Enhanced H-bonding and π-stacking in DNA: A potent duplex-stabilising and mismatch sensing nucleobase analogue

Chenguang Lou,[a] Andre Dallmann,[c] Pietro Marafini,[b] Rachel Gao[a] and Tom Brown[b]*

[a] School of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ, UK

[b] Department of Chemistry, University of Oxford, Chemistry Research Laboratory, 12 Mansfield Road, Oxford, OX1 3TA, UK

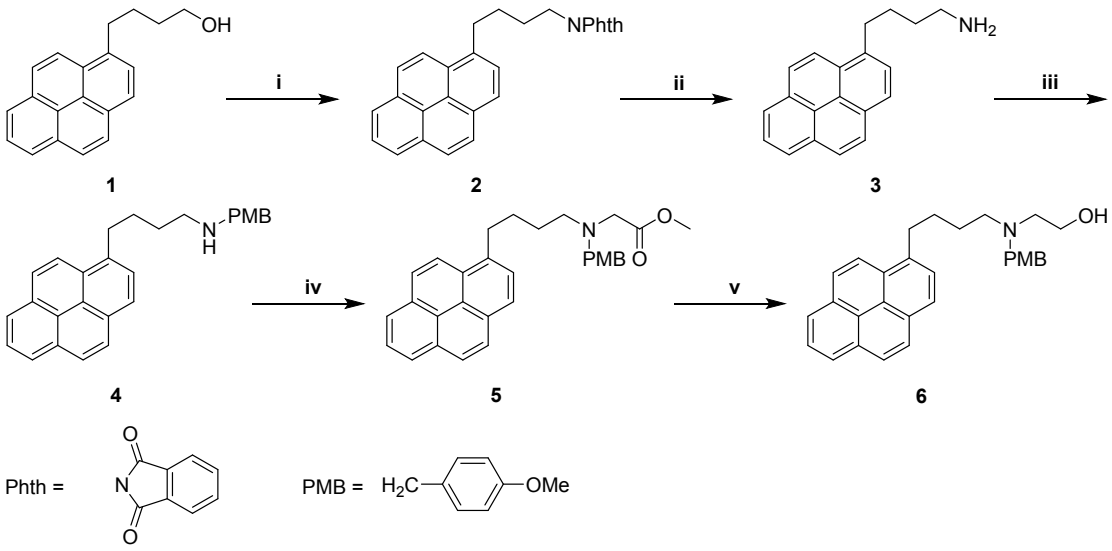
[c] Institute of Structural Biology, Helmholtz Zentrum München, 85764 Neuherberg, Germany and Center for Integrated Protein Science Munich

*To whom correspondence should be sent: Tel: +44(0)1865 275413, e-mail: tom.brown@chem.ox.ac.uk

S1 General Experimental

All reagents used were purchased from Aldrich, Fluka, Molekula, Apollo or Lancaster and used without purification. DNA phosphoramidite monomers, G-clamp monomer, solid supports and additional reagents were purchased from Link Technologies Ltd, Glen Research or Applied Biosystems Ltd. Dichloromethane (DCM), acetonitrile (CH_3CN), *N*,*N*-diisopropylethylamine (DIPEA), triethylamine (Et_3N) and pyridine were distilled over calcium hydride. All reactions were carried out under an argon atmosphere using glassware that had been dried at 120 ˚C overnight. Column chromatography was carried out under pressure using Fisher Scientific DAVISIL 60A 30-70 micron silica. Thin layer chromatography (TLC) was performed using Merck Kieselgel 60 F_{254} (0.22 mm thickness, aluminium backed). Compounds were visualized at 254 nm or stained with 10 % sulfuric acid in EtOH. ^{1}H-NMR spectra were measured at 400 MHz on a Bruker DPX 400 spectrometer. ^{13}C-NMR spectra were measured at 100 MHz on the same spectrometers. Chemical shifts are given in ppm and *J* values are given in Hz. All assignments for ^{1}H-NMR and ^{13}C-NMR have been confirmed by H-H COSY, HMQC and HMBC. ^{31}P NMR spectra were recorded on a Bruker AV 300 spectrometer at 121 MHz. DMSO-d^6, $CDCl_3$ or CD_3CN were used as solvents. Low resolution mass spectra were recorded in acetonitrile using the electrospray technique on a Fisons VG platform instrument. High resolution mass spectra were recorded in acetonitrile or methanol using the electrospray technique on a Bruker APEX III FT-ICR mass spectrometer. HPLC grade CH_3CN was used as the solvent.

S2. Synthesis of 4-(1-pyrenyl)-N-(4-methoxybenzyl)-N-(ethoxy)-1-butylamine:



Scheme S1. *Reagents and conditions:* (i) phthalimide, PPh_3, DEAD, THF, RT, 1 h; (ii) H_2NNH_2-H_2O, EtOH/DCM (1:1), RT, 72 h, 87 % for two steps from **1** to **3**; (iii) 4-methoxybenzyl chloride, NaH, DMF, 0 ˚C, 24 h, 43 %; (iv) methyl bromoacetate, NaH, DMF, 0 ˚C, 1 h; (v) $LiBH_4$, THF, RT, 72 h, 49 % for two steps from **4** to **6**.

4-(1-pyrenyl)-1-phthalimidobutane (2)

4-(1-pyrenyl)-1-butanol (5.00 g, 18.2 mmol) was dissolved in anhydrous THF (50 mL). Triphenylphosphine (5.75 g, 21.9 mmol) and phthalimide (3.22 g, 21.9 mmol) was then added, followed by diethyl azodicarboxylate (DEAD, 3.45 mL, 21.9 mmol, dissolved in 1 mL THF). The reaction mixture was stirred at room temperature for 1 h before methanol (2 mL) was added to quench the reaction. After most of the solvents was removed *in vacuo*, the residue was partitioned between DCM (120 mL) and water (60 mL). The organic layer was separated and washed with water (3 × 50 mL), saturated sodium bicarbonate (50 mL), brine (50 mL), then dried (on sodium sulfate), filtered and the solvent removed *in vacuo* to give a yellow solid (7.15 g, crude) upon purification by silica gel column chromatography (50 % DCM in petroleum ether), which was used for next step without further purification. R_f = 0.35 (ethyl acetate/hexane, 1:3)

4-(1-pyrenyl)-1-butylamine (3)

2 (7.15 g, crude) was stirred with hydrazine monohydrate (in excess, 3.5 mL, 72.1 mmol) in EtOH/DCM solution (250 mL, 1:1) at room temperature for 72 h. After filtration, the filtrate was evaporated *in vacuo* to remove the solvents and the residue was purified on silica gel column chromatography (3.5 % methanol in DCM plus 1% Et_3N) to give a syrup (4.34 g, 87 % for two steps). R_f = 0.19 (methanol/DCM, 1:15, plus 5 % Et_3N) ^{1}H NMR: (400 MHz, DMSO-*d*^{6}) δ ppm 8.41-7.89 (m, 9H, C<u>H</u>-Ar), 3.40-3.28 (m, C<u>H</u>$_2$, 2H), 3.14 (s, N<u>H</u>$_2$, 2H), 2.66 (t, *J* = 7.0 Hz, C<u>H</u>$_2$, 2H), 1.88-1.76 (m, 2H, C<u>H</u>$_2$), 1.61-1.50 (m, 2H, C<u>H</u>$_2$). ^{13}C NMR: (100 MHz, DMSO-*d*^{6}) δ ppm 137.0 (<u>C</u>-Ar), 136.2 (<u>C</u>-Ar), 129.0 (<u>C</u>H-Ar), 127.5 (<u>C</u>H-Ar), 127.1 (<u>C</u>H-Ar), 127.0 (<u>C</u>H-Ar), 126.4 (<u>C</u>H-Ar), 126.1 (<u>C</u>H-Ar), 124.9 (<u>C</u>H-Ar), 124.7 (<u>C</u>H-Ar), 123.5 (<u>C</u>H-Ar), 41.1 (<u>C</u>H$_2$), 32.5 (<u>C</u>H$_2$), 32.3 (<u>C</u>H$_2$), 28.8 (<u>C</u>H$_2$). *m/z* (%) LRMS [ES$^+$, MeCN]: 274.3 ([M+H]$^+$, 100 %). HRMS [ES$^+$]: $C_{20}H_{20}N$ calculated 274.1590 found 274.1585.

4-(1-pyrenyl)-*N*-(4-methoxybenzyl)-1-butylamine (4)

A solution of **3** (6.12 g, 22.4 mmol) and 4-methoxybenzyl chloride (3.20 mL, 23.6 mmol) in anhydrous DMF (150 mL) was cooled to 0 ˚C before sodium hydride (1.79 g, 44.7 mmol, 60 % dispersion in mineral oil) was added portion-wise. The reaction mixture was stirred at 0 ˚C for 24 h. Water (5 mL) was added to quench the reaction. After most of the solvents were removed *in vacuo*, the residue was redissolved in DCM (150 mL) and washed with water (3 × 60 mL), saturated sodium bicarbonate (60 mL) and brine (60 mL), then dried (on sodium sulfate), and filtered. The filtrate was evaporated *in vacuo* to give a yellow gum (3.81 g, 43 %) upon purification by silica gel column chromatography (6 % methanol in DCM). R_f = 0.29 (methanol/DCM, 1:15, plus 5 % Et_3N) ^{1}H NMR: (400 MHz, DMSO-*d*^{6}) δ ppm 8.37-7.88 (m, 9H, C<u>H</u>-Ar), 7.21 (d, *J* = 8.4 Hz, 2H, C<u>H</u>-Ar), 6.83 (d, *J* = 8.4 Hz, 2H, C<u>H</u>-Ar), 3.71 (s, 3H, C<u>H</u>$_3$), 3.61 (s, 2H, C<u>H</u>$_2$), 3.30 (t, *J* = 7.6 Hz, 2H, C<u>H</u>$_2$), 2.58-2.48 (m, 2H, C<u>H</u>$_2$), 1.85-1.74 (m, 2H C<u>H</u>$_2$), 1.62-1.53 (m, 2H, C<u>H</u>$_2$). ^{13}C NMR: (100 MHz, DMSO-*d*^{6}) δ ppm 158.0 (<u>C</u>-Ar), 137.0 (<u>C</u>-Ar), 132.4 (<u>C</u>-Ar), 129.1 (<u>C</u>H-Ar), 127.4 (<u>C</u>H-Ar), 127.1 (<u>C</u>H-Ar), 126.4 (<u>C</u>H-Ar), 126.1 (<u>C</u>H-Ar), 124.9 (<u>C</u>H-Ar), 124.7 (<u>C</u>H-Ar), 123.5 (<u>C</u>H-Ar), 113.4

(CH-Ar), 54.9 ($\underline{C}H_3$), 52.2 ($\underline{C}H_2$), 48.2 ($\underline{C}H_2$), 32.5 ($\underline{C}H_2$), 29.2 ($\underline{C}H_2$), 29.1 ($\underline{C}H_2$). *m/z* (%) LRMS [ES^+, MeCN]: 394.2 ($[M+H]^+$, 100 %). HRMS[ES^+]: $C_{28}H_{28}NO$ calculated 394.2165 found 394.2171.

4-(1-pyrenyl)-*N*-(4-methoxybenzyl)-*N*-(methylethanoyl)-1-butylamine (5)

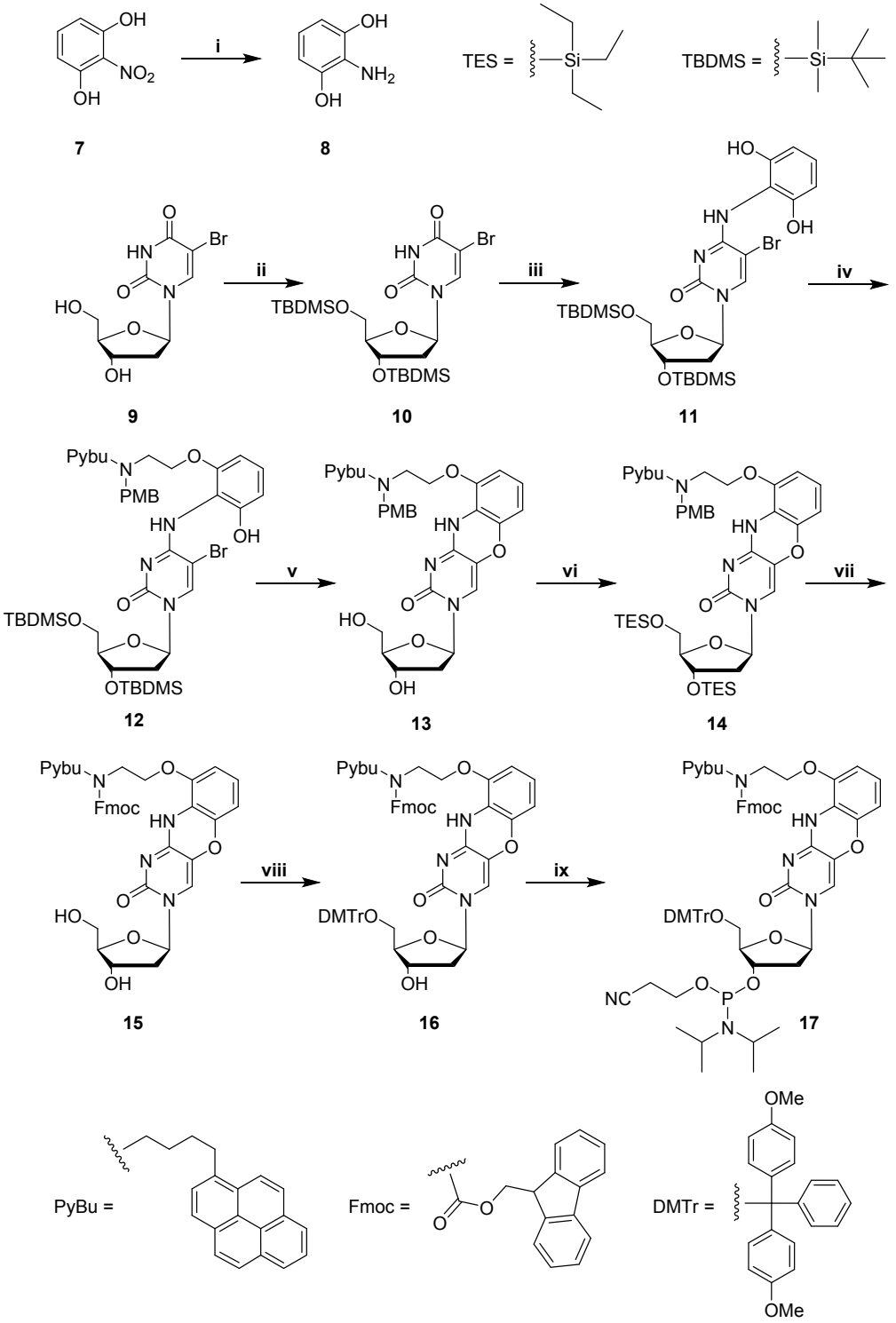
A solution of **4** (3.81 g, 9.68 mmol) and methyl bromoacetate (2.00 mL, 21.0 mmol) in anhydrous DMF (60 mL) was cooled to - 5 ˚C (ice/methanol bath) before sodium hydride (0.775 g, 19.4 mmol, 60 % dispersion in mineral oil) was added portion-wise. The reaction mixture was stirred at 0 ˚C for 1 h. Methanol (3 mL) was added dropwise to quench the reaction. After most of the solvent was removed *in vacuo*, the residue was partitioned between DCM (100 mL) and water (60 mL). The organic layer was washed with water (2 × 50 mL), saturated sodium bicarbonate (50 mL), brine (50 mL), then dried over anhydrous sodium sulfate. After filtration, the solution was evaporated *in vacuo* to give a yellow oil (4.30 g, crude) upon purification by silica gel column chromatography (14 % ethyl acetate in petroleum ether), which was used for next step without further purification. R_f = 0.55 (ethyl acetate/hexane, 1:2).

