Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry. This journal is © The Royal Society of Chemistry 2014

## Recognition of double-stranded DNA using energetically activated duplexes with interstrand zippers of 1-, 2- or 4-pyrenyl-functionalized O2'-alkylated RNA monomers

Saswata Karmakar, Andreas S. Madsen, Dale C. Guenther, Bradley C. Gibbons and Patrick J.

Hrdlicka\*

## hrdlicka@uidaho.edu

## **ELECTRONIC SUPPLEMENTARY INFORMATION**

General experimental section	<b>S</b> 2
Synthesis of 2-pyrenemethanol and 4-pyrene methanol (Schemes S1 and S2)	<b>S</b> 3
MALDI-MS of modified ONs (Table S1)	S4
Representative thermal denaturation curves (Fig. S1)	<b>S</b> 5
DNA selectivity of <b>B1-B6</b> (Table S2)	<b>S</b> 5
Additional data and discussion of mismatch specificity (Tables S3 and S4)	<b>S</b> 6
Absorbance spectra of <b>X1-Z5</b> in absence or presence of cDNA or cRNA (Figs. S2-S4)	<b>S</b> 8
Steady-state fluorescence emission spectra of of <b>X1-Z5</b> in absence or presence of cDNA or cRNA (Figs. S5-S7)	S11
Thermodynamic parameters of X/Y/Z-modified duplexes (Tables S5 and S6)	S14
Absorbance spectra of representative Invaders (Fig. S8)	S16
Pseudorotational phase angles of selected nucleotides in <b>B2:D4</b> duplexes (Table S7)	S17
Pseudorotational phase angles of selected nucleotides in <b>B2:B5</b> duplexes (Table S8)	S17
Structural base pair and dinucleotide step parameters for duplexes <b>D1:D4, B2:D4, B2:B5</b> (Type A) and <b>B2:B5</b> (Type B) (Figs. S9-S14)	S18
Energy minimized structures of X2:X5 and Z2:Z5 (Type B) (Fig. 15)	S24
References	S25
<sup>1</sup> H, COSY, <sup>13</sup> C, DEPT, and HSQC spectra of nucleosides <b>2Y/2Z/3Y/3Z/4Y/4Z</b>	S26

General experimental section. Reagents and solvents were commercially available, of analytical grade and used without further purification. Petroleum ether of the distillation range 60-80 °C was used. Solvents were dried over activated molecular sieves: THF (3Å); CH<sub>2</sub>Cl<sub>2</sub>, 1,2dichloroethane, N,N'-diisopropylethylamine and anhydrous DMSO (4Å). Anhydrous diethyl ether, DMF and benzene were obtained from commercial suppliers (<10 ppm water). Water content of anhydrous solvents was verified on Karl-Fisher apparatus. Reactions were conducted under argon whenever anhydrous solvents were used. Reactions were monitored by TLC using silica gel coated plates with a fluorescence indicator (SiO<sub>2</sub>-60, F-254) which were visualized a) under UV light and/or b) by dipping in 5% conc. H<sub>2</sub>SO<sub>4</sub> in absolute ethanol (v/v) followed by heating. Silica gel column chromatography was performed with Silica gel 60 (particle size 0.040-0.063 mm) using moderate pressure (pressure ball). Evaporation of solvents was carried out under reduced pressure at temperatures below 45 °C. After column chromatography, appropriate fractions were pooled, evaporated and dried at high vacuum for at least 12h to give the obtained products in high purity (>95%) as ascertained by 1D NMR techniques. Chemical shifts of <sup>1</sup>H NMR (500 MHz), <sup>13</sup>C NMR (125.6 MHz), and/or <sup>31</sup>P NMR (121.5 MHz) are reported relative to deuterated solvent or other internal standards (80% phosphoric acid for <sup>31</sup>P NMR). Exchangeable (ex) protons were detected by disappearance of signals upon  $D_2O$  addition. Assignments of NMR spectra are based on 2D spectra (COSY, HSQC) and DEPT-spectra. Quaternary carbons are not assigned in <sup>13</sup>C NMR but verified from HSQC and DEPT spectra (absence of signals). MALDI-HRMS spectra of compounds were recorded on a Waters Q-TOF Premiere mass spectrometer using 2,5-dihydroxybenzoic acid as a matrix and polyethylene glycol (PEG 600) as an internal calibration standard.



