (s, s CH_3 str), 1224 (s, C-C-O str), 1118 (s, C-C-O str), 1026 (s, C-O-C str); β anomer δ_H (400 MHz, *d*-DMSO) 11.7 (1H, s, N**H**), 8.93 (1H, s, **H**^{6}), 8.73 (4H, s, C**H**^{Ar}), 6.62 (1H, d, *J* = 5.5 Hz, **H**$^{1'}$), 6.05 (1H, t, *J* = 5.0 Hz, **H**$^{3'}$), 5.30 (1H, t, *J* = 5.5, **H**$^{2'}$), 5.07-5.12 (3H, m, **H**$^{4'}$ plus **H**$^{5'}$), 4.53-4.70 (4H, m, **H**$^{1''}$ plus **H**$^{2''}$), 3.00 (3H, s, Me), 2.81 (3H, s, Me); β anomer δ_C (100 MHz, *d*-DMSO) 170.1 (s, **C**O), 169.6 (s, **C**O), 167.8 (s, **C**O), 160.5 (s, **C**^{4}), 150.2 (s, **C**^{2}), 144.9 (d, **C**^{6}), 134.6 (d, **C**H^{Ar}), 131.6 (s, **C**^{Ar}), 123.1 (d, **C**H^{Ar}), 87.8 (d, **C**$^{1'}$), 79.4 (d, **C**$^{4'}$), 78.2 (d, **C**$^{2'}$), 69.7 (s, **C**^{5}), 70.4 (d, **C**$^{3'}$), 67.2 (t, **C**$^{1''}$), 63.1 (t, **C**$^{5'}$), 37.4 (t, **C**$^{2''}$), 19.9 (q, **C**H$_3$), 20.4 (q, **C**H$_3$); *m/z* LRMS (ES$^+$, MeOH) 650 (M+Na$^+$, 100%); *m/z* HRMS (ES$^+$, MeOH) found 650.0249, $C_{23}H_{22}N_3O_{10}INa$ requires 650.0242.

2'-O-(2-phthalimidoethyl)-5'-O-(4,4'-dimethoxytrityl)-5-iodouridine, 10

Compound **8** (6.76 g, 10.8 mmol) was dissolved in anhydrous methanol (85 ml), treated with sodium methoxide (1.69 g, 32 mmol) and stirred at rt for 1 hour under argon. Dowex 50 ion exchange resin (pyridinium form) was added and the mixture was stirred for 30 minutes at rt.

The resin was removed by filtration and the resin washed extensively with methanol. The filtrate was then reduced to dryness *in vacuo* to yield **9** as a yellow foam (5.82 g, >99 %) which was used without further purification. Compound **9** (5.82 g, 10.7 mmol) was dissolved in anhydrous pyridine (50 ml) and to this was added 4',4'-dimethoxytrityl chloride (4.83 g, 14.3 mmol). The reaction mixture was stirred at rt, for 3 hours. Toluene (5 ml) was added and the resulting solution reduced to dryness *in vacuo* to yield an orange foam. This residue was redissolved in DCM and the organic layer washed with water, saturated sodium bicarbonate solution and water. The combined organic layers were dried over sodium sulfate and concentrated to dryness *in vacuo*. The resulting yellow foam was purified by column chromatography (1:1 ethyl acetate:toluene with 1% pyridine), to yield **10** (6.85 g, 76 %) as a pale yellow foam.