4-(1-pyrenyl)-*N*-(4-methoxybenzyl)-*N*-(ethoxy)-1-butylamine (6)

A solution of **5** (4.30 g, crude) in anhydrous THF (60 mL) was cooled to 0 ˚C before $LiBH_4$ (0.603 g, 27.7 mmol) was added portionwise. The reaction mixture was warmed to room temperature and stirred at the same temperature for 72 h before methanol (5 mL) was added dropwise to quench the reaction. The solvents were removed *in vacuo* and the residue was partitioned between DCM (90 mL) and water (60 mL). The organic layer was separated and washed with water (3 × 40 mL), saturated sodium bicarbonate (40 mL) and brine (40 mL), then dried (on sodium sulfate), filtered and the solvent removed *in vacuo* to give a clear oil (2.08 g, 49 % for two steps) upon purification by silica gel column chromatography (3 % methanol in DCM). R_f = 0.41 (methanol/DCM, 1:10) ^{1}H NMR: (400 MHz, DMSO-d^6) δ ppm 8.36-8.02 (m, 8H, CH-Ar), 7.90 (d, *J* = 7.8 Hz, 1H, CH-Ar), 7.15 (d, *J* = 8.4 Hz, 2H,

C<u>H</u>-Ar), 6.78 (d, *J* = 8.4 Hz, 2H, C<u>H</u>-Ar), 4.29 (s, 1H, O<u>H</u>), 3.68 (s, 3H, C<u>H</u>$_3$), 3.49-3.41 (m, 4H, C<u>H</u>$_2$, C<u>H</u>$_2$), 3.28 (t, *J* = 7.5 Hz, 2H, C<u>H</u>$_2$), 2.53-2.42 (m, 4H, C<u>H</u>$_2$, C<u>H</u>$_2$), 1.82-1.71 (m, 2H C<u>H</u>$_2$), 1.64-1.54 (m, 2H, C<u>H</u>$_2$). ^{13}C NMR: (100 MHz, DMSO-*d*^{6}) δ ppm 158.0 (<u>C</u>-Ar), 137.1 (<u>C</u>-Ar), 131.5 (<u>C</u>-Ar), 130.9 (<u>C</u>-Ar), 130.4 (<u>C</u>-Ar), 129.7 (<u>C</u>H-Ar), 127.4 (<u>C</u>H-Ar), 127.1 (<u>C</u>H-Ar), 126.4 (<u>C</u>H-Ar), 126.1 (<u>C</u>H-Ar), 124.9 (<u>C</u>H-Ar), 124.7 (<u>C</u>H-Ar), 123.5 (<u>C</u>H-Ar), 113.3 (<u>C</u>H-Ar), 59.2 (<u>C</u>H$_2$), 57.8 (<u>C</u>H$_2$), 55.6 (<u>C</u>H$_2$), 54.9 (<u>C</u>H$_3$), 53.4 (<u>C</u>H$_2$), 32.4 (<u>C</u>H$_2$), 29.2 (<u>C</u>H$_2$), 26.5 (<u>C</u>H$_2$). *m/z* (%) LRMS [ES$^+$, MeCN]: 438.2 ([M+H]$^+$, 100 %). HRMS [ES$^+$]: $C_{30}H_{32}NO_2$ calculated 438.2428 found 438.2436.

S3: Synthesis of *N*-(pyrenylbutyl)-G-clamp phosphoramidite monomer:



Scheme S2. *Reagents and conditions:* (i) Pd/C, H_2, MeOH, 50 ˚C, 12 h, 93 %; (ii) TBDMSCl, imidazole, DMF, RT,1.5 h, 94 %; (iii) PPh_3, CCl_4, DCM, reflux, 16 h; **8**, DBU, DCM, RT,18 h, 47 % for two steps; (iv) **6**, PPh_3, DEAD, DCM, RT, 2 h, 81 %; (v) KF, ethanol, 82 ˚C, 24 h, 54 %; (vi) TESCl, imidazole, DMF, RT, 24 h, 86 %; (vii) α-chloroethyl chloroformate, DCM, RT, 1 h; methanol, reflux, 2 h; Fmoc-OSu, pyridine, RT, 0.5 h, 48 % for three steps; (viii) DMTrCl, pyridine, RT, 6 h, 74 %; (ix) 2-cyanoethyl-*N*,*N*-diisopropyl chlorophosphine, DIPEA, DCM, 0.5 h, RT, 79 %.

2-aminoresorcinol (8)

A suspension of 2-nitroresorcinol (**7**, 1.55 g, 10.0 mmol) and Pd/C (0.155 g, 10 % Pd on carbon) in methanol (20 mL) was stirred under hydrogen atmosphere at 50 ˚C for 12 h. The reaction mixture was filtered to remove the catalysts. After the solvent was removed *in vacuo*, the residue was purified by silica gel column chromatography (45 % ethyl acetate in petroleum ether) to afford a pale yellow solid (1.17 g, 93 %) R_f = 0.26 (ethyl acetate/hexane, 1:1) ^{1}H NMR: (400 MHz, DMSO-*d*^{6}) δ ppm 8.82 (s, 2H, O<u>H</u>), 6.30-6.21 (m, 3H, C<u>H</u>-Ar), 3.83 (s, 2H, NH_2). ^{13}C NMR: (100 MHz, DMSO-*d*^{6}) δ ppm 144.9 (<u>C</u>-Ar), 123.8 (<u>C</u>-Ar), 115.8 (<u>C</u>H-Ar), 106.6 (<u>C</u>H-Ar). LRMS [ES$^+$, MeOH]: 126.1 ([M+H]$^+$, 100 %). HRMS [ES$^+$]: $C_6H_8NO_2$ calculated 126.0550 found 126.0547.

2′-deoxy-3′,5′-*O*-di-(*tert*-butyldimethylsilyl)-5-bromouridine (10)

A solution of 5-bromo-2′-deoxyuridine (4.00 g, 13.0 mmol), *tert*-butyldimethylsilyl chloride (TBDMSCl, 4.40 g, 29.2 mmol) and imidazole (4.00 g, 58.7 mmol) in DMF (50 mL) was stirred room temperature for 1.5 h. The reaction mixture was evaporated *in vacuo* to remove solvents. The residue was redissolved in DCM (100 mL) and washed with water (3 × 50 mL), saturated sodium bicarbonate (50 mL) and brine (50 mL), then dried (on sodium sulfate), filtered and the solvent removed *in vacuo* to give a clear oil (6.57 g, 94 %) upon purification by silica gel column chromatography (14 % ethyl acetate in petroleum ether). R_f = 0.75 (ethyl acetate/hexane, 1:1) ^{1}H NMR: (400 MHz, DMSO-*d*^{6}) δ ppm 11.76 (s, 1H, N<u>H</u>), 7.89 (s, 1H, C<u>H</u>$^{\mathbf{6}}$), 6.01 (t, *J* = 6.7 Hz, 1H, C<u>H</u>$^{\mathbf{1'}}$), 4.31-4.25 (m, 1H, C<u>H</u>$^{\mathbf{3'}}$), 3.77-3.70 (m, 2H, C<u>H</u>$^{\mathbf{4'}}$, C<u>H</u>H$^{\mathbf{5'}}$), 3.64 (dd, *J* = 10.6, 2.6 Hz, 1H, CH<u>H</u>$^{\mathbf{5'}}$), 2.20-2.12 (m, 1H, C<u>H</u>H$^{\mathbf{2'}}$), 2.08-2.01 (m, 1H, CH<u>H</u>$^{\mathbf{2'}}$), 0.82 (s, 9H, C<u>H</u>$_3$, C<u>H</u>$_3$, C<u>H</u>$_3$), 0.79 (s, 9H, C<u>H</u>$_3$, C<u>H</u>$_3$, C<u>H</u>$_3$), 0.02 (d, *J* = 1.07 Hz, 6H, C<u>H</u>$_3$, C<u>H</u>$_3$), 0.00 (s, 6H, C<u>H</u>$_3$, C<u>H</u>$_3$). ^{13}C NMR: (100 MHz, DMSO-*d*^{6}) δ ppm 159.0 (<u>C</u>-Ar), 149.6 (<u>C</u>-Ar), 139.4 (<u>C</u>H$^{\mathbf{6}}$), 87.0 (<u>C</u>H$^{\mathbf{4'}}$), 84.8 (<u>C</u>H$^{\mathbf{1'}}$), 71.7 (<u>C</u>H$^{\mathbf{3'}}$), 62.5 (<u>C</u>H$_2^{\mathbf{5'}}$), 39.9 (<u>C</u>H$_2^{\mathbf{2'}}$), 25.5 (<u>C</u>H$_3$), 25.6 (<u>C</u>H$_3$), -5.44 (<u>C</u>H$_3$). LRMS [ES$^+$, MeCN]: 557.2 ([M+Na]$^+$,

42 %), 559.1 ([M+Na]$^+$, 50 %), 1093 [[2M+Na]$^+$, 100 %], 1095.60 [[2M+Na]$^+$, 62 %]. HRMS [ES$^+$]: $C_{21}H_{40}BrN_2O_5Si_2$ calculated 537.1639 found 537.1634.

2′-deoxy-3′,5′-*O*-di-(*tert*-butyldimethylsilyl)-4-*N*-(2,6-dihydroxyphenyl)--5-bromouridine (11)

10 (0.514 g, 0.960 mmol) was dissolved in a mixture of anhydrous DCM and CCl_4 (8 mL, 1:1) and to the solution triphenylphosphine was added (0.374 g , 1.42 mmol). The reaction mixture was refluxed under argon atmosphere for 4 h. Extra triphenylphosphine (0.400 g, 1.52 mmol) was added and the resultant suspension was further refluxed for 12 h before cooling down to room temperature. To this was added **8** (0.404 g, 3.23 mmol) and 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU, 0.485 mL, 3.24 mmol). The resulting solution was stirred under argon atmosphere for 18 h. DCM (50 mL) was added to dilute the solution and the organic layer was washed with 5 % citric acid solution (30 mL), dried over anhydrous sodium sulfate and filtered. The filtrate was evaporated *in vacuo* to remove solvents and the residue was purified by silica gel column chromatography (20 % ethyl acetate in petroleum ether) to give a pale yellow solid (0.288 g, 47 %). R_f = 0.61 (ethyl acetate/hexane, 1:1) ^{1}H NMR: (400 MHz, DMSO-*d*^{6}) δ ppm 9.75-9.65 (b, 2H, OH), 8.14-7.97 (b, 1H, NH), 7.91 (s, 1H, CH^{6}), 6.83 (t, *J* = 8.2 Hz, 1H, CH-Ar), 6.29 (d, *J* = 8.1 Hz, 2H, CH-Ar), 6.01 (t, *J* = 6.6 Hz, 1H, CH$^{1'}$), 4.30-4.25 (m, 1H, CH$^{3'}$), 3.80-3.72 (m, 2H, CH$^{4'}$, CHH$^{5'}$), 3.65 (dd, *J* = 11.1, 3.0 Hz, 1H, CHH$^{5'}$), 2.13-2.01 (m, 2H, CH$_2^{2'}$), 0.83 (s, 9H, CH$_3$, CH$_3$, CH$_3$), 0.79 (s, 9H, CH$_3$, CH$_3$, CH$_3$), 0.04 (d, *J* = 3.00 Hz, 6H, CH$_3$, CH$_3$), 0.00 (s, 6H, CH$_3$, CH$_3$). ^{13}C NMR: (100 MHz, DMSO-*d*^{6}) δ ppm 158.3 (C-Ar), 152.9 (C-Ar), 152.7 (C-Ar), 140.9 (CH^{6}), 127.3 (CH-Ar), 113.4 (C-Ar), 107.2 (CH-Ar), 87.1 (CH$^{4'}$), 85.5 (CH$^{1'}$), 71.8 (CH$^{3'}$), 62.5 (CH$_2^{5'}$), 40.6 (CH$_2^{2'}$), 25.9 (CH$_3$), 25.6 (CH$_3$), -5.41 (CH$_3$). LRMS [ES$^+$, MeCN]: 664.1 ([M+Na]$^+$, 58 %), 666.2 ([M+Na]$^+$, 60 %), 705.1 ([M+MeCN+Na]$^+$, 82 %), 707.3 ([M+MeCN+Na]$^+$, 100 %). HRMS [ES$^+$]: $C_{27}H_{45}BrN_3O_6Si_2$ calculated 644.2010 found 644.2004.

2′-Deoxy-3′,5′-*O*-di-(*tert*-butyldimethylsilyl)-4-*N*-[(2-{*N*-[4-(1-pyrenyl)butyl]-*N*-(4-methoxybenzyl)amino}ethoxy)-6-hydroxyphenyl]-5-bromocytidine (12):

11 (0.168 g, 0.261 mmol) was suspended in anhydrous DCM (5 mL) before triphenylphosphine (0.103 g, 0.393 mmol) and DEAD (0.062 mL, 0.394 mmol) were added. The resultant solution was stirred at room temperature for 0.5 h and to this **6** (0.131 g, 0.299 mmol, dissolved in 2 mL anhydrous DCM) was added dropwise. The reaction mixture was further stirred at room temperature for 2 h before water (1.0 mL) was added to quench the reaction. The solution was evaporated *in vacuo* to remove most of solvents and the residure was purified by silica gel column chromatography (26 % ethyl acetate in petroleum ether) to give a clear oil (0.225 g, 81 %). R_f = 0.49 (ethyl acetate/hexane, 2:1) ^{1}H NMR: (400 MHz, DMSO-*d*^{6}) δ ppm 9.63 (s, 1H, O<u>H</u>), 8.31-8.01 (m, 9H, C<u>H</u>-Ar, N<u>H</u>), 7.90 (s, 1H, C<u>H</u>$^{\mathbf{6}}$), 7.85 (d, *J* = 7.8 Hz, 1H, C<u>H</u>-Ar), 7.08 (d, *J* = 8.5 Hz, 2H, C<u>H</u>-Ar), 7.03 (t, *J* = 8.3 Hz, 1H, C<u>H</u>-Ar), 6.72 (d, *J* = 8.6 Hz, 2H, C<u>H</u>-Ar), 6.52 (d, *J* = 8.3 Hz, 1H, C<u>H</u>-Ar), 6.46 (d, *J* = 8.3 Hz, 1H, C<u>H</u>-Ar), 6.06 (t, *J* = 6.6 Hz, 1H, C<u>H</u>$^{\mathbf{1'}}$), 4.30-4.24 (m, 1H, C<u>H</u>$^{\mathbf{3'}}$), 4.00-3.95 (m, 2H, C<u>H</u>$_2$), 3.81-3.76 (m, 1H, C<u>H</u>$^{\mathbf{4'}}$), 3.72 (dd, *J* = 11.37, 3.17 Hz, 1H, C<u>H</u>H$^{\mathbf{5'}}$), 3.68-3.61 (m, 1H, C<u>H</u>H$^{\mathbf{5'}}$), 3.65 (s, 3H, C<u>H</u>$_3$), 3.47 (s, 2H, C<u>H</u>$_2$), 3.22 (t, *J* = 7.5 Hz, 2H, C<u>H</u>$_2$), 2.76-2.65 (m, 2H, C<u>H</u>$_2$), 2.52-2.44 (m, 2H, C<u>H</u>$_2$), 2.17-2.04 (m, 1H, C<u>H</u>H$^{\mathbf{2'}}$), 2.01-1.91 (m, 1H, CH<u>H</u>$^{\mathbf{2'}}$), 1.78-1.67 (m, 2H, C<u>H</u>$_2$), 1.57-1.48 (m, 2H, C<u>H</u>$_2$), 0.83 (s, 9H, C<u>H</u>$_3$, C<u>H</u>$_3$, C<u>H</u>$_3$), 0.81 (s, 9H, C<u>H</u>$_3$, C<u>H</u>$_3$, C<u>H</u>$_3$), 0.03 (d, *J* = 4.02 Hz, 6H, C<u>H</u>$_3$, C<u>H</u>$_3$), 0.00 (d, *J* = 2.80 Hz, 6H, C<u>H</u>$_3$, C<u>H</u>$_3$). ^{13}C NMR: (100 MHz, DMSO-*d*^{6}) δ ppm 158.9 (<u>C</u>-Ar), 158.0 (<u>C</u>-Ar), 156.5 (<u>C</u>-Ar), 155.0 (<u>C</u>-Ar), 153.7 (<u>C</u>-Ar), 153.0 (<u>C</u>-Ar), 140.5 (<u>C</u>H$^{\mathbf{6}}$), 137.0 (<u>C</u>-Ar), 131.4 (<u>C</u>-Ar), 130.9 (<u>C</u>-Ar), 130.4 (<u>C</u>-Ar), 129.7 (<u>C</u>H-Ar), 127.7 (<u>C</u>H-Ar), 127.4 (<u>C</u>H-Ar), 127.3 (<u>C</u>H-Ar), 127.0 (<u>C</u>H-Ar), 126.4 (<u>C</u>H-Ar), 126.0 (<u>C</u>H-Ar), 124.8 (<u>C</u>H-Ar), 124.7 (<u>C</u>H-Ar), 124.5 (<u>C</u>H-Ar), 123.5 (<u>C</u>H-Ar), 114.2 (<u>C</u>-Ar), 113.3 (<u>C</u>H-Ar), 109.3 (<u>C</u>H-Ar), 103.5 (<u>C</u>H-Ar), 87.0 (<u>C</u>H$^{\mathbf{4'}}$), 85.3 (<u>C</u>H$^{\mathbf{1'}}$), 71.9 (<u>C</u>H$^{\mathbf{3'}}$), 67.2 (<u>C</u>H$_2$), 62.5 (<u>C</u>H$_2^{\mathbf{5'}}$), 57.4 (<u>C</u>H$_2$), 54.8 (<u>C</u>H$_3$), 52.8 (<u>C</u>H$_2$), 51.5 (<u>C</u>H$_2$), 40.7 (<u>C</u>H$_2^{\mathbf{2'}}$), 32.4 (<u>C</u>H$_2$), 28.9 (<u>C</u>H$_2$), 26.6 (<u>C</u>H$_2$), 25.8 (<u>C</u>H$_3$), 25.6 (<u>C</u>H$_3$), 14.5 (<u>C</u>H$_3$). LRMS [ES$^+$,

MeCN]: 1085.6 ([M+Na]$^+$, 100 %). HRMS [ES$^+$]: $C_{57}H_{74}BrN_4O_7Si_2$ calculated 1063.4259 found 1063.4260.