4,5,9,10-tetrahydropyrene 2-pyrenemethanol

**Scheme S1**. Synthesis of 2-pyrenemethanol. Reagents and conditions: (a) Br<sub>2</sub>, DMF, rt, 6h; (b) DDQ, benzene, reflux, 4h; (c) (i) *n*-BuLi, THF/ether, -78 °C for 2h; (ii) DMF, -30 to -50 °C for 1h, then rt for 15h; (d) NaBH<sub>4</sub>, THF, rt, overnight.



**Scheme S2**. Synthesis of 4-pyrenemethanol. Reagents and conditions: (a) Br<sub>2</sub>, AcOH, rt, 1h; (b) DDQ, benzene, reflux, 4h; (c) (i) *n*-BuLi, THF/ether, -78 °C for 2h; (ii) DMF, -50 to -30 °C for 1h, then rt for 15h; (d) NaBH<sub>4</sub>, THF, rt, overnight.

ONs	Sequence	Calc. $m/z$ [M+H]	Found $m/z$ [M+H]
Y1	5'-GYG ATA TGC	2969	2969
Y2	5'-GTG A <u>Y</u> A TGC	2969	2969
<b>Y3</b>	5'-GTG ATA <u>Y</u> GC	2969	2969
<b>Y4</b>	3'-CAC <u>Y</u> AT ACG	2898	2898
<b>Y5</b>	3'-CAC TA <u>Y</u> ACG	2898	2898
<b>Y6</b>	3'-CAC $\underline{\mathbf{Y}}A\underline{\mathbf{Y}}$ ACG	3114	3114
<b>Z</b> 1	5'-G <u>Z</u> G ATA TGC	2969	2969
<b>Z</b> 2	5'-GTG A <u>Z</u> A TGC	2969	2969
<b>Z</b> 3	5'-GTG ATA <u>Z</u> GC	2969	2969
<b>Z4</b>	3'-CAC <u>Z</u> AT ACG	2898	2898
Z5	3'-CAC TA <u>Z</u> ACG	2898	2898
Z6	3'-CAC <u>Z</u> A <u>Z</u> ACG	3114	3114

Table S1. MALDI-MS of ONs modified with monomers  $\mathbf{Y}$  and  $\mathbf{Z}$ .<sup>a</sup>

<sup>a</sup> For structures of monomers **Y** and **Z**, see Figure 1 in the main manuscript.



**Figure S1**. Representative thermal denaturation curves of **X/Y/Z**-modified duplexes and reference duplex **D1**:**D2**. For experimental conditions, see Table 1.

			$\Delta\Delta T_{\rm m}/{\rm mod}$ (DNA-RNA) [°C						
ON	Duplex	<u>B</u> =	X	Y	Z				
<b>B1</b>	5'-G <u>B</u> G ATA TGC		+7.0	+7.5	+6.0				
<b>B2</b>	5'-GTG A <u>B</u> A TGC		+9.0	+6.5	+7.0				
<b>B3</b>	5'-GTG ATA <u>B</u> GC		+8.0	+8.5	+9.0				
<b>B4</b>	3'-CAC <u>B</u> AT ACG		+8.0	+7.5	+5.0				
B5	3'-CAC TA <u>B</u> ACG		+10.5	+6.5	+7.0				
<b>B6</b>	3'-CAC <u><b>B</b></u> A <u><b>B</b></u> ACG		+7.3	+6.0	+5.5				

Table S2. DNA selectivity of B1-B6.<sup>a</sup>

<sup>a</sup> DNA selectivity defined as  $\Delta\Delta T_{\rm m}/{\rm mod}$  (DNA-RNA) =  $\Delta T_{\rm m}/{\rm mod}$  (vs cDNA) -  $\Delta T_{\rm m}/{\rm mod}$  (vs cRNA).

Additional discussion of binding specificity. X2/Y2/Z2 display less efficient discrimination of mismatched RNA targets compared to reference strand D1, except when a mismatched G is opposite of the pyrene-functionalized monomers (Table S3).

ONs with two modifications positioned as next-nearest neighbors (**B6**-series) very efficiently discriminate DNA targets with a single mismatched nucleotide opposite of the central adenosine residue (Table S4). Similar observations have been observed with analogous ONs modified with other O2'-pyrene-functionalized RNA monomers or N2'-pyrene-functionalized 2'-*N*-methyl-2'-amino-DNA,<sup>S1</sup> which suggests that this probe architecture may be a general strategy toward improving binding specificity. However, this hypothesis must be evaluated across different sequence contexts.