R_f (1:1 ethyl acetate:toluene) 0.47; ν_{max} (film, cm^{-1}) 3223 (w, CH^{Ar} str), 3023 (w, CH^{Ar} str), 2926 (w, CH^{Ar} str), 1770 (s, C=O str), 1718 (s, C=O str), 1508 (m, as CH_3), 1395 (s, s CH_3 str), 1250 (s, C-C-O str), 1176 (s, C-O-C str), 1096 (s, C-O-C str), 1030 (s, C-O-C str); δ_H (400 MHz, *d*-DMSO) 11.87 (1H, s, N**H**), 8.16 (1H, s, $\mathbf{H}^{6}$), 8.02 (4H, s, $C\mathbf{H}^{Ar}$), 7.35-7.64 (9H, m, $C\mathbf{H}^{Ar}$), 7.12 (4H, d, J = 8.5 Hz, $C\mathbf{H}^{Ar}$), 5.96 (1H, d, J = 3.0 Hz, $\mathbf{H}^{1'}$), 5.36 (1H, app d, J = 5.5 Hz, $\mathbf{H}^{3'}$), 4.35 (1H, s, $\mathbf{H}^{2'}$), 3.99-4.15 (5H, m, $\mathbf{H}^{4'}$, $\mathbf{H}^{5'}$, $\mathbf{H}^{1''}$), 3.92 (6H, s, OMe), 3.37-3.42 (2H, m, $\mathbf{H}^{2''}$); δ_C (100 MHz, *d*-DMSO) 167.6 (s, **C**O), 159.4 (s, $\mathbf{C}^{4}$), 158.1 (s, $\mathbf{C}^{2}$), 149.1 (s, $\mathbf{C}^{Ar}$), 143.7 (d, $\mathbf{C}^{6}$), 143.3, 136.4, 134.5 (s, $\mathbf{C}^{Ar}$), 134.4 (d, $\mathbf{C}H^{Ar}$), 130.5 (s, $\mathbf{C}^{Ar}$), 128.8, 128.0, 127.3, 127.1, 126.8, 125.8, 124.4, 122.0, 112.4 (d, $\mathbf{C}H^{Ar}$), 86.3 (d, $\mathbf{C}^{1'}$), 85.0 (s, $\mathbf{C}^{Ar}$), 82.4 (d, $\mathbf{C}^{3'}$), 79.5 (d, $\mathbf{C}^{4'}$), 69.2 (s, $\mathbf{C}^{5}$), 67.8 (d, $\mathbf{C}^{2'}$), 65.7 (t, $\mathbf{C}^{5'}$), 62.3 (t, $\mathbf{C}^{1''}$), 54.1 (q, OMe), 36.1 (t, $\mathbf{C}^{2''}$); *m/z* LRMS (ES^+, MeOH) 868 ($M+Na^+$, 100%); *m/z* HRMS (ES^+, MeOH) found 868.1314, $C_{40}H_{36}N_3O_{10}I$ requires 868.1337.

2'-O-(2-phthalimidoethyl)-5'-O-(4,4'-dimethoxytrityl)-5-(3-trifluoroacetamidoprop-1-ynyl)uridine, 11

Compound **10** (6.85 g, 8.11 mmol) was dissolved in anhydrous DMF (130 ml) and to this was added copper iodide (0.46 g, 2.42 mmol), *N*-propargyltrifluoroacetamide(17) (1.44 g, 9.73 mmol) and distilled triethylamine (3.26 ml, 23.4 mmol). The reaction mixture was covered in aluminium foil to exclude light and stirred at rt. After 10 mins, tetrakis(triphenylphosphine)palladium(0) (0.92 g, 0.80 mmol) was added and the reaction was stirred under argon, at rt for a further 5 hours. The reaction mixture was reduced to

dryness *in vacuo* to afford upon purification by column chromatography (1:1 ethyl acetate:toluene with 1% pyridine), **11** (6.34 g, 90 %) as a pale yellow foam.