3-(2′-deoxy-β-D-ribofuranosyl)-9-*N*-[(2-{*N*-[4-(1-pyrenyl)butyl]-*N*-(4-methoxybenzyl)amino}ethoxy)-1,3-diaza-2-oxophenoxazine (13):

12 (0.203 g, 0.191 mmol) was dissolved in anhydous ethanol (10 mL) and to the solution was added potassium fluoride (0.111 g, 1.91 mmol). The reaction mixture was stirred at 82 ˚C for 24 h under argon atmosphere. The solution was allowed to cool down to room temperature before being concentrated to dryness. The residue was then purified by silica gel column chromatography (6 % methanol in DCM) to give a pale yellow powder (0.078 g, 54 %). R_f = 0.34 (methanol/DCM, 1:10) ^{1}H NMR: (400 MHz, DMSO-*d*^{6}) δ ppm 8.37-8.05 (m, 9H, C<u>H</u>-Ar, C<u>H</u>^{4}), 7.92 (d, *J* = 7.8 Hz, 1H, C<u>H</u>-Ar), 7.57 (s, 1H, N<u>H</u>), 7.21 (d, *J* = 8.4 Hz, 2H, C<u>H</u>-Ar), 6.84 (d, *J* = 8.4 Hz, 2H, C<u>H</u>-Ar), 6.76 (t, *J* = 8.3 Hz, 1H, C<u>H</u>^{7}), 6.58 (d, *J* = 8.3 Hz, 1H, C<u>H</u>^{8}), 6.42 (d, *J* = 8.1 Hz, 1H, C<u>H</u>^{6}), 6.19 (t, *J* = 6.7 Hz, 1H, C<u>H</u>$^{1'}$), 5.25 (d, *J* = 3.8 Hz, 1H, 3'-OH), 5.12 (t, *J* = 4.4 Hz, 1H, 5'-OH), 4.31-4.23 (m, 1H, C<u>H</u>$^{3'}$), 4.03 (t, *J* = 5.6 Hz, 2H, C<u>H</u>$_2$), 3.87-3.81 (m, 1H, C<u>H</u>$^{4'}$), 3.72 (s, 3H, C<u>H</u>$_3$), 3.69-3.51 (m, 2H, C<u>H</u>$_2^{5'}$), 3.38 (s, 2H, C<u>H</u>$_2$), 3.32 (t, *J* = 7.52 Hz, 2H, C<u>H</u>$_2$), 2.85-2.74 (m, 2H, C<u>H</u>$_2$), 2.59 (t, *J* = 6.2 Hz, 2H, C<u>H</u>$_2$), 2.18-1.97 (m, 2H, C<u>H</u>$_2^{2'}$), 1.87-1.74 (m, 2H, C<u>H</u>$_2$), 1.70-1.55 (m, 2H, C<u>H</u>$_2$). ^{13}C NMR: (100 MHz, DMSO-*d*^{6}) δ ppm 158.1 (<u>C</u>-Ar), 137.0 (<u>C</u>-Ar), 131.2 (<u>C</u>-Ar), 130.9 (<u>C</u>-Ar), 130.4 (<u>C</u>-Ar), 129.8 (<u>C</u>H-Ar), 127.7 (<u>C</u>H-Ar), 127.4 (<u>C</u>H^{4}), 127.3 (<u>C</u>H-Ar), 127.0 (<u>C</u>H-Ar), 126.4 (<u>C</u>H-Ar), 126.1 (<u>C</u>H-Ar), 124.8 (<u>C</u>H-Ar), 124.7 (<u>C</u>H-Ar), 123.5 (<u>C</u>H-Ar), 123.2 (<u>C</u>H-Ar), 113.4 (<u>C</u>H-Ar), 113.3 (<u>C</u>H-Ar), 108.4 (<u>C</u>H-Ar), 107.6 (<u>C</u>H-Ar), 87.3 (<u>C</u>H$^{4'}$), 84.8 (<u>C</u>H$^{1'}$), 70.4 (<u>C</u>H$^{3'}$), 67.4 (<u>C</u>H$_2$), 61.3 (<u>C</u>H$_2^{5'}$), 57.6 (<u>C</u>H$_2$), 54.9 (<u>C</u>H$_3$), 53.3 (<u>C</u>H$_2$), 51.7 (<u>C</u>H$_2$), 40.0 (<u>C</u>H$_2^{2'}$), 32.4 (<u>C</u>H$_2$), 29.1 (<u>C</u>H$_2$), 26.4 (<u>C</u>H$_2$). LRMS [ES$^+$, MeCN]: 775.2 ([M+Na]$^+$, 100 %). HRMS [ES$^+$]: $C_{45}H_{45}N_4O_7$ calculated 753.3283 found 753.3277.

3-[2′-deoxy--3′,5′-*O*-di-(triethylsilyl)-β-D-ribofuranosyl]-9-*N*-[(2-{*N*-[4-(1-pyrenyl) butyl]-*N*-(4-methoxybenzyl)amino}ethoxy)-1,3-diaza-2-oxophenoxazine (14):

A solution of **13** (0.591 g, 0.785 mmol), chlorotriethylsilane (0.434 g, 2.59 mmol) and imidazole (0.353 g, 5.18 mmol) in DMF (15 mL) was stirred at room temperature for 24 h. The reaction mixture was diluted with DCM (120 mL) and washed with water (3 × 60 mL), saturated sodium bicarbonate (60 mL) and brine (60 mL). After the solution was dried over sodium sulfate and filtered, the filtrate was then evaporated *in vacuo* to remove solvents and the residue was purified by silica gel column chromatography (1.4 % methanol in DCM) to afford a yellow foam (0.661 g, 86%) R_f = 0.53 (methanol/DCM, 1:10) ^{1}H NMR: (400 MHz, DMSO-d^6) δ ppm 8.32-8.00 (m, 9H, C<u>H</u>-Ar, C<u>H</u>^{4}), 7.87 (d, *J* = 7.8 Hz, 1H, C<u>H</u>-Ar), 7.41 (s, 1H, N<u>H</u>), 7.16 (d, *J* = 8.5 Hz, 2H, C<u>H</u>-Ar), 6.79 (d, *J* = 8.6 Hz, 2H, C<u>H</u>-Ar), 6.74 (t, *J* = 8.3 Hz, 1H, C<u>H</u>^{7}), 6.55 (d, *J* = 8.3 Hz, 1H, C<u>H</u>^{8}), 6.29 (d, *J* = 8.1 Hz, 1H, C<u>H</u>^{6}), 6.11 (t, *J* = 6.3 Hz, 1H, C<u>H</u>$^{1'}$), 4.38-4.32 (m, 1H, C<u>H</u>$^{3'}$), 3.99 (t, *J* = 5.9 Hz, 2H, C<u>H$_2$</u>), 3.81-3.75 (m, 2H, C<u>H</u>$^{4'}$, C<u>H</u>H$^{5'}$), 3.71-3.65 (m, 1H, CH<u>H</u>$^{5'}$), 3.68 (s, 3H, C<u>H$_3$</u>), 3.53 (s, 2H, C<u>H$_2$</u>), 3.26 (t, *J* = 7.5 Hz, 2H, C<u>H$_2$</u>), 2.77 (t, *J* = 5.9 Hz, 2H, C<u>H$_2$</u>), 2.54 (d, *J* = 6.7 Hz, 2H, C<u>H$_2$</u>), 2.09 (t, *J* = 5.7 Hz, 2H, C<u>H$_2$</u>$^{2'}$), 1.82-1.72 (m, 2H, C<u>H$_2$</u>), 1.65-1.56 (m, 2H, C<u>H$_2$</u>), 0.97-0.86 (m, 18H, C<u>H$_3$</u>), 0.68-0.51 (m, 12H, C<u>H$_2$</u>). ^{13}C NMR: (100 MHz, DMSO-d^6) δ ppm 158.1 (<u>C</u>-Ar), 137.0 (<u>C</u>-Ar), 131.2 (<u>C</u>-Ar), 130.9 (<u>C</u>-Ar), 130.4 (<u>C</u>-Ar), 129.8 (<u>C</u>H-Ar), 127.3 (<u>C</u>H^{4}), 127.0 (<u>C</u>H-Ar), 126.3 (<u>C</u>H-Ar), 126.0 (<u>C</u>H-Ar), 124.8 (<u>C</u>H-Ar), 124.6 (<u>C</u>H-Ar), 123.4 (<u>C</u>H-Ar), 123.3 (<u>C</u>H-Ar), 113.4 (<u>C</u>H-Ar), 108.5 (<u>C</u>H-Ar), 107.4 (<u>C</u>H-Ar), 86.5 (<u>C</u>H$^{4'}$), 84.5 (<u>C</u>H$^{1'}$), 70.9 (<u>C</u>H$^{3'}$), 67.4 (<u>C</u>H$_2$), 61.7 (<u>C</u>H$_2$$^{5'}$), 57.6(<u>C</u>H$_2$), 54.9 (<u>C</u>H$_3$), 53.3 (<u>C</u>H$_2$), 51.7 (<u>C</u>H$_2$), 40.0 (<u>C</u>H$_2$$^{2'}$), 32.4 (<u>C</u>H$_2$), 29.0 (<u>C</u>H$_2$), 26.4 (<u>C</u>H$_2$). LRMS [ES$^+$, MeCN]: 981.2 ([M+H]$^+$, 60 %), 1003.4 ([M+Na]$^+$, 100 %). HRMS [ES$^+$]: $C_{57}H_{73}N_4O_7Si_2$ calculated 981.5012 found 981.5017.

3-(2′-deoxy-β-D-ribofuranosyl)-9-*N*-[(2-{*N*-[4-(1-pyrenyl)butyl]-*N*-(9-fluorenylmethoxycarbonyl)amino}ethoxy)-1,3-diaza-2-oxophenoxazine (15):

To a solution of **14** (0.661 g, 0.673 mmol) in DCM (20 mL) α-chloroethyl chloroformate (0.147 mL, 1.35 mmol) was added slowly at room temperature. The resultant solution was stirred at the same temperature for 1 h. After removal of the solvent, the residue was dissolved in anhydrous methanol (30 mL) and then refluxed for 2 h. The solution was cooled down to room temperature before being evaporated *in vacuo* to dryness. The residue was redissolved in anhydrous pyridine (23 mL), to which Fmoc-OSu (0.454 g, 1.35 mmol) was added. After the reaction mixture was stirred at room temperature for 0.5 h, it was diluted with DCM (150 mL) and washed with water (3 × 70 mL), citric acid solution (70 mL, 2 %), sodium bicarbonate (70 mL) and brine (70 mL). The organic layer was then dried over anhydrous sodium sulfate and filtered. The filtrate was evaporated *in vacuo* to remove solvent and the residue was purified by silica gel column chromatography (3.25 % methanol in DCM) to afford a yellow foam (0.274 g, 48 %). R_f = 0.35 (methanol/DCM, 1:10) ^{1}H NMR: (400 MHz, $CDCl_3$) δ ppm 8.22-7.89 (m, 9H, C<u>H</u>-Ar, C<u>H</u>^{4}), 7.74 (d, *J* = 7.6 Hz, 1H, C<u>H</u>-Ar), 7.65 (d, J = 7.4 Hz, 2H, C<u>H</u>-Ar), 7.56-7.43 (m, 3H, N<u>H</u>, C<u>H</u>-Ar, C<u>H</u>-Ar), 7.33-7.25 (m, 2H, C<u>H</u>-Ar), 7.33-7.25 (m, 2H, C<u>H</u>-Ar), 6.56-6.48 (m, 1H, C<u>H</u>^{7}), 6.26-5.90 (m, 3H, C<u>H</u>^{8}, C<u>H</u>^{6}, C<u>H</u>$^{1'}$), 4.65-4.51 (m, 4H, 3'-O<u>H</u>, C<u>H</u>$_2$, C<u>H</u>$^{3'}$), 4.26-4.06 (m, 2H, 5'-O<u>H</u>, C<u>H</u>), 4.04-3.98 (m, 1H, C<u>H</u>$^{4'}$), 3.92-3.82 (m, 2H, C<u>H</u>$_2$$^{5'}$), 3.80-3.72 (m, 2H C<u>H</u>$_2$), 3.48-3.33 (m, 2H, C<u>H</u>$_2$), 3.20-3.06 (m, 2H, C<u>H</u>$_2$), 3.03-2.90 (m, 2H, C<u>H</u>$_2$), 2.49-2.37 (m, 1H, C<u>H</u>H$^{2'}$), 2.29-2.18 (m, 1H, CH<u>H</u>$^{2'}$), 1.67-1.50 (m, 2H, C<u>H</u>$_2$), 1.43-1.32 (m, 2H, C<u>H</u>$_2$). ^{13}C NMR: (100 MHz, $CDCl_3$) δ ppm 143.9 (<u>C</u>-Ar), 142.6 (<u>C</u>-Ar), 141.3 (<u>C</u>-Ar), 136.2 (<u>C</u>-Ar), 131.4 (<u>C</u>-Ar), 130.8 (<u>C</u>-Ar), 129.8 (<u>C</u>-Ar), 128.5 (<u>C</u>-Ar), 127.6 (<u>C</u>H-Ar), 127.5 (<u>C</u>H-Ar), 127.2 (<u>C</u>H^{4}), 126.6 (<u>C</u>H-Ar), 127.0 (<u>C</u>H-Ar), 126.6 (<u>C</u>H-Ar), 125.8 (<u>C</u>H-Ar), 124.9 (<u>C</u>H-Ar), 124.8 (<u>C</u>H-Ar), 124.6 (<u>C</u>H-Ar), 123.3 (<u>C</u>H-Ar), 119.8 (<u>C</u>H-Ar), 108.4 (<u>C</u>H-Ar), 106.5 (<u>C</u>H-Ar), 87.3 (<u>C</u>H$^{4'}$), 86.9 (<u>C</u>H$^{1'}$), 70.2 (<u>C</u>H$^{3'}$), 66.7 (<u>C</u>H$_2$), 61.6 (<u>C</u>H$_2$$^{5'}$), 53.4 (<u>C</u>H$_2$), 48.6 (<u>C</u>H$_2$), 47.3 (<u>C</u>H), 47.2 (<u>C</u>H$_2$), 40.8 (<u>C</u>H$_2$$^{2'}$), 33.0 (<u>C</u>H$_2$), 28.5 (<u>C</u>H$_2$), 26.9 (<u>C</u>H$_2$). LRMS [ES$^+$, MeCN]: 855.5 ([M+H]$^+$, 80 %), 877.4 ([M+Na]$^+$, 100 %). HRMS[ES$^+$]: $C_{52}H_{46}N_4O_8Na$ calculated 877.3208 found 877.3197.