				RNA: 3'-CA	AC U <u>B</u> U A	lCG		
		-	$T_{\rm m}/^{\circ}{ m C}$	$\Delta T_{ m m}/^{\circ}{ m C}$				
ON	Sequence	<u>B</u> =	А	С	G	U		
D1	5'-GTG ATA TGC		26.5	<-16.5	-4.5	<-16.5		
$\mathbf{X2}^{\mathrm{b}}$	5'-GTG A <u>X</u> A TGC		31.0	-17.5	-3.5	-9.5		
Y2	5'-GTG A <u>Y</u> A TGC		34.0	-13.5	-8.5	-13.5		
Z2	5'-GTG A <u>Z</u> A TGC		27.5	-13.0	-7.0	-8.0		

Table S3. Discrimination of mismatched RNA targets by X2/Y2/Z2 and reference ONs.<sup>a</sup>

<sup>a</sup> For conditions of thermal denaturation experiments, see Table 1.  $T_{\rm m}$ 's of fully matched duplexes are shown in bold.  $\Delta T_{\rm m}$  = change in  $T_{\rm m}$  relative to fully matched DNA:RNA duplex.

<sup>b</sup> Data previously reported in reference S1. Included to facilitate direct comparison.

					DNA : 5'-GT	G A <u>B</u> A ACG	
			$T_{\rm m} [^{\circ} \rm C]$			$\Delta T_{\rm m} [^{\circ}{\rm C}]$	
ON	Sequence	<u>B</u> =	Т		А	С	G
D2	3'-CAC TAT ACG		29.5	-	-17.0	-15.5	-9.0
<b>X6</b>	3'-CAC <u>X</u> A <u>X</u> ACG		43.5		-24.0	-17.0	-14.0
¥6	3'-CAC $\underline{\mathbf{Y}}A\underline{\mathbf{Y}}$ ACG		45.5		-24.0	-16.0	-14.0
Z6	3'-CAC $\underline{Z}$ A $\underline{Z}$ ACG		32.5		-21.0	-16.0	-12.0

Table S4. Discrimination of mismatched DNA targets by X6/Y6/Z6 and reference ONs.<sup>a</sup>

<sup>a</sup> For conditions of thermal denaturation experiments, see Table 1.  $T_{\rm m}$ 's of fully matched duplexes are shown in bold.  $\Delta T_{\rm m}$  = change in  $T_{\rm m}$  relative to fully matched DNA:DNA duplex.

<sup>b</sup> Data previously reported in reference S1. Included to facilitate direct comparison.



Figure S2. Absorbance spectra of single-stranded X1-X5 and their corresponding duplexes with DNA/RNA targets. Spectra were recorded at T = 5 °C using each strand at 1.0  $\mu$ M concentration in  $T_{\rm m}$  buffer.



**Figure S3**. Absorbance spectra of single-stranded **Y1-Y5** and their corresponding duplexes with DNA/RNA targets. For conditions, see Fig. S2.



**Figure S4**. Absorbance spectra of single-stranded **Z1-Z5** and their corresponding duplexes with DNA/RNA targets. For conditions, see Fig. S2.



**Figure S5**. Steady-state fluorescence emission spectra of **X**-modified ONs and their corresponding duplexes with DNA/RNA targets. Spectra were recorded at T = 5 °C using  $\lambda_{ex} = 350$  nm and each strand at 1.0  $\mu$ M concentration in  $T_m$  buffer. Some, but not all, of these spectra have been previously published in reference S1. Spectra are included to facilitate direct comparison.



Figure S6. Steady-state fluorescence emission spectra of Y-modified ONs and their corresponding duplexes with DNA/RNA targets.  $\lambda_{ex} = 345$  nm. For other conditions, see Fig. S5.



**Figure S7**. Steady-state fluorescence emission spectra of **Z**-modified ONs and their corresponding duplexes with DNA/RNA targets.  $\lambda_{ex} = 340$  nm. For other conditions, see Fig. S5.