R_f (2:1 ethyl acetate:toluene), 0.29; ν_{max} (film, cm^{-1}) 3291 (w, CH^{Ar} str), 3220 (w, CH^{Ar} str), 3075 (w, CH^{Ar} str), 2941 (w, CH^{Ar} str), 2841 (w, CH^{Ar} str), 1772 (w, C=O str), 1706 (s, C=O str), 1508 (m, as CH_3), 1394 (m, s CH_3 str), 1249 (s, C-C-O str), 1174 (s, C-O-C str), 1069 (s, C-O-C str), 1031 (s, C-O-C str); δ_H (400 MHz, *d*-DMSO) 11.8 (1H, s, N**H**), 10.17 (1H, t, *J* = 5.5 Hz, N**H**), 8.05 (1H, s, $\mathbf{H}^6$), 8.00 (4H, s, C$\mathbf{H}^{Ar}$), 7.35-7.60 (9H, m, C$\mathbf{H}^{Ar}$), 7.08 (4H, dd, *J* = 2.0, 9.0 Hz, C$\mathbf{H}^{Ar}$), 5.88 (1H, d, *J* = 4.0 Hz, $\mathbf{H}^{1'}$), 5.32 (1H, d, *J* = 6.5 Hz, $\mathbf{H}^{3'}$), 4.32-4.37 (2H, m, $\mathbf{H}^{2'}$ plus $\mathbf{H}^{4'}$), 4.21 (2H, d, *J* = 5.5 Hz, C≡CC$\mathbf{H}_2$), 3.95-4.14 (4H, m, $\mathbf{H}^{5'}$ plus $\mathbf{H}^{1''}$), 3.92 (6H, s, OMe), 3.46 (1H, dd, *J* = 11.0, 5.5 Hz, $\mathbf{H}^{2''}$), 3.26 (1H, dd, *J* = 10.5, 2.5Hz, $\mathbf{H}^{2''}$); δ_C (100 MHz, *d*-DMSO) 167.8 (s, **C**O), 161.5 (s, $\mathbf{C}^5$), 158.0 (s, $\mathbf{C}^2$), 149.3 (s, $\mathbf{C}^{Ar}$), 143.7 (d, $\mathbf{C}^6$), 137.3, 135.6, 135.1 (s, $\mathbf{C}^{Ar}$), 134.3 (d, $\mathbf{C}H^{Ar}$), 131.4 (s, $\mathbf{C}^{Ar}$), 129.3, 129.6, 128.8, 128.1, 127.5, 128.6, 125.3, 122.9 (d, $\mathbf{C}H^{Ar}$), 117.2 (q, $\mathbf{C}F_3$), 114.3, 113.2, 113.2 (d, $\mathbf{C}H^{Ar}$), 98.1 (s, C≡$\mathbf{C}CH_2$), 87.6 (d, $\mathbf{C}^{1'}$), 87.3 (s, $\mathbf{C}$≡CCH_2), 85.8 (s, $\mathbf{C}^{Ar}$), 83.1 (d, $\mathbf{C}^{3'}$), 80.3 (d, $\mathbf{C}^{4'}$), 75.0 (s, $\mathbf{C}^{Ar}$), 68.6 (d, $\mathbf{C}^{2'}$), 66.6 (t, $\mathbf{C}^{5'}$), 63.1 (t, $\mathbf{C}^{1''}$), 55.0 (q, OMe), 37.0 (t, $\mathbf{C}^{2''}$), 29.3 (t, C≡$\mathbf{C}CH_2$); *m/z* LRMS (ES^+, CH_3CN) 891 ($M+Na^+$, 100%); *m/z* HRMS (ES^+, CH_3CN) found 891.2477, $C_{45}H_{39}F_3N_4O_{11}Na$ requires 891.2459.

2'-O-(2-phthalimidoethyl)-3'-O-(2-cyanoethoxy-N,N-diisopropylphosphino)-5'-O-(4,4'-dimethoxytrityl)-5-(3-trifluoroacetamidoprop-1-ynyl)uridine, 1

Compound **11** (2.48 g, 2.86 mmol) was dissolved in anhydrous THF (25 ml). DIPEA (1.31 ml, 7.5 mmol) was added, followed by the dropwise addition of 2-cyanoethoxy-(*N,N*-diisopropylamino)chlorophosphine (0.83 ml, 3.72 mmol), and the reaction mixture was stirred under argon at rt for 5 hours. After dilution with ethyl acetate the reaction mixture was washed with saturated KCl, dried over sodium sulfate and reduced to dryness *in vacuo*. The resulting yellow foam was purified by column chromatography (1:2 ethyl acetate:toluene with 1% pyridine) to give **1** as a colourless foam (2.12 g, 70%).