3-(5′-*O*-(4,4′-dimethoxytrityl)-2′-deoxy-β-D-ribofuranosyl)-9-*N*-[(2-{*N*-[4-(1-pyrenyl)butyl]-*N*-(9-fluorenylmethoxycarbonyl)amino}ethoxy)-1,3-diaza-2-oxophenoxazine (16):

DMTrCl (0.650 g, 1.92 mmol) in anhydrous pyridine (2 mL) was added dropwise to a solution of **15** (0.274 g, 0.320 mmol) in anhydrous pyridine (20 mL). The reaction mixture was stirred at room temperature for 6 h. Methanol (0.5 mL) was added to quench the reaction before it was diluted by DCM (150 mL). The organic layer was washed with water (3 × 70 mL), sodium bicarbonate (70 mL) and brine (70 mL). The solution was dried over anhydrous sulfate and filtered. The filtrate was evaporated *in vacuo* to remove solvents and the residue was purified by silica gel column chromatography (1.75 % methanol in DCM with 0.5 % pyridine) to give a pale yellow foam (0.274 g, 74 %). ^{1}H NMR: (400 MHz, $CDCl_3$) δ ppm 8.28-7.92 (m, 9H, C<u>H</u>-Ar, C<u>H</u>^{4}), 7.86-7.76 (m, 1H, C<u>H</u>-Ar), 7.74-7.64 (m 2H, C<u>H</u>-Ar), 7.60-7.45 (m, 5H, N<u>H</u>, C<u>H</u>-Ar, C<u>H</u>-Ar, C<u>H</u>-Ar, C<u>H</u>-Ar), 7.43-7.18 (m, 11H, C<u>H</u>-Ar), 6.90-6.82 (m, 4H, C<u>H</u>-Ar), 6.67 (t, *J* = 8.3 Hz, 1H, C<u>H</u>^{7}), 6.38-6.01 (m, 3H, C<u>H</u>^{8}, C<u>H</u>^{6}, C<u>H</u>$^{1'}$), 4.69-4.59 (m, 2H, C<u>H</u>$_2$), 4.59-4.53 (m, 1H, C<u>H</u>$^{3'}$), 4.25-4.12 (m, 2H, C<u>H</u>, C<u>H</u>$^{4'}$), 3.79-3.70 (m, 6H, C<u>H</u>$_3$), 3.52-3.27 (m, 6H, C<u>H</u>$_2^{5'}$, C<u>H</u>$_2$, C<u>H</u>$_2$), 3.27-3.00 (m, 4H, C<u>H</u>$_2$), 2.72-2.64 (m, 1H, C<u>H</u>H$^{2'}$), 2.29-2.19 (m, 1H, CH<u>H</u>$^{2'}$), 1.87-1.37 (m, 4H, C<u>H</u>$_2$). ^{13}C NMR: (100 MHz, $CDCl_3$) δ ppm 158.6 (<u>C</u>-Ar), 156.5 (<u>C</u>-Ar), 144.5 (<u>C</u>-Ar), 144.0 (<u>C</u>-Ar), 142.8 (<u>C</u>-Ar), 141.4 (<u>C</u>-Ar), 135.8 (<u>C</u>-Ar), 135.6 (<u>C</u>-Ar), 131.4 (<u>C</u>-Ar), 130.9 (<u>C</u>-Ar), 130.1 (<u>C</u>H-Ar), 130.0 (<u>C</u>H-Ar), 129.8 (<u>C</u>-Ar), 129.2 (<u>C</u>H-Ar), 128.6 (<u>C</u>-Ar), 128.1 (<u>C</u>H-Ar), 128.0 (<u>C</u>H-Ar), 127.9 (<u>C</u>H-Ar), 127.6 (<u>C</u>H-Ar), 127.5 (<u>C</u>H-Ar), 127.3 (<u>C</u>H-Ar), 127.1 (<u>C</u>H^{4}), 126.9 (<u>C</u>H-Ar), 126.6 (<u>C</u>H-Ar), 125.8 (<u>C</u>H-Ar), 125.0 (<u>C</u>H-Ar), 124.8 (<u>C</u>H-Ar), 124.7 (<u>C</u>H-Ar), 123.3 (<u>C</u>H-Ar), 119.9 (<u>C</u>H-Ar), 113.3 (<u>C</u>H-Ar), 113.2(<u>C</u>H-Ar), 108.4 (<u>C</u>H-Ar), 106.4 (<u>C</u>H-Ar), 86.5 (<u>C</u>H$^{4'}$), 86.3 (<u>C</u>H$^{1'}$), 71.9 (<u>C</u>H$^{3'}$), 66.9 (<u>C</u>H$_2$), 63.5 (<u>C</u>H$_2^{5'}$), 55.2 (<u>C</u>H$_3$), 53.5 (<u>C</u>H$_2$), 48.6 (<u>C</u>H$_2$), 47.4 (<u>C</u>H), 47.2 (<u>C</u>H$_2$), 41.8 (<u>C</u>H$_2^{2'}$), 33.1 (<u>C</u>H$_2$), 28.6 (<u>C</u>H$_2$), 26.9 (<u>C</u>H$_2$). LRMS [ES$^+$, MeCN]: 1179.5 ([M+Na]$^+$, 100 %).

3-{3′-*O*-[2-cyanoethyl(*N*,*N*-diisopropyl-amino)phosphanyl]-5′-*O*-(4,4′-dimethoxytrityl)-2′-deoxy-β-D-ribofuranosyl}-9-*N*-[(2-{*N*-[4-(1-pyrenyl)butyl]-*N*-(9-fluorenylmethoxycarbonyl)amino}ethoxy)-1,3-diaza-2-oxophenoxazine (17): *N*,*N*-Diisopropylethylamine (0.082 mL, 0.459 mmol) was added to a solution of **16** (0.274 g, 0.236 mmol) in anhydrous DCM (10 mL), followed by 2-cyanoethyl-*N*,*N*-diisopropyl chlorophosphine (0.074 mL, 0.331 mmol). The mixture was stirred under argon for 0.5 h at room temperature before saturated potassium chloride (10 mL) and anhydrous DCM (25 mL) were added. The organic layer was separated under an argon atmosphere and the solvent was removed *in vacuo*. The residue was purified under an argon atmosphere by silica gel column chromatography (50 % ethyl acetate in DCM with 0.5 % pyridine) to yield the diastereomeric product as a white foam (0.254 g, 79 %). R_f = 0.69, 0.54 (ethyl acetate/DCM, 1:1). ^{31}P NMR (121 MHz, DMSO-d^6) δ ppm 148.9, 148.5. *m/z* (%) LRMS [ES$^+$, MeCN]: 1380 ([M+Na]$^+$, 100 %).

S4 Oligonucleotide synthesis, purification and analysis

Oligonucleotide synthesis was carried out on an Applied Biosystems 394 automated DNA/RNA synthesizer using a standard 0.2 μmol or 1.0 μmol phosphoramidite cycle of acid-catalyzed detritylation, coupling, capping and iodine oxidation. All X-pyrene and G-clamp phosphoramidite monomers were dissolved in anhydrous DCM or acetonitrile to a concentration of 0.1 M immediately prior to use. The coupling time for normal A, G, C, and T monomers was 35 s and this was extended to 6 min for all modified monomers. Stepwise coupling efficiencies and overall yields were determined by automated trityl cation conductivity monitoring and in all cases were >98.0 %. The oligonucleotides attached to the synthesis columns were treated with 20 % diethylamine in acetonitrile for 20 min then washed with acetonitrile (5 x 1 mL). This procedure removes cyanoethyl groups from the phosphotriesters and scavenges the

resultant acrylonitrile, preventing cyanoethyl adducts being formed at the secondary amines of X-pyrene. After this, cleavage of oligonucleotides from the solid support and deprotection were achieved by exposure to concentrated aqueous ammonia for 5 h at 55 ˚C.

Purification of oligonucleotides was carried out by reversed-phase HPLC on a Gilson system using a Brownlee Aquapore column (C8, 8 mm x 250 mm, 300 Å pore) with a gradient of CH_3CN in NH_4OAc increasing from 0 % (buffer A) to 50 % (buffer B) over 20 min with a flow rate of 4 mL/min (buffer A: 0.1 M NH_4OAc, pH 7.0; buffer B: 0.1 M NH_4OAc with 50 % CH_3CN, pH 7.0). Elution of oligonucleotides was monitored by ultraviolet absorption at 298 nm. After HPLC purification, oligonucleotides were desalted using NAP-10 Sephadex columns (GE Healthcare) according to the manufacturer's instructions.

Mass spectra of oligonucleotides were recorded on a Bruker micrOTOF™ II focus ESI-TOF MS instrument in ES^- mode. Capillary electrophoresis (CE) analysis was conducted on a Beckman Coulter P/ACE™ MDQ Capillary Electrophoresis System, using the 32 Karat Software MDQ UV application, at a concentration of ~4 OD/mL. An ssDNA 100-R gel with Tris-Borate-7 M Urea system was used (representative CE traces in Figure S2). Analytical HPLC traces were recorded on Bruker micrOTOF™ II focus ESI-TOF MS instrument (representative HPLC traces in Figure S2).

S5. UV spectra of oligonucleotides:

The representative UV absorption curves of synthetic oligonucleotide containing X-pyrene (ODN-1) and its Watson Crick complementary strand (bulged duplex formation, with C on the unpaired position) are shown in Figure S8. 260 nm was used in the duplex UV-melting studies.

UV Scans (200-475 nm) were carried out on three duplex samples (BD1, Y = X-pyrene; ND1, Y = G-clamp or C; Figure S9).

S6. Ultraviolet duplex melting studies:

The representative UV absorption curves of synthetic oligonucleotides are shown in Figure S1, Figure S4, Figure S6, Figure S10, Figure S11, Figure S12, Figure S13 and Figure S14. All the duplex melting curves shown in this paper were recorded at 260 nm. The extinction coefficient of the X-pyrene at 260 nm is 29300 M^{-1} cm^{-1}.[1-2] This was used to calculate oligonucleotide concentrations.

To determine duplex melting temperatures (T_m), UV melting studies were carried out on a Varian Cary 400 scan UV-visible spectrophotometer using Hellma SUPRASIL synthetic quartz 10 mm path length cuvettes, monitoring at 260 nm with a complimentary DNA strand concentration of 2.5 μM and a volume of 1.2 mL. Samples were prepared as follows: The modified sequences and their corresponding complementary strand were mixed in a 1:1 ratio in 2 mL Eppendorf tubes then lyophilized before dissolved in 1.2 mL of the buffer solution (5 mM sodium phosphate, 1 mM $MgCl_2$, pH 7.2 containing 140 mM KCl). The samples were then filtered into the cuvettes with Kinesis regenerated cellulose 13 mm, 0.45 μm syringe filters. The UV melting protocol involved initial denaturation by heating to 84 ˚C at 10 ˚C/min followed by annealing by cooling to 18 ˚C at 1.0 ˚C/min, then maintaining at 18 ˚C for 2 min before starting the melting experiment which involved heating from 18 ˚C to 84 ˚C at 1.0 ˚C/min, holding at 84 ˚C for 2 min then cooling to 18 ˚C at 1.0 ˚C/min. Three successive melting curves were measured before fast annealing from 84 ˚C to 20 ˚C at 10 ˚C/min. T_m values were calculated using Cary Win UV thermal application software, taking an average of the three melting curves.

S7. Fluorescence studies:

The experiments were carried out using a BMG Labtech CLARIOstar plate reader with a Greiner 96 wells fluorescence microplate. In order to maximise the emission intensities the excitation wavelength used were 344 nm for X-pyrene and 364 nm for G-clamp, while the detection was at 372-612 nm for X-pyrene and 390-612 nm for G-clamp. All samples were prepared with a concentration of 0.15 μM of probe strand (containing either X-pyrene or G-clamp) and 0.2 μM of complementary strand in a phosphate buffer (5 mM sodium phosphate, 1 mM $MgCl_2$, pH 7.2 containing 140 mM KCl). 240 μL of solution per well were used and each sample was divided in 4 wells. Before the analysis, the samples were denatured by heating until 95 °C and then they were left to slowly cool down to room temperature overnight, in order to allow a slow annealing. All the spectra reported are an average of four spectra from which the average blank spectrum was subtracted and they are normalized to the ssDNA spectrum present in each graph.

Titrations reported in Figure S15 B were done by adding increasing quantities (from 1 μL to 4 μL in 0.5 μL steps) of the complementary oligonucleotide to ODN-1 containing a bulged G to a solution of the duplex BD1-3′G (with the following initial concentrations: 0.15 μM of probe strand containing X-pyrene and 0.2 μM of complementary strand in a phosphate buffer with 5 mM sodium phosphate, 140 mM KCl, 1 mM $MgCl_2$, pH 7.2). For the blank and the solution of ODN-1 alone, buffer was added each time to obtain the same volume as the BD1-3′G solution.

S8. Circular dichroism:

CD spectra were measured on a Jasco J-815 spectropolarimeter at 2.5 μM concentration of duplexes in phosphate buffer (5 mM sodium phosphate, 1 mM $MgCl_2$, pH 7.2 containing 140 mM KCl).[3] Spectra were recorded at 50 nm min^{-1}

(accumulation = 5) with a response time of 1 s, a bandwidth of 1 nm and a data pitch of 0.2 nm. A buffer baseline was subtracted from each spectrum which started at 320 nm.

CD spectra were obtained from two bulge duplexes (where Y = X-pyrene or G-clamp) and three normal ones containing X-pyrene, G-clamp and C (Figure S19). Despite the negative phase being more prominent and moved slightly towards lower wavelength (~15 nm), the incorporation of X-pyrene has no significant effect on DNA conformation in BD1 and ND1; the CD spectra of both bulged and normal duplexes containing X-pyrene are consistent with the B-conformation as in the standard duplex (Y = C, ND1). Interestingly, G-clamp modification resulted in a weak pseudo-positive Cotton effect between 233 nm and 252 nm in both duplex contexts compared to the other duplexes (where Y = X-pyrene or C). This is very similar to what was observed on tC^{O}.[4] Overall, the CD spectra indicated that the B-conformation is maintained in bulged and normal duplexes when G-clamp is introduced.

S9. Error calculations:

Melting temperature errors were calculated using the confidence intervals derivable with Student's t distribution with 95 % confidence and the appropriate degrees of freedom. For ΔT_m the errors were calculated using the normal error propagation formula without the covariance factor. In graphs errors are represented by error bars.

S10. Azide labeling:

NHS ester of 6-azidohexanoate (1 mg) in DMF (80 μL) was added post-synthetically to the lyophilized X-pyrene modified oligonucleotide (0.1 μmol, ODN-1) in Na_2CO_3/$NaHCO_3$ buffer (80 μL, 0.5 M, pH 8.75) for 4 h at 55 ˚C.[5] The fully labeled oligonucleotides were desalted on NAP-25 Sephadex columns (GE Healthcare) and purified by reverse-phase HPLC (ODN-12, MS, requires: 3517, found 3517).

When the secondary amine on the X-pyrene side-chain was functionalized by acylation with 6-azidohexanoic acid NHS ester after oligonucleotide synthesis, a large decrease in T_m (−16.4 ˚C) was observed in the stability of the BD1 bulged duplex (Table S1 and Figure S1). This is probably because conversion of the secondary amine to an amide prevents X-pyrene forming the extra hydrogen bond in the major groove with the complementary guanine base (Figure 1B), suggesting it plays a pivotal role in duplex stabilization.

S11. NMR spectroscopy

Two samples of 1.2 mM X-pyrene DNA duplex with T or C as the widowed base respectively, were dissolved in 90% H_2O/10% ^2H_2O or 100% ^2H_2O respectively. Final sample volume for both samples was 250 µl (Shigemi). Samples were heated to 95 ˚C for 5 minutes and then cooled slowly to room temperature to promote dimer formation.