		$\Delta H^{293}[\Delta \Delta H^{293}] \text{ (kJ/mol)}$								
ON	ZP	Sequence	upper ON	lower ON	probe	$\Delta H_{\rm rec}$				
87.1			vs cDNA	VS CDNA	duplex	(KJ/MOI)				
XI X5	+4	$3'-G\underline{\mathbf{X}}G$ ATA TGC $3'-CAC TA\underline{\mathbf{X}}$ ACG	-281±4 [+25]	-299±2 [+7]	-279±1 [+27]	+5				
X1 X4	+2	5'-G <u>X</u> G ATA TGC 3'-CAC <u>X</u> AT ACG	-281±4 [+25]	-300±7 [+6]	-248±8 [+58]	-27				
X2 X5	+1	5'-GTG A <u>X</u> A TGC 3'-CAC TA <u>X</u> ACG	-305±1 [+1]	-299±2 [+7]	-244±3 [+62]	-54				
X2 X4	-1	5'-GTG A <u>X</u> A TGC 3'-CAC <u>X</u> AT ACG	-305±1 [+1]	-300±7 [+6]	-296±6 [+10]	-3				
X3 X5	-1	5'-GTG ATA <u>X</u> GC 3'-CAC TA <u>X</u> ACG	-270±7 [+36]	-299±2 [+7]	-280±5 [+26]	+17				
X3 X4	-3	5'-GTG ATA <u>X</u> GC 3'-CAC <u>X</u> AT ACG	-270±7 [+36]	-300±7 [+6]	-309±4 [-3]	+45				
Y1 Y5	+4	5'-G <u>¥</u> G ATA TGC 3'-CAC TA <u>¥</u> ACG	-292±14 [+14]	-306±15 [±0]	-277±2 [+29]	-15				
Y1 Y4	+2	5'-G <u>¥</u> G ATA TGC 3'-CAC <u>¥</u> AT ACG	-292±14 [+14]	-271±20 [+35]	-238±3 [+68]	-19				
Y2 Y5	+1	5'-GTG A <u>Y</u> A TGC 3'-CAC TA <u>Y</u> ACG	-309±10 [-3]	-306±15 [±0]	-257±5 [+49]	-52				
Y2 Y4	-1	5'-GTG A <u>Y</u> A TGC 3'-CAC <u>Y</u> AT ACG	-309±10 [-3]	-271±20 [+35]	-307±6 [+1]	+31				
Y3 Y5	-1	5'-GTG ATA <u>Y</u> GC 3'-CAC TA <u>Y</u> ACG	-317±10 [-11]	-306±15 [±0]	-288±2 [+18]	-29				
Y3 Y4	-3	5'-GTG ATA <u>Y</u> GC 3'-CAC <u>Y</u> AT ACG	-317±10 [-11]	-271±20 [+35]	-276±4 [+30]	-6				
Z1 Z5	+4	5'-G <b>Z</b> G ATA TGC 3'-CAC TA <b>Z</b> ACG	-309±8 [-3]	-253±26 [+53]	-252±4 [+54]	-4				
Z1 Z4	+2	5'-G <u>Z</u> G ATA TGC 3'-CAC <u>Z</u> AT ACG	-309±8 [-3]	-332±7 [-26]	-251±5 [+55]	-84				
Z2 Z5	+1	5'-GTG A <u>Z</u> A TGC 3'-CAC TA <u>Z</u> ACG	-260±8 [+46]	-253±26 [+53]	N/A	-				
Z2 Z4	-1	5'-GTG A <u>Z</u> A TGC 3'-CAC <u>Z</u> AT ACG	-260±8 [+46]	-332±7 [-26]	-283±9 [+23]	-3				
Z3 Z5	-1	5'-GTG ATA <u>Z</u> GC 3'-CAC TA <u>Z</u> ACG	-280±4 [+26]	-253±26 [+53]	-251±2 [+55]	+24				
Z3 Z4	-3	5'-GTG ATA <u>Z</u> GC 3'-CAC <u>Z</u> AT ACG	-280±4 [+26]	-332±7 [-26]	-247±1 [+59]	-59				

**Table S5**. Change in enthalpy upon duplex formation ( $\Delta H$ ) and change in enthalpy upon probe recognition of iso-sequential dsDNA target **D1**:**D4** ( $\Delta H_{rec}$ ).<sup>a</sup>

<sup>a</sup>  $\Delta\Delta H$  is measured relative to  $\Delta H$  for **D1:D4** = -306 kJ/mol.  $\Delta H_{rec} = \Delta H$  (upper strand vs cDNA) +  $\Delta H$  (lower strand vs cDNA) -  $\Delta H$  (probe duplex) -  $\Delta H$  (dsDNA target). "±" denotes standard deviation. N/A = the lack of a clear lower base line prevented determination of this value.