R_f (1:1 ethyl acetate:toluene) 0.39, 0.48; δ_P (121 MHz, *d*-DMSO) 150.2, 149.5; δ_H (400 MHz, *d*-DMSO) 11.83 (1H, s, N**H**), 10.17 (1H, dt, *J* = 11.5, 5.5 Hz, N**H**), 8.14 (1H, d, *J* = 6.5 Hz, $\mathbf{H}^6$), 8.03 (4H, s, C$\mathbf{H}^{Ar}$), 7.37-7.65 (9H, m, C$\mathbf{H}^{Ar}$), 7.08-7.11 (4H, dd, *J* = 5.0, 8.5 Hz, $\mathbf{C}H^{Ar}$), 5.92 (1H, t, *J* = 5.0 Hz, $\mathbf{H}^{1'}$), 4.53 (1H, dd, *J* = 11.6, 4.5 Hz, $\mathbf{H}^{2'}$), 3.97-4.31 (4H, m, $\mathbf{H}^{3'}$, $\mathbf{H}^{4'}$, O$\mathbf{C}H_2$), 3.94-4.04 (2H, m, C$\mathbf{H}_2$), 3.93 (6H, s, O$\mathbf{C}H_3$), 3.75-3.91 (2H, m, $CH_2C\mathbf{H}_2CN$), 3.64-

3.71 (2H, m, NC**H**$(CH_3)_2$), 3.46-3.55 (2H, m, $\mathbf{H}^{2''}$), 3.44 (1H, dd, *J* = 11.0, 2.5 Hz, $\mathbf{H}^{5'}$), 3.35 (1H, dd, *J* = 10.5, 2.5 Hz, $\mathbf{H}^{5'}$), 2.95 (1H, t, *J* = 6.0 Hz, C$\mathbf{H}_2$CN), 2.78 (1H, t, *J* = 6.0 Hz, C$\mathbf{H}_2$CN), 1.10-1.28 (12H, m, *i*Pr-C$\mathbf{H}_3$); δ_C (100 MHz, *d*-DMSO) 167.5 (s, **C**O), 163.1 (s, **C**O), 157.9 (s, **C**O), 143.8 (d, $\mathbf{C}^6$), 137.0, 134.9 (s, $\mathbf{C}^{Ar}$), 134.1 (d, $\mathbf{C}H^{Ar}$), 131.2 (s, $\mathbf{C}^{Ar}$), 129.5, 128.6, 128.0, 127.6, 127.3 (d, $\mathbf{C}H^{Ar}$), 126.4 (s, **C**N), 125.1 (s, $\mathbf{C}^{Ar}$), 122.7 (d, $\mathbf{C}H^{Ar}$), 114.2 (q, **C**F_3), 113.0 (d, $\mathbf{C}H^{Ar}$), 100.3 (s, $\mathbf{C}^{Ar}$), 88.1 (s, $\mathbf{C}^{Ar}$), 87.7 (d, $\mathbf{C}^{1'}$), 83.3 (s, $\mathbf{C}^{Ar}$), 82.7 (d, $\mathbf{C}^{3'}$), 79.9 (d, $\mathbf{C}^{4'}$), 72.4 (s, $\mathbf{C}^5$), 70.1 (d, $\mathbf{C}^{2'}$), 66.5 (t, $\mathbf{C}^{5'}$), 62.5 (t, $\mathbf{C}^{1''}$), 57.7 (t, O**C**$H_2$$CH_2$CN), 54.1 (q, O**C**$H_3$), 39.6 (d, N**C**H$(CH_3)_2$), 36.5 (t, $\mathbf{C}^{2''}$), 29.1 (t, **C**H_2), 24.1 (q, **C**H_3), 19.6 (t, **C**H_2CN); *m/z* LRMS (ES^+, CH_3CN) 1069 ($M+H^+$, 100 %).