Data collection

Identical sets of spectra were acquired for both samples at 278 K, 298 K and 310 K in H_2O and ^2H_2O, on a Bruker Avance III 600 MHz spectrometer equipped with a room temperature probe and cryogenic probe, respectively. Spectra were processed with Topspin and NMRPipe[6] and analyzed using CARA[7]. Acquisition parameters as described in Table S6 were used. For the NOESY experiments in ^2H_2O and H_2O, two spectra each were acquired with 100 ms and 200 ms mixing time respectively. We acquired one TOCSY spectrum with 40 ms mixing times respectively. The data were zero filled to 512 × 2K complex data points, followed by apodization using Lorentz-to-Gauss transformation and cosine functions in t_2 and t_1, respectively, before Fourier transformation.

S12. Modelling

The models, based on the NMR data, were prepared using Hyperchem 7 (Hypercube, Inc.) and Avogadro (an open-source molecular builder and visualization tool version 1.1.1 http://avogadro.openmolecules.net/). Geometrical optimization was done in Hyperchem using the Polak-Ribiere algorithm, RMS gradient of 0.01 Kcal/(Å·mol). All the models were rendered with PyMOL (Molecular Graphics System, Version 1.5.0.3 Schrödinger, LLC).

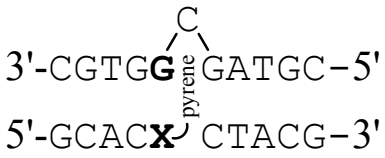


Table S1. UV-melting analysis on the bulged duplex (azide labeling)[a]

Azide labelling	Before	After
T_m /°C	66.2	49.8 (-16.4)

[a] T_m values are an average of three melting temperatures (˚C). The value in parenthesis is the difference between the T_m before azide labeling and after.

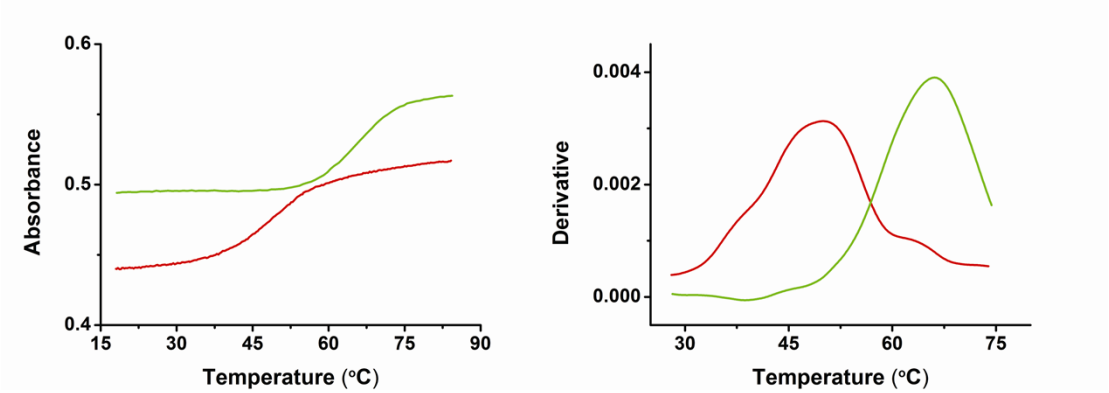


Figure S1. Representative UV melting curves (left) and derivatives (right) of the bulged duplex containing X-pyrene (green) and the one with 6-azidohexanoic modification (red). The experiments were recorded on 260 nm in 5 mM NaH_2PO_4/Na_2HPO_4 buffer (pH 7.2, containing 140 mM KCl. 1mM $MgCl_2$, The concentration of modified oligo: Watson-Crick complementary strand = 2.5 μM:2.5 μM.

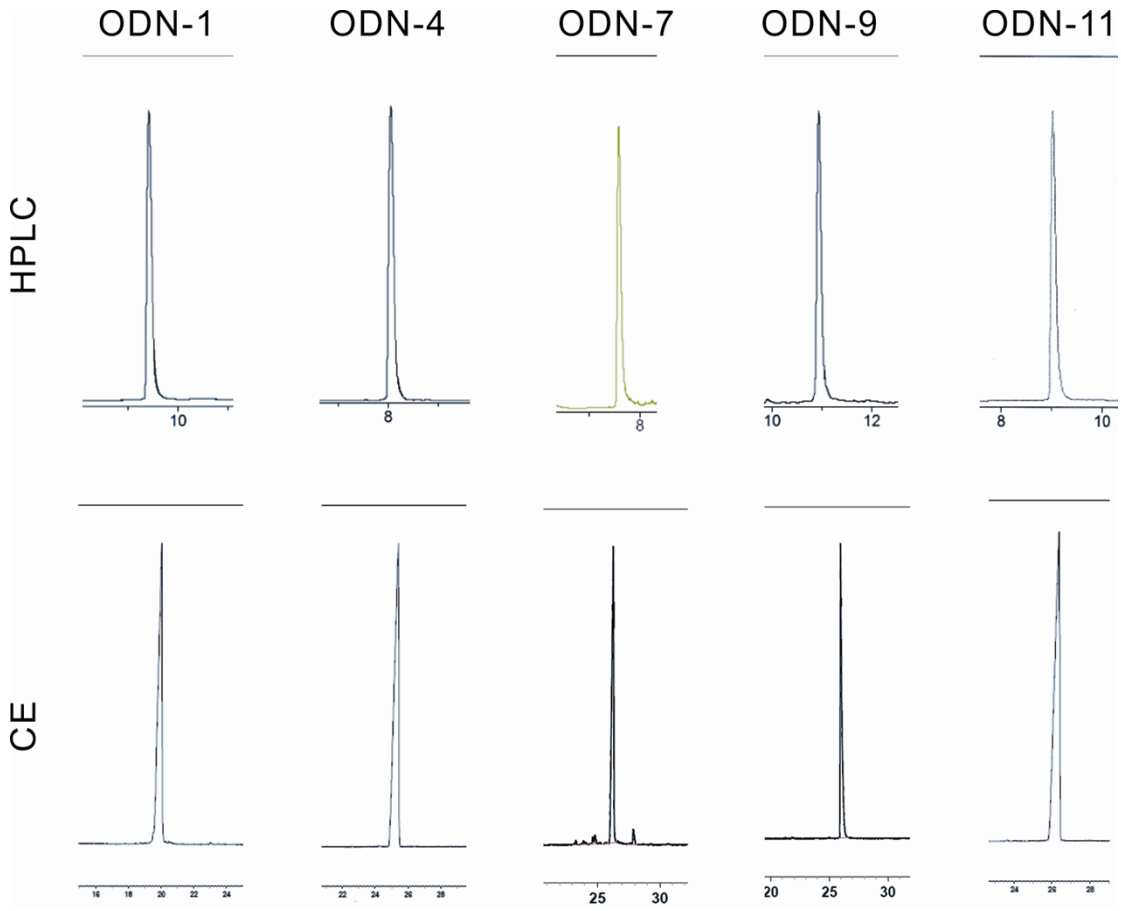


Figure S2 Representative analytic HPLC and CE traces of synthesized oligonucleotides.

		5'-intercalation		3'-intercalation	
CYC form	**i**	3'-CGTG<u>GGATGC</u>-5' 5'-GCAC<u>XCTACG</u>-3'	T_m = 55.7 °C	3'-CGTG<u>GGATGC</u>-5' 5'-GCAC<u>XCTACG</u>-3'	T_m = 61.7 °C
	ii	3'-CGTG<u>GGATGC</u>-5' 5'-GCAC<u>XCTACG</u>-3'	T_m = 58.1 °C	3'-CGTG<u>GGATGC</u>-5' 5'-GCAC<u>XCTACG</u>-3'	T_m = 66.2 °C
AYA form	**iii**	3'-CGTT<u>AGTATGC</u>-5' 5'-GCAA<u>XATACG</u>-3'	T_m = 37.2 °C	3'-CGTTG<u>ATATGC</u>-5' 5'-GCAA<u>XATACG</u>-3'	T_m = 49.0 °C

Figure S3. Melting temperatures of 5′- and 3′- X-pyrene bulged duplexes in two different neighboring environments (CYC and AYA, where Y = X-pyrene). ODN-1 (CYC form) and ODN-11 (5′-GCAAYATACG, AYA form) were used to construct X-pyrene bulged duplexes. Unpaired positions are in cyan and intercalating pyrene moieties in red. Different flanking neighborhoods are underlined in green.

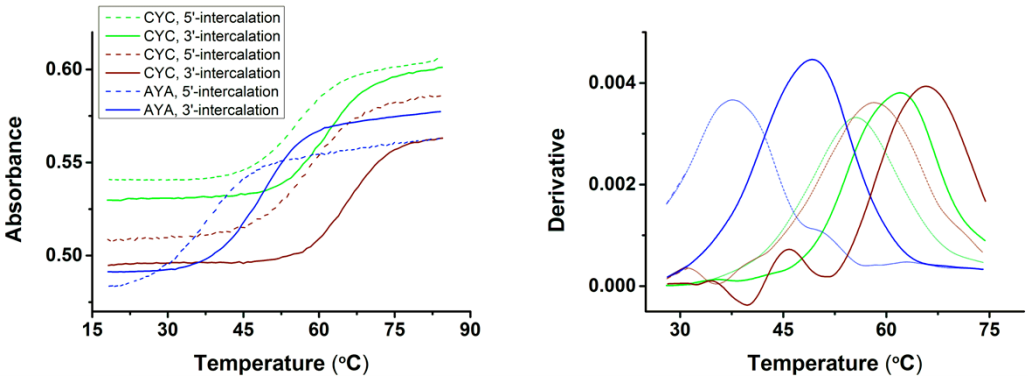


Figure S4. UV melting curves (left) and derivatives (right) of 5′- and 3′- X-pyrene bulged duplexes in two different neighboring environments (Figure S3). The experiments were carried out at pH 7.2 under the same condition as in Figure S1 (CYC form: **i**, green; **ii**, wine; AYA form, **iii**, blue; dash line, 5′-intercalation; solid line, 3′-intercalation).

	5'-intercalation		3'-intercalation	
DNA/RNA hybrid	3'-CGUG^{pyrene}GGAUGC–5' 5'-GCAC<u>XCTACG</u>–3'	T_m = 57.6 °C	3'-CGUGG^{pyrene}GAUGC–5' 5'-GCAC<u>X</u>CTACG–3'	T_m = 60.1 °C

Figure S5. Melting temperatures of 5′- and 3′- X-pyrene bulged DNA/RNA hybrid duplexes. Unpaired positions are in cyan and intercalating pyrene moieties in red. The flanking neighborhood is underlined in green.

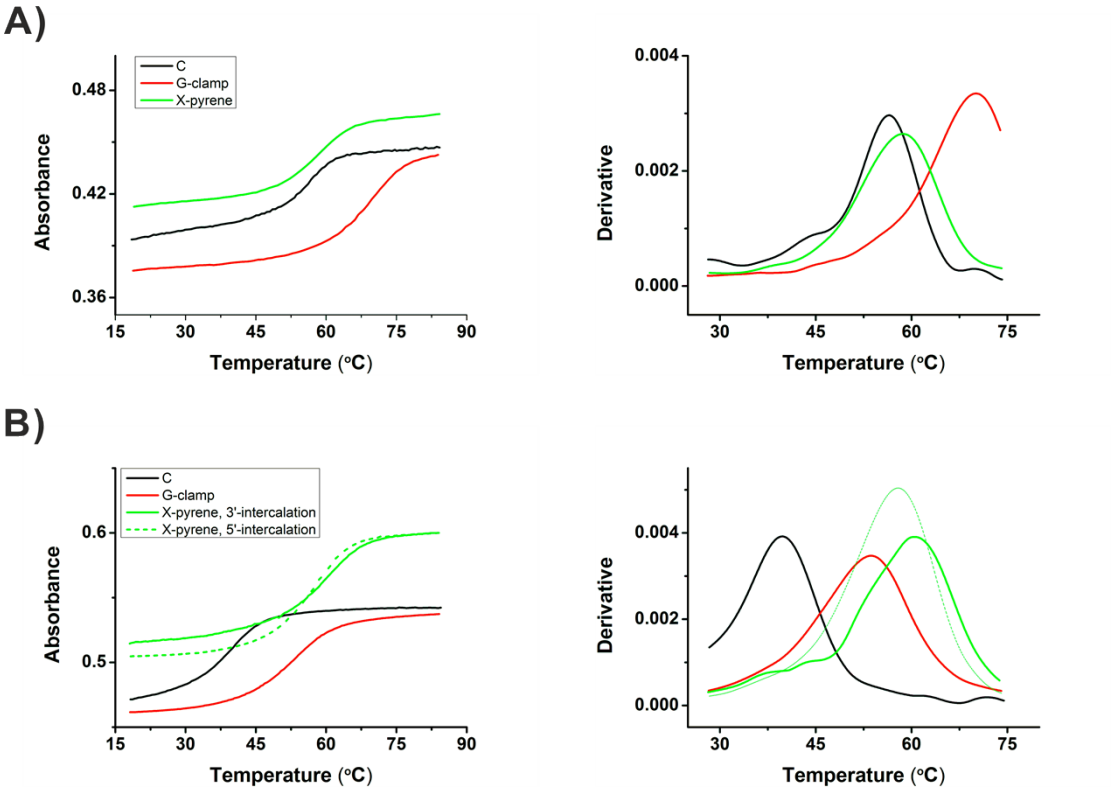


Figure S6. UV melting curves (left) and derivatives (right) of DNA/RNA hybrids in Figure 2 and Figure S5 (Y = C, black; Y = G-clamp, red; Y = X-pyrene, green). The experiments were carried out at pH 7.2 under the same condition as in Figure S1: A) ND5; B) BD5 (solid line) and X-pyrene 5′-bulged duplex (dash line).

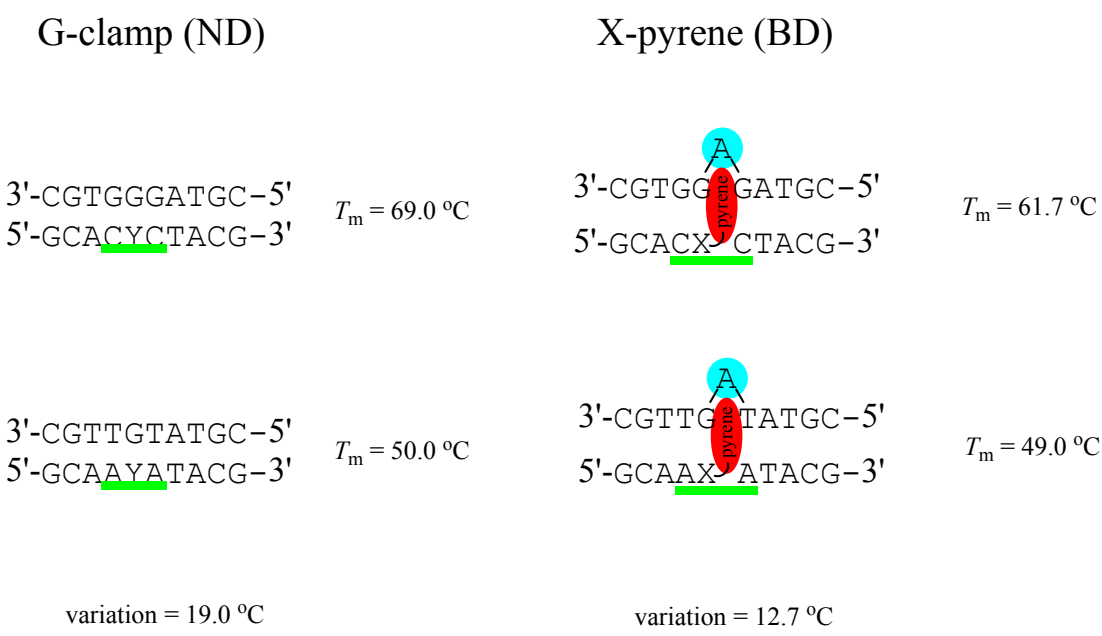


Figure S7. Comparison of melting temperatures of 3′- X-pyrene bulged duplexes (BD) to G-clamp normal standard duplexes (ND) in two extreme neighboring environments (CYC and AYA, where Y = X-pyrene or G-clamp). The T_m values of duplexes containing G-clamp are quoted from Ortega's work and used for calculating variation.[8] ODN-1 (CYC form) and ODN-11 (AYA form) were used to construct X-pyrene bulged duplexes and the melting study was performed under the same condition as theirs (pH 7.2, containing 140 mM KCl. 1 mM $MgCl_2$, the concentration of modified oligo: Watson-Crick complementary strand = 2.5 μM:2.5 μM.). T_m values of bulged duplexes can be referred to Figure S3. Unpaired positions are in cyan and intercalating pyrene moieties in red. Different flanking neighborhoods are underlined in green.