			- <i>T</i> <sup>293</sup>			
Duplex	Zipper	Sequence	upper strand vs cDNA	lower strand vs cDNA	probe duplex	-T <sup>293</sup> ΔS <sub>rec</sub> (kJ/mol)
X1 X5	+4	5'-G <u>X</u> G ATA TGC 3'-CAC TA <u>X</u> ACG	235±3 [-30]	247±6 [-18]	224±1 [-41]	-7
X1 X4	+2	5'-G <u>X</u> G ATA TGC 3'-CAC <u>X</u> AT ACG	235±3 [-30]	254±6 [-11]	203±8 [-62]	+21
X2 X5	+1	5'-GTG A <u>X</u> A TGC 3'-CAC TA <u>X</u> ACG	251±1 [-14]	247±6 [-18]	204±3 [-61]	+29
X2 X4	-1	5'-GTG A <u>X</u> A TGC 3'-CAC <u>X</u> AT ACG	251±1 [-14]	254±6 [-11]	245±5 [-20]	-5
X3 X5	-1	5'-GTG ATA <u>X</u> GC 3'-CAC TA <u>X</u> ACG	220±7 [-45]	247±6 [-18]	225±5 [-40]	-23
X3 X4	-3	5'-GTG ATA <u>X</u> GC 3'-CAC <u>X</u> AT ACG	220±7 [-45]	254±6 [-11]	256±3 [-9]	-47
Y1 Y5	+4	5'-G <u>¥</u> G ATA TGC 3'-CAC TA <u>¥</u> ACG	244±14 [-21]	251±14 [-14]	221±2 [-44]	+9
Y1 Y4	+2	5'-G <u>Y</u> G ATA TGC 3'-CAC <u>Y</u> AT ACG	244±14 [-21]	225±20 [-40]	193±3 [-72]	-3
Y2 Y5	+1	5'-GTG A <u>¥</u> A TGC 3'-CAC TA <u>¥</u> ACG	253±10 [-12]	251±14 [-14]	214±5 [-51]	+25
Y2 Y4	-1	5'-GTG A <u>¥</u> A TGC 3'-CAC <u>¥</u> AT ACG	253±10 [-12]	225±20 [-40]	252±5 [-13]	-39
Y3 Y5	-1	5'-GTG ATA <u>Y</u> GC 3'-CAC TA <u>Y</u> ACG	266±11 [+1]	251±14 [-14]	230±2 [-35]	+22
Y3 Y4	-3	5'-GTG ATA <u>Y</u> GC 3'-CAC <u>Y</u> AT ACG	266±11 [+1]	225±20 [-40]	223±4 [-42]	+3
Z1 Z5	+4	5'-G <b>Z</b> G ATA TGC 3'-CAC TA <b>Z</b> ACG	265±8 [±0]	205±24 [-60]	203±3 [-62]	+2
Z1 Z4	+2	5'-G <u>Z</u> G ATA TGC 3'-CAC <u>Z</u> AT ACG	265±8 [±0]	291±7 [+26]	210±5 [-55]	+81
Z2 Z5	+1	5'-GTG A <u>Z</u> A TGC 3'-CAC TA <u>Z</u> ACG	212±8 [-53]	205±24 [-60]	N/A	N/A
Z2 Z4	-1	5'-GTG A <u>Z</u> A TGC 3'-CAC <u>Z</u> AT ACG	212±8 [-53]	291±7 [+26]	242±11 [-23]	-4
Z3 Z5	-1	5'-GTG ATA <u>Z</u> GC 3'-CAC TA <u>Z</u> ACG	234±3 [-31]	205±24 [-60]	205±2 [-60]	-31
Z3 Z4	-3	5'-GTG ATA <u>Z</u> GC 3'-CAC <u>Z</u> AT ACG	234±3 [-31]	291±7 [+26]	205±1 [-60]	+55

**Table S6**. Change in entropy at 293K upon duplex formation  $(-T^{293}\Delta S)$  and change in entropy upon probe recognition of iso-sequential dsDNA target **D1:D4**  $(-T^{293}\Delta S_{rec})$ .<sup>a</sup>

<sup>a</sup>  $\Delta(T^{293}\Delta S)$  is measured relative to  $-T^{293}\Delta S$  for **D1:D4** = 265 kJ/