Preparation of synthetic oligonucleotides

All oligonucleotides were synthesised on an Applied Biosystems 394 automated DNA/RNA synthesiser using the standard 0.2 μmole phosphoramidite cycle of acid-catalysed detritylation, coupling, capping and iodine oxidation. Stepwise coupling efficiencies and overall yields were determined by the automated trityl cation conductivity monitoring facility and in all cases were >98.0%. All β-Cyanoethyl phosphoramidite monomers were dissolved in anhydrous acetonitrile to a concentration of 0.1 M immediately prior to use. Standard DNA phosphoramidites, solid supports and additional reagents were purchased from Link Technologies or Applied Biosystems Ltd. The monomers described in this publication were treated as follows: After purification by column chromatography the monomer was dissolved in anhydrous acetonitrile and filtered through a Millipore Millex®-FH syringe filter (0.45 μm, 25mm). The solvent was then removed and the monomer redissolved in anhydrous DCM. Aliquots corresponding to 100 μmoles were transferred to ABI monomer reagent bottles and dried in a desiccator overnight under high vacuum before being stored under a positive pressure of argon at –20 °C. Oligonucleotides were cleaved and deprotected at rt for 36 hours using 2 ml of 10% $MeNH_2$ in water containing 2.5 mg/ml phenol. Purification was carried out by reversed phase HPLC on a Gilson system using an ABI Aquapore column (C8), 8 mm x 250 mm, pore size 300 Å. The system was controlled by Gilson 7.12 software

and the following protocol was used: Run time 30 minutes, flow rate 4ml per min, binary system, gradient: Time in mins (% buffer B);0 (0); 3(0); 5(20); 21 (100); 25(100); 27 (0); 30(0). Elution buffer A 0.1 M ammonium acetate, pH 7.0, buffer B 0.1 M ammonium acetate with 35% acetonitrile pH 7.0. Elution of oligonucleotides was monitored by ultraviolet absorption at 295 nm. After HPLC purification oligonucleotides were desalted using disposable NAP 10 Sephadex columns (Pharmacia), aliquoted into eppendorf tubes and stored at –20 ºC. Purified oligonucleotides were analysed by MALDI-TOF MS using a ThermoBioAnalysis Dynamo MALDI-TOF mass spectrometer in positive ion mode (Table S1).

Chemical synthesis

The bis-aminouridine phosphoramidite monomer **1** was synthesized from D-ribose and 5-iodouacil in 11-steps (scheme 1). Commercially available D-ribose was first converted into the methyl glycoside by stirring at rt overnight with 1% methanolic hydrochloric acid in methanol (formed by the addition of acetyl chloride to anhydrous methanol). The 3' and 5' hydroxyl groups were then protected with TIPDS using Markiewicz reagent (26) in pyridine to afford compound **3** as an anomeric mixture (α:β *ca* 2:1) in 72% yield. Introduction of the protected 2'aminoethoxy substituent was achieved in three steps. Firstly alkylation of the 2' position with an excess of methyl bromoacetate and sodium hydride in DMF at –5 ^{o}C gave **4** (61%). Secondly the ester moiety of **4** was reduced with $LiBH_4$ in THF to give alcohol **5** (86%) and thirdly the hydroxyl group of **5** was substituted by phthalimide under Mitsunobu conditions (27) (91%). The phthalimide-protected 2'aminoethoxy moiety was stable to the remaining steps of the synthesis and is compatible with oligonucleotide synthesis and deprotection (28). Treatment of compound **6** with acetic anhydride, acetic acid and concentrated sulfuric acid gave the tri-*O*-acetyl compound **7**. This was converted to the

protected nucleoside **8** (α and β-anomers 1:3 ratio) by reaction with 5-iodouracil under Vorbrüggen conditions (29). The β-anomer of **8** was deacetylated with sodium methoxide in methanol to give nucleoside **9,** then protected selectively at the 5'-position by reaction with 4,4'-dimethoxytrityl chloride in pyridine to afford compound **10** in 90 % yield. Addition of the 3-trifluoroacetamidoprop-1-yne (30) moiety was achieved by Sonogashira coupling (31,32) to give compound **11** which was converted to the target monomer **1** by treatment with 2-cyanoethoxy-(*N,N*-diisopropylamino)chlorophosphine in an argon atmosphere (70 %). The above procedure was employed in the synthesis of multi-gram quantities of phosphoramidite **1** and this monomer was used to synthesise a series of triplex forming oligonucleotides (TFOs) using standard solid-phase oligonucleotide synthesis conditions. Oligonucleotides were cleaved from the solid support and deprotected using 10% methylamine in water containing 2.5 mg/ml phenol. The phenol scavenger prevents cyanoethylation of the N(3)-position of BAU with acrylonitrile liberated by deprotection of the phosphodiesters.