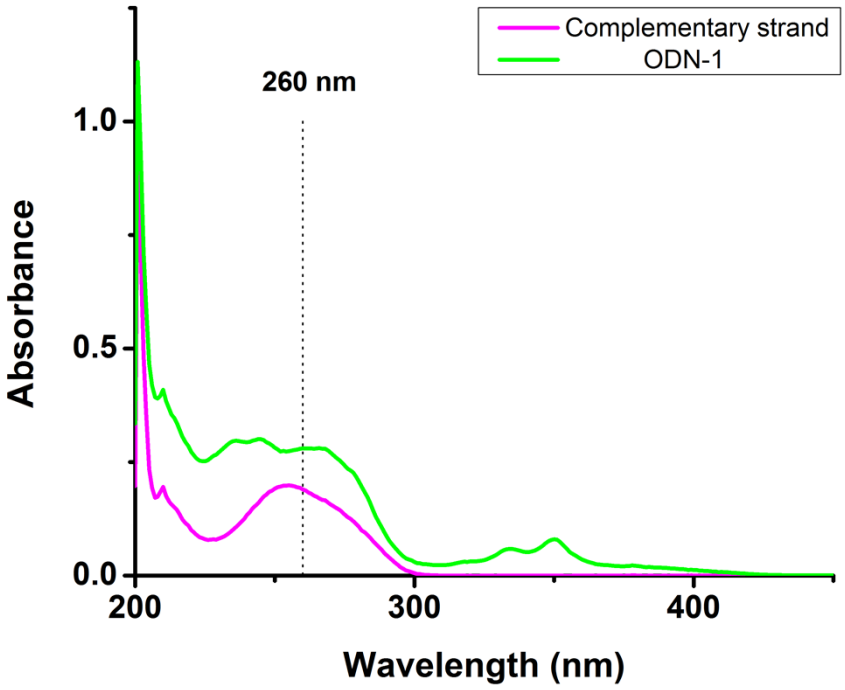


Figure S8: Representative UV absorption curves of synthesized oligonucleotides (ODN-1 and its complementary strand 5′-CGTAGCGGTGC) within the range of 200 - 450 nm.

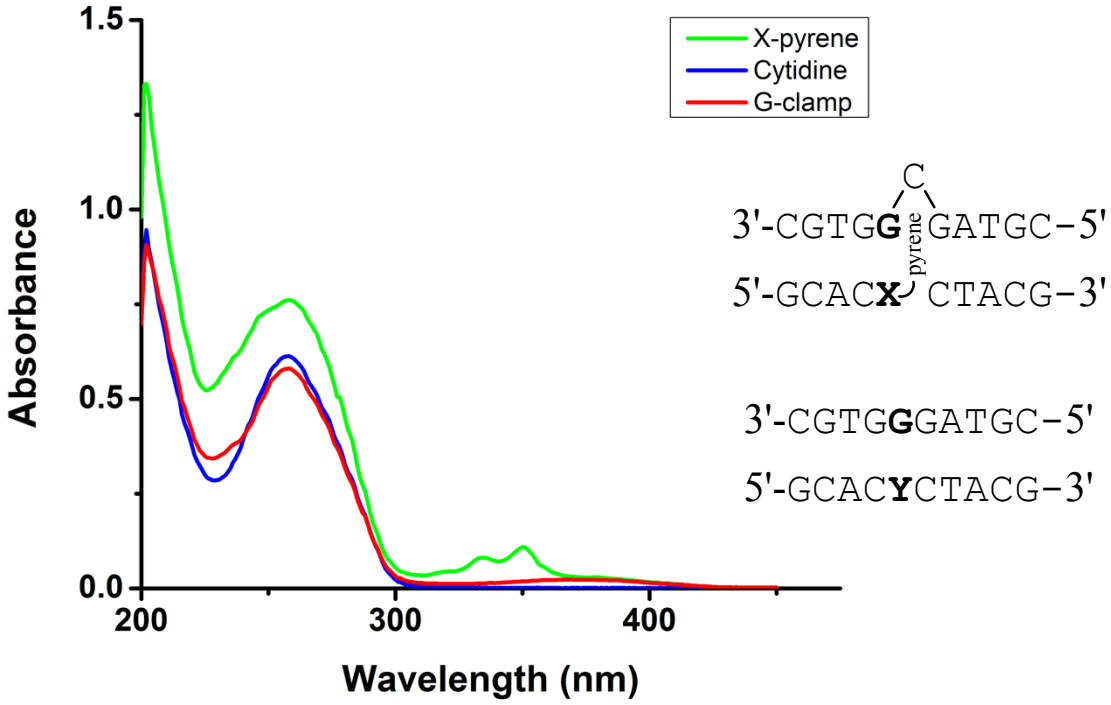


Figure S9: Representative UV absorption curves of ND1 (Y = G-clamp or C, 2.5 μM) and X-pyrene bulged duplex (BD1, 2.5 μM) within the range of 200 - 475 nm: BD1 is in green, ND1 containing G-clamp in red and ND1 containing C in blue.

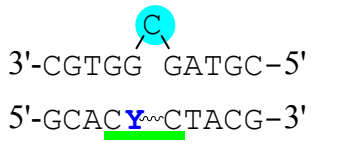
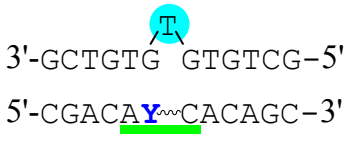
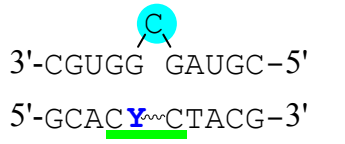
Target	Entry	ND	BD
DNA	1	3'-CGTGGGATGC-5' 5'-GCACYCTACG-3'	
	2	3'-GCTGTGGTGTCG-5' 5'-CGACAYCACAGC-3'	
RNA	5	3'-CGUGGGAUGC-5' 5'-GCACYCTACG-3'	

Table S2. UV-melting analysis on above duplexes in Entry 1, 2 and 5[a]

Target strand	**Y**	X-pyrene	G-clamp	C
ssDNA	**ND1**	60.8 (+11.3)	65.4 (+15.9)	49.5
	ND2	65.1 (+7.9)	65.3 (+8.1)	57.2
	BD1	66.2 (+32.2)	51.8 (+17.8)	34.0
	BD2	64.1 (+15.3)	55.2 (+6.4)	48.8
ssRNA	**ND5**	58.4 (+2.1)	69.9 (+13.6)	56.3
	BD5	60.1 (+20.4)	53.6 (+13.9)	39.7

[a] T_m values are an average of three melting temperatures (˚C). Values in parentheses are the differences between the T_m of duplexes containing X-pyrene or G-clamp and the corresponding T_m on C.

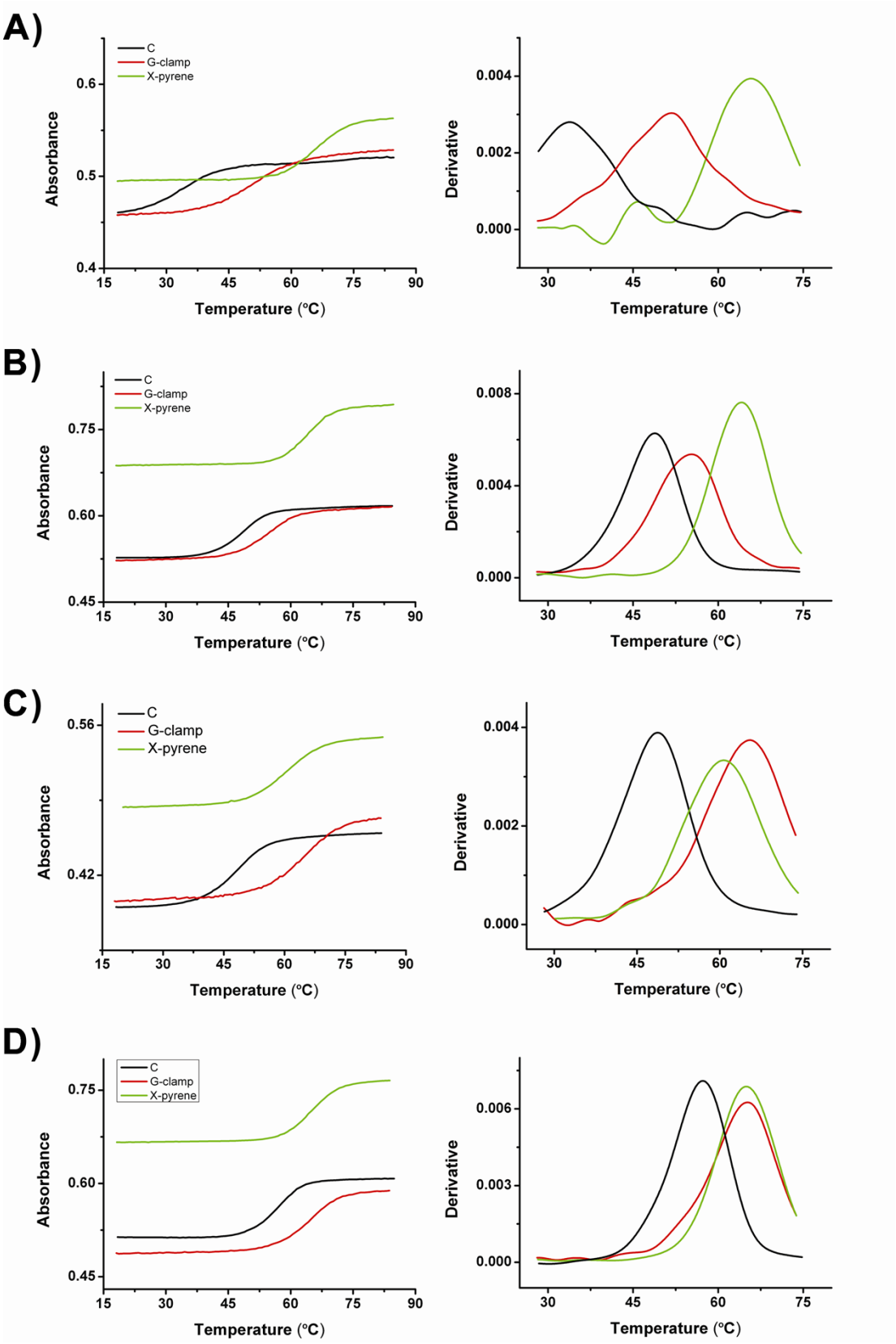


Figure S10. Representative UV melting curves (left) and derivatives (right) of BD1-2 and ND1-2 above Table S2 (Y = C, black; Y = G-clamp, red; Y = X-pyrene, green).The experiments were carried out at pH 7.2 under the same condition as in Figure S1: A) BD1; B) BD2; C) ND1; D) ND2.

Table S3 UV-melting analysis on duplexes in Entry 3-4 in Figure 2[a]

	ND (C)	**BD (X-pyrene)**	**ΔT_m**
3 (clustered)	50.6 (±0.4)	70.4 (±0.4)	+19.8 (±0.5)
4 (dispersed)	44.9 (±0.2)	74.7 (±0.2)	+29.8 (±0.3)

[a] T_m values are an average of three melting temperatures (˚C). The experiments were carried out at pH 7.2 under the same condition as in Figure S1.

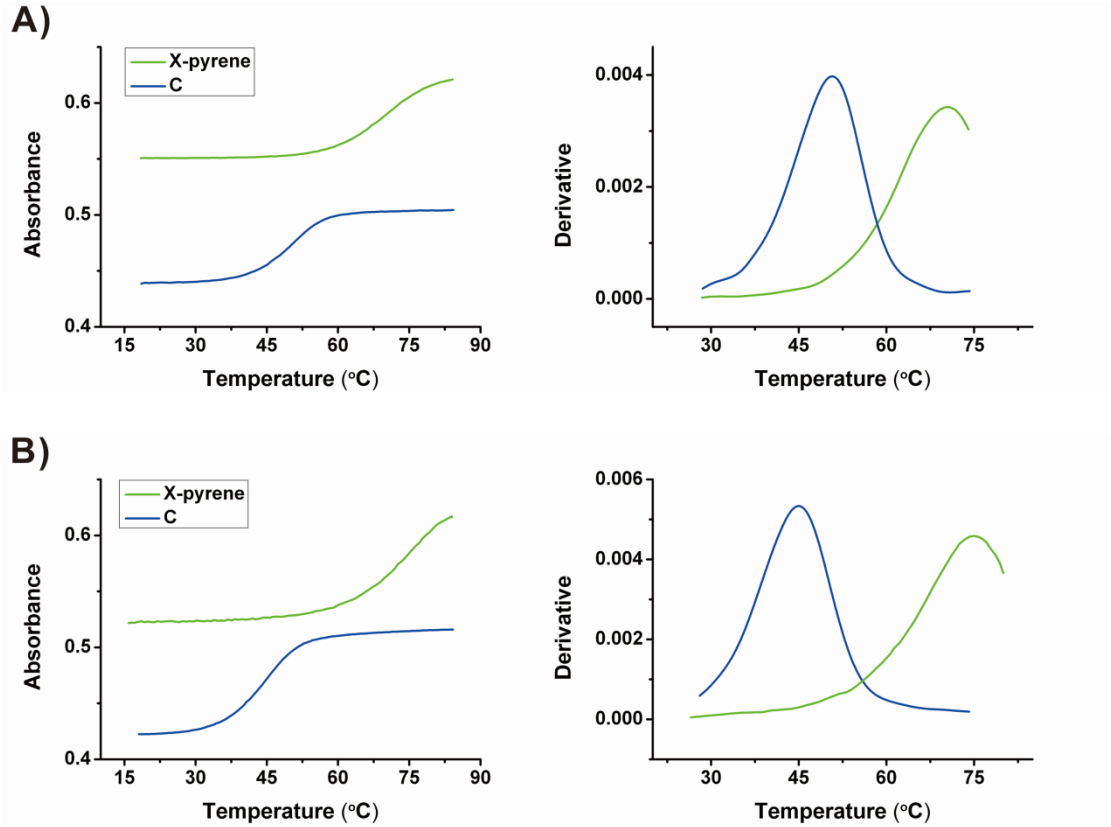


Figure S11. UV melting curves (left) and derivatives (right) of duplexes (entries 3 and 4, Figure 2) A) clustered--ND3 (blue) and BD3 (green); B) dispersed--ND4 (blue) and BD4 (green). The experiments were carried out at pH 7.2 under the same condition as in Figure S1.

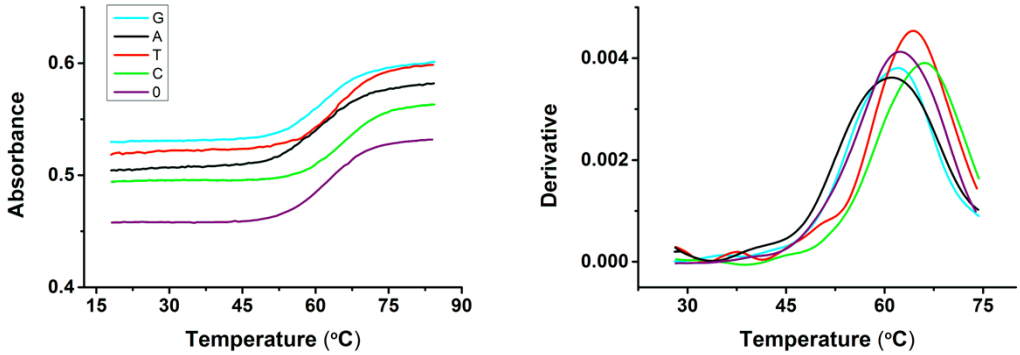


Figure S12. UV melting curves (left) and derivatives (right) of the above bulged duplex (Z = G, cyan; Z = A, black; Z = T, red; Z = C, green; Z = 0, tetrahydrofuran, purple). The experiments were carried out at pH 7.2 under the same condition as in Figure S1.

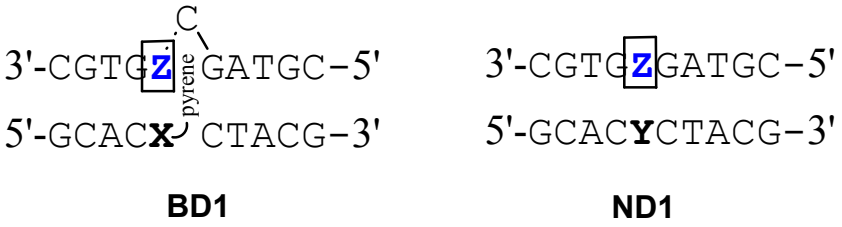


Table S4. UV-melting analysis on above duplexes

Z	**G**	**A**	**C**	**T**
X-pyrene	66.2	31.5	30.4	30.4
G-clamp	65.4	38.2	27.1	32.0
C	49.5	28.8	21.6	27.2

[a] T_m values are an average of three melting temperatures (˚C).

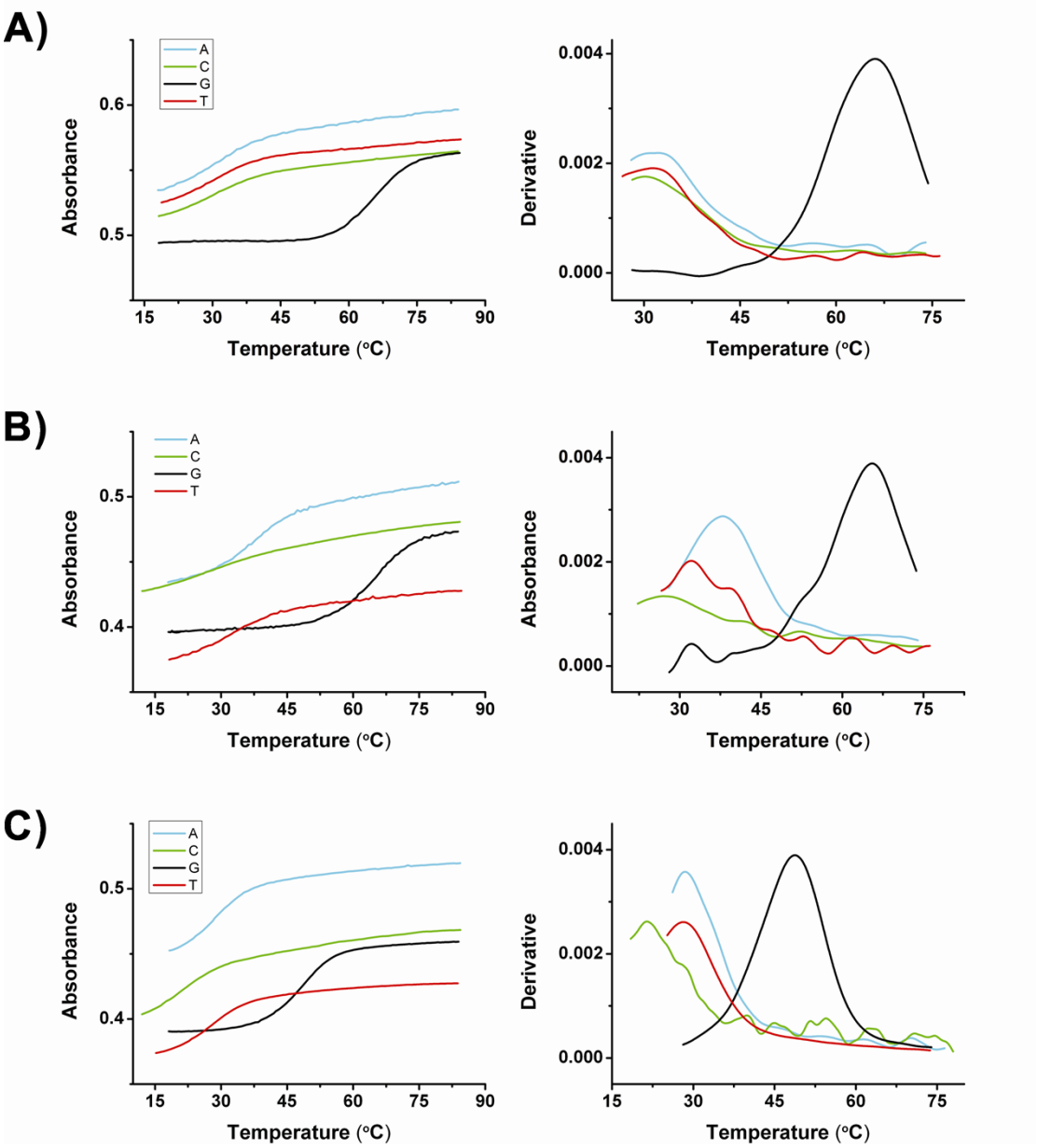


Figure S13. Representative UV melting curves (left) and derivatives (right) of the bulged duplex and normal duplexes above Table S4 (Z = A, cyan; Z = G, black; Z = T, red; Z = C, green). The experiments were carried out at pH 7.2 under the same condition as in Figure S1: A) BD1, containing X-pyrene; B) ND1, Y = G-clamp; C) ND1, Y = C.

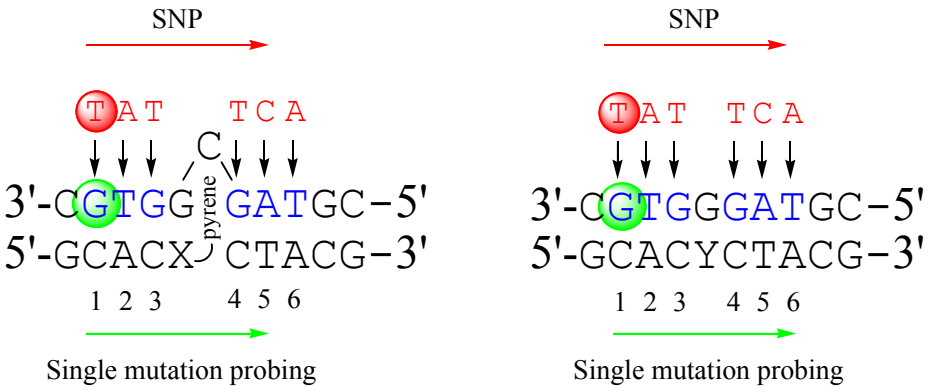


Table S5. UV-melting analysis on above duplexes[a]

Mutating position	**X-pyrene**	**G-clamp**	**C**
1 (TC)	55.7 (-10.5)	54.7 (-10.7)	31.1 (-18.4)
2 (AA)	27.7 (-38.5)	50.1 (-15.3)	32.4 (-17.1)
3 (TC)	29.8 (-36.4)	36.2 (-29.2)	29.0 (-20.5)
4 (TC)	46.8 (-19.4)	39.1 (-26.3)	31.0 (-18.5)
5 (CT)	49.5 (-16.7)	49.8 (-15.6)	36.3 (-13.2)
6 (AA)	53.4 (-12.8)	55.9 (-9.5)	36.2 (-13.3)

[a] T_m values are an average of three melting temperatures (°C). Values in parentheses are the difference between the T_m of fully paired or bulged duplexes and the corresponding T_m with single mismatch introduced. The experiments were carried out at pH 7.2 under the same condition as in Figure S1.

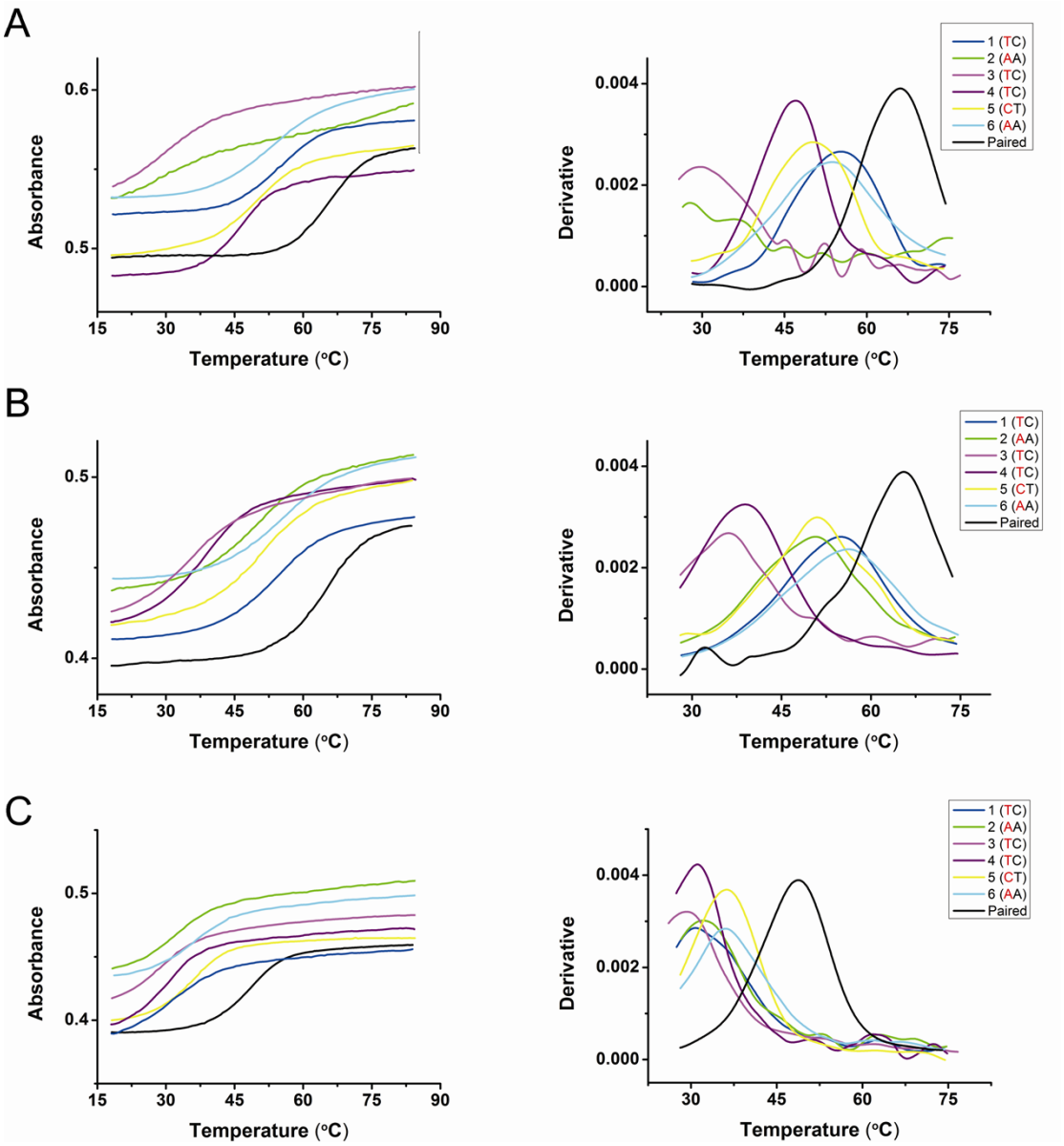


Figure S14. Representative UV melting curves (left) and derivatives (right) of bulged duplex containing a single mismatched base pair and normal duplexes above Table S5 (position 1, TC mismatch, blue; position 2, AA mismatch, green; position 3, TC mismatch, magenta; position 4, TC mismatch, purple; position 5, CT mismatch, yellow; position 6, AA mismatch, cyan; fully paired, black). The experiments were carried out at pH 7.2 under the same condition as in Figure S1: A) BD, Y = X-pyrene; B) ND, Y = G-clamp; C) ND, Y = C.

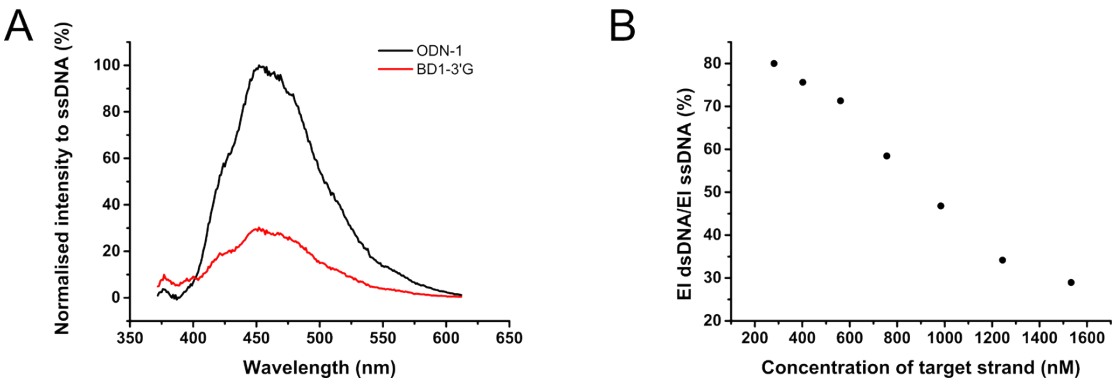


Figure S15. Fluorescence of X-pyrene and G-rich complementary strand. **A)** Fluorescence emission spectra of an X-pyrene containing oligonucleotide (ODN-1) and its duplex with the complementary strand presenting a G in the unpaired position (BD1-3′G). The target strand is in a large excess compared to the probe strand (almost 11 equivalents). The spectra are an average of four spectra from which the average blank spectrum was subtracted and they are normalized to the ssDNA spectrum. **B)** Titration curve of ODN-1 with the target strand containing a G in the unpaired position. The graph plots the ratio between the emission intensity of the BD1-3′G duplex and the emission intensity of the ODN-1 alone measured at 454 nm versus the concentration of the target strand.

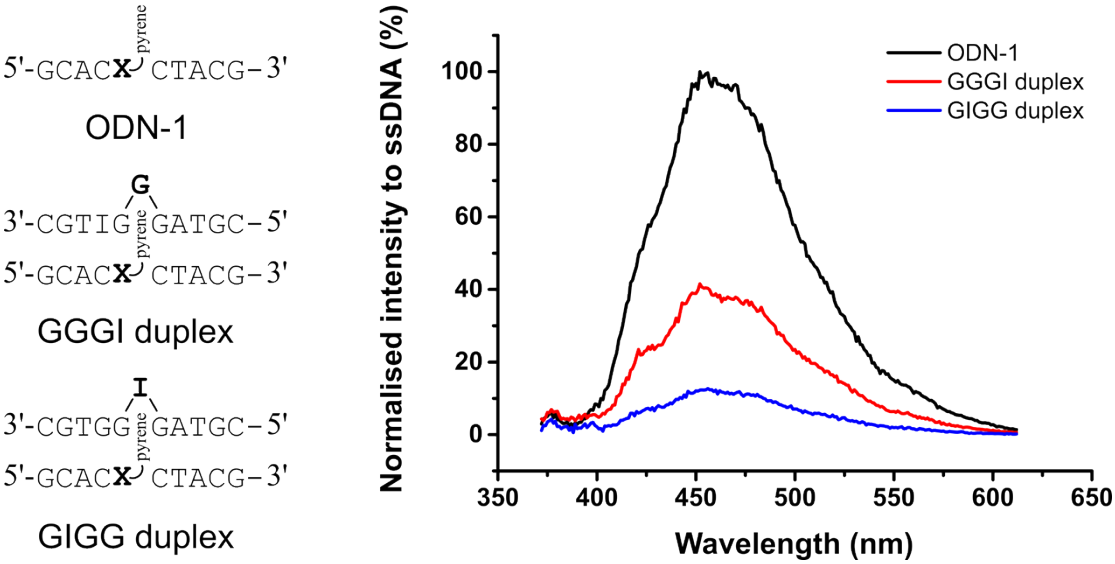


Figure S16. Fluorescence of X-pyrene and inosine containing complementary strands. Fluorescence emission spectra of an X-pyrene containing oligonucleotide (ODN-1) and its duplexes with the complementary strands presenting an inosine (I) at the 3′ end (GGGI duplex) and in the middle (GIGG duplex) of the G-rich core. The target strand is in the same concentration as the experiments in Figure 9 (200 nM). The spectra are an average of four spectra from which the average blank spectrum was subtracted and they are normalized to the ssDNA one.