Table S1. Positive ion MALDI-TOF mass spectra of modified oligonucleotides. (F = 6-amidohexylfluorescein)

TFO	**Sequence (5'-3')** **X = (BAU)**	**Calculated mass ($M+H^+$)**	**Observed mass ($M+H^+$)**
1	**F-**TCTCTCTT**X**TCCTCCTCC	5781.9	5782.4
2	TCTC**X**CTT**X**TCC**X**CCTCC	5573.8	5574.6
3	TC**X**C**X**CTT**X**TCC**X**CC**X**CC	5770.0	5768.6
4	**X**C**X**C**X**T**X**T**X**TCT	4034.9	4035.5

Table S2. T_m values (°C) determined for the melting of the triplexes shown in Figure 2a. The T_ms 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. The values in parentheses are the T_ms derived from the annealing profiles. f indicates a fast rate of heating and cooling (0.1 °C/sec) while s indicates a slow rate (0.2 °C/min). – indicates that the T_m was too low to measure (< 25 °C). *The melting and annealing T_ms for the BAU.AT complex were 69.0 °C and 66.0 °C when the rate of temperature change was reduced to 0.067 °C/min.

pH	N =		A	G	C	T	BAU*
5.0	N.AT	f	58.5 (56.9)	55.3 (54.7)	55.8 (55.4)	69.3 (58.5)	77.5 (58.4)
		s	57.3 (57.6)	54.7 (54.4)	55.3 (55.2)	63.8 (63.7)	70.8 (64.4)
	N.TA	f	55.6 (54.9)	61.3 (56.8)	54.2 (54.2)	54.1 (54.2)	55.0 (55.7)
		s	54.7 (54.8)	58.8 (58.9)	53.4 (53.7)	53.8 (53.9)	55.9 (55.9)
	N.GC	f	63.8 (59.9)	58.7 (57.8)	73.1 (60.3)	58.4 (57.8)	66.5 (55.9)
		s	61.8 (61.9)	58.1 (58.1)	67.4 (65.9)	57.8 (57.8)	62.8 (63.0)
	N.CG	f	55.8 (56.0)	57.3 (56.9)	58.1 (57.7)	58.4 (58.2)	57.2 (56.1)
		s	55.7 (55.5)	56.8 (56.6)	57.6 (57.7)	58.1 (58.2)	56.7 (56.8)
5.5	N.AT	f	46.0 (45.1)	40.8 (39.3)	41.4 (40.4)	56.9 (54.3)	64.6 (55.3)
		s	44.4 (44.2)	39.0 (38.8)	39.8 (39.4)	54.5 (54.7)	59.5 (59.2)
	N.TA	f	42.2 (41.1)	48.9 (47.4)	40.1 (39.4)	40.4 (39.6)	43.8 (42.5)
		s	40.7 (40.4)	47.3 (47.2)	38.7 (38.6)	39.2 (38.8)	41.8 (41.3)
	N.GC	f	49.6 (48.2)	45.4 (44.1)	58.8 (55.7)	45.1 (44.1)	53.5 (52.0)
		s	48.4 (48.0)	44.1 (43.9)	56.6 (56.7)	43.7 (43.9)	51.8 (51.6)
	N.CG	f	41.9 (41.0)	43.5 (42.4)	44.8 (43.0)	45.4 (44.7)	43.3 (42.5)
		s	40.7 (40.4)	41.9 (41.6)	43.7 (43.4)	44.5 (44.3)	42.0 (41.9)
6.0	N.AT	s	- (-)	- (-)	- (-)	34.8 (34.8)	41.2 (41.0)
	N.TA	s	- (-)	- (-)	- (-)	- (-)	- (-)
	N.GC	s	- (-)	- (-)	34.8 (34.9)	- (-)	30.0 (30.3)
	N.CG	s	- (-)	- (-)	- (-)	- (-)	- (-)