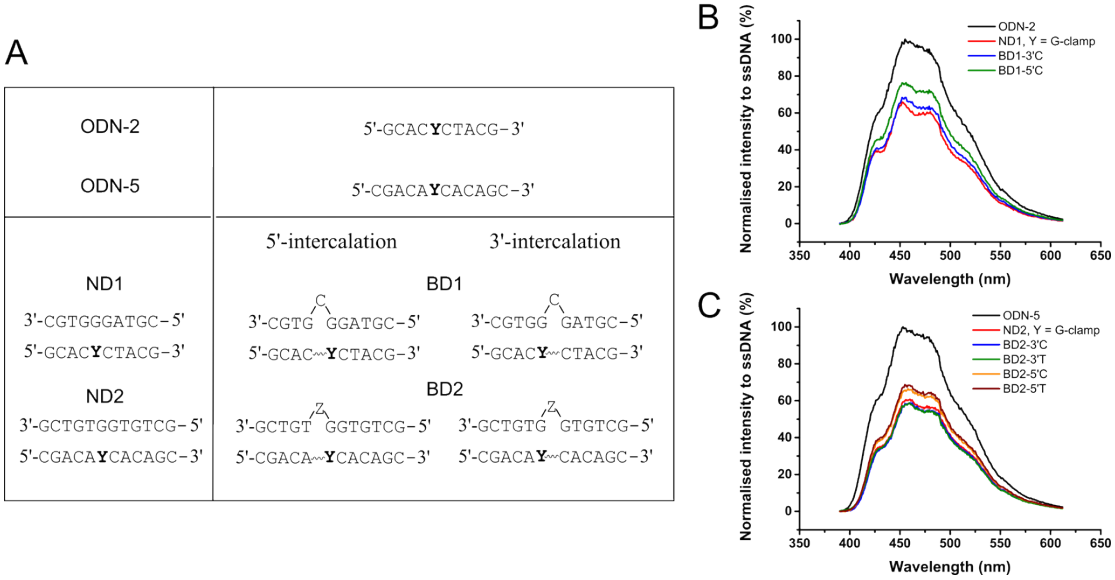


Figure S17. Fluorescence of G-clamp containing oligonucleotides. **A)** Oligonucleotides and corresponding duplexes involved in the fluorescence study; **B)** Fluorescence emission spectra of a G-clamp containing oligonucleotide (ODN-2) and its three duplexes (ND1, BD1-3′C, BD1-5′C); **C)** Fluorescence emission spectra of a G-clamp containing oligonucleotide (ODN-5) and its five duplexes (ND2, BD2-3′C, BD2-3′T, BD2-5′C, BD2-5′T). The spectra are an average of four spectra from which the average blank spectrum was subtracted. All the spectra are normalized to the ssDNA one.

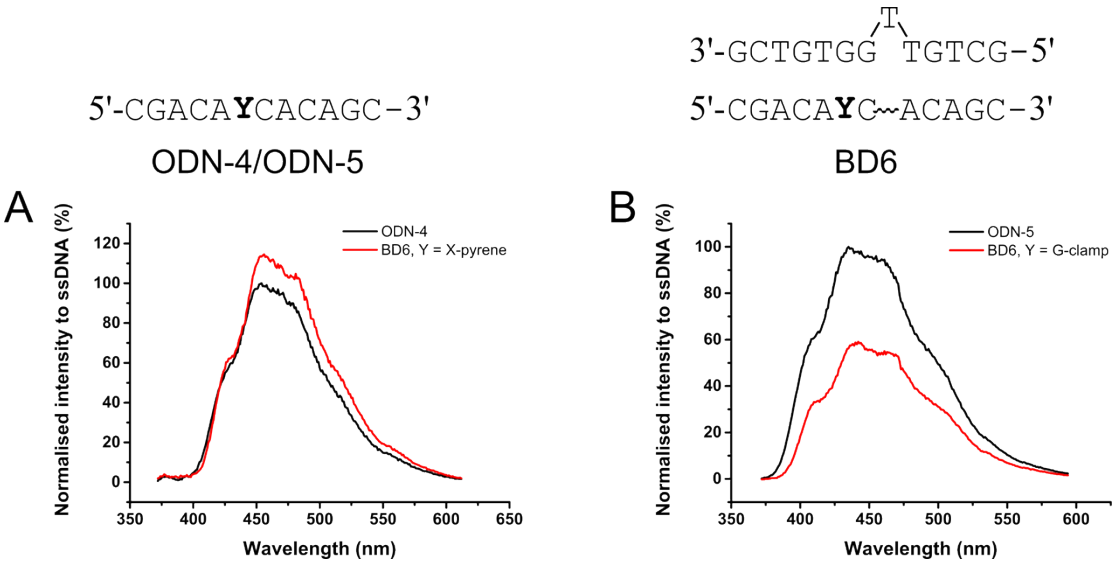


Figure S18. Fluorescence of an X-pyrene and a G-clamp oligonucleotide with the unpaired base 1 bp away. **A)** (Y = X-pyrene) Fluorescence emission spectra of a X-pyrene containing oligonucleotide (ODN-4) and its duplex with the unpaired base 1 bp away from X-pyrene on the 3′ side (BD6). **B)** (Y = G-clamp) Fluorescence emission spectra of a G-clamp containing oligonucleotide (ODN-5) and its duplex with the unpaired base 1 bp away from G-clamp on the 3′ side (BD6). The spectra are an average of four spectra from which the average blank spectrum was subtracted. All the spectra are normalized to the ssDNA one.

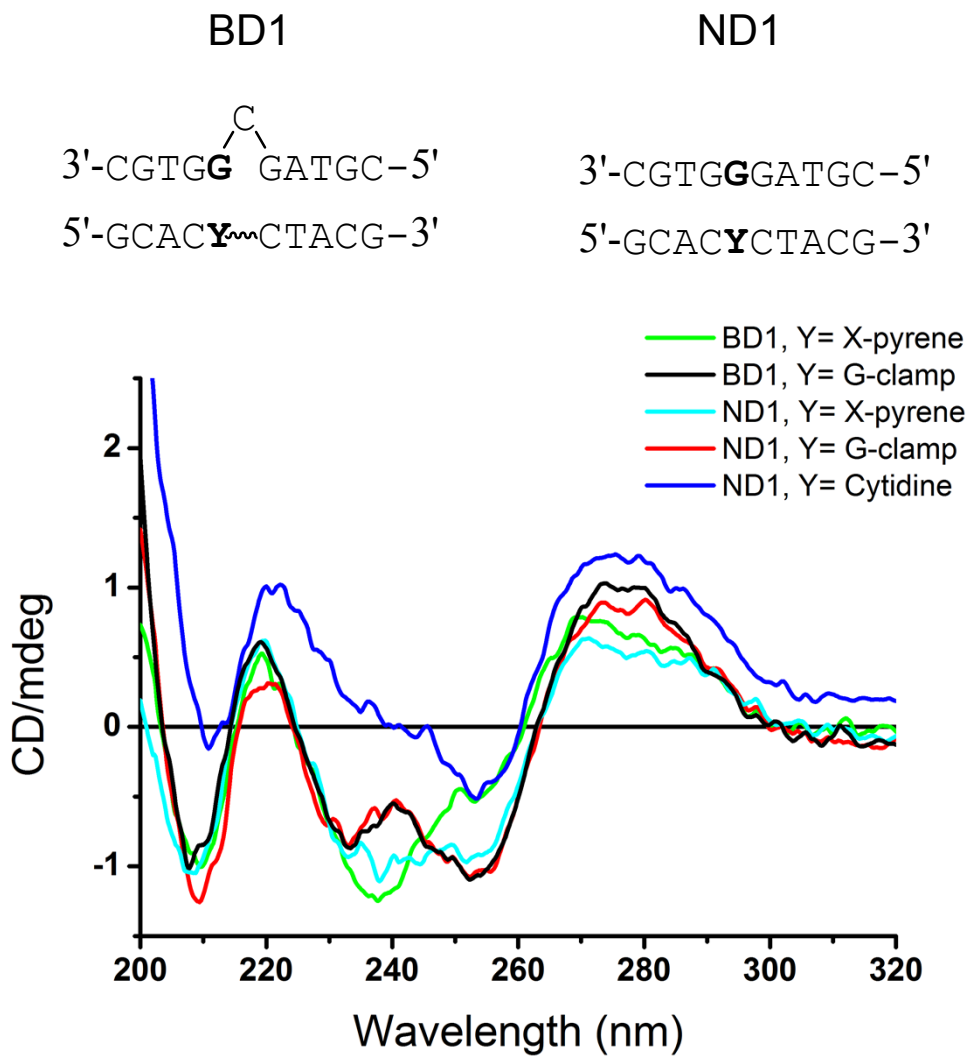


Figure S19. CD spectra of bulge duplex and normal duplexes (Y = X-pyrene, G-clamp or C) were recorded from 200 nm to 320 nm. The signals were smoothed by Savitzky-Golay method (20 pts, polynomial order = 2) in Origin 8.6 software.

Table S6: Acquisition parameters for the identical sets of spectra at temperatures 278 K, 298 K and 310 K for both X-pyrene DNA duplexes measured.

Spectrum	Time domain size (F2/F1)	Acquisition times F2/ F1 (ms)	Sweepwidth F2/F1 (kHz)	Repetition delay (s)	No. of scans	Total experiment time (h)
NOESY D_2O	2048/512	113.9/28.5	8.99/8.99	2	16	5.0
COSY D_2O	2048/256	67.0/8.4	15.29/15.29	1	32	2.5
TOCSY D_2O	2048/256	67.0/8.4	15.29/15.29	1	16	1.25
HSQC D_2O	8192/512	525.9/16.2	7.79/15.85	1	96	21
NOESY H_2O	2048/512	75.8/18.9	13.52/13.52	2	16	5.5
SF-HMQC H_2O	1642/128	49.9/28.1	16.45/2.28	0.2	6144	60

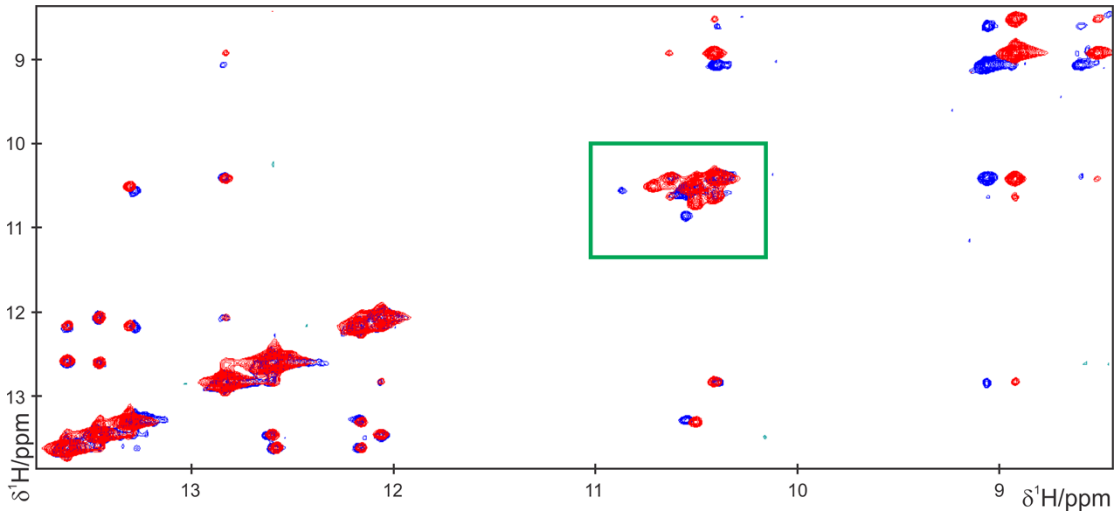


Figure S20. Overlay of the Imino-NOESY spectra for BD2 (red) and BD2-C (blue) at 278K. The overlap between the two spectra is considerable and indicates that no significant changes are introduced when changing the widowed base from T to C. This is also reflected in the D_2O NOESY (data not shown). The green box indicates the region where the low intensity, second conformation imino protons are best observed.

Table S7: Resonance assignments at 310 K for the X-pyrene DNA duplex with T as the widowed base (referenced to the HOD-signal at 4.81 ppm). The resonances of the Pyrene moiety and the linker could not be assigned due to severe broadening and resonance overlap.

Residue	H1'	H2'	H2''	H3'	H1/H2/H3/H41/H42	H5/H7	H6/H8
G1	6.041	2.685	2.81	4.88	-	-	8.014
C2	6.114	2.159	2.547	4.869	8.222/6.577	5.417	7.574
T3	5.688	2.152	2.435	4.896	13.875	1.682	7.38
G4	5.876	2.595	2.654	4.969	12.422	-	7.914
T5	5.416	1.45	1.804	4.698	13.567	1.375	6.927
G6	5.445	2.596	2.415	5.101	10.757	-	7.611
T7	6.608	2.521	2.591	5.176	-	2.093	8.006
G8	6.146	2.548	2.761	4.633	10.673	-	7.664
T9	5.727	2.101	2.501	4.835	13.089	1.223	7.21
G10	5.974	2.562	2.761	4.918	12.322	-	7.844
T11	6.063	2.13	2.489	4.896	13.715	1.378	7.345
C12	5.739	2.028	2.401	4.852	8.552/7.042	5.702	7.51
G13	6.175	2.629	2.401	4.701	-	-	7.964
C14	5.767	1.894	2.396	4.709	-	5.943	7.655
G15	5.528	2.72	2.81	5.016	12.86	-	7.983
A16	6.239	2.668	2.901	5.04	7.815	-	8.19
C17	5.503	1.714	2.155	4.753	8.127/6.546	5.156	7.105
A18	6.123	2.627	2.889	5.05	7.444	-	7.803
X19	6.13	2.336	2.532	5.061	-	-	-
C20	5.251	2.248	2.349	4.932	-	5.963	7.752
A21	6.153	2.764	2.859	5.083	7.483	-	8.283
C22	5.228	1.918	2.211	4.787	8.09/6.453	5.355	7.291
A23	5.974	2.726	2.862	5.034	7.593	-	8.154
G24	5.809	2.489	2.649	4.955	12.838	-	7.694
C25	6.146	2.139	2.218	4.465	-	5.439	7.432

X=GPY	H4	H6	H7	H8
	7.159	6.545	6.974	6.563

Table S8. Oligonucleotide sequences and MS analysis

Code	Oligonucleotide sequence (5′-3′)	MS (Da)	
		Calc.	Found
ODN-1	GCACYCTACG	3378	3378
ODN-4	CGACAYCACAGC	3990	3990
ODN-7	GCACYYATCG	3784	3783
ODN-9	GCYATACYAG	3808	3808
ODN-11	GCAAYATACG	3426	3426
ODN-12	GCACZCTACG	3517	3517

Note: Mass spectra of modified oligonucleotides were recorded on a Bruker micrOTOF™ II focus ESI-TOF MS instrument in ES^- mode. Y = X-pyrene, Z= 6-azidohexanoic labeled X-pyrene.

References:

1. H. Du, R. C. A. Fuh, J. Z. Li, L. A. Corkan and J. S. Lindsey, *Photochem. Photobiol.*, 1998, **68**, 141-142.
2. K. Y. Lin, R. J. Jones and M. Matteucci, *J. Am. Chem. Soc.*, 1995, **117**, 3873-3874.
3. P. Kočalka, A. H. El-Sagheer and T. Brown, *ChemBioChem*, 2008, **9**, 1280-1285.
4. P. Sandin, K. Borjesson, H. Li, J. Martensson, T. Brown, L. M. Wilhelmsson and B. Albinsson, *Nucleic Acids Res.*, 2008, **36**, 157-167.
5. M. Shelbourne, X. Chen, T. Brown and A. H. El-Sagheer, *Chem. Commun.*, 2011, **47**, 6257-6259.
6. F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer and A. Bax, *J. Biomol. NMR*, 1995, **6**, 277-293.
7. R. L. J. Keller, *The Computer Aided Resonance Assignment Tutorial*, Cantina Verlag press, Goldau, 2004.
8. J. A. Ortega, J. R. Blas, M. Orozco, A. Grandas, E. Pedroso and J. Robles, *Org. Lett.*, 2007, **9**, 4503-4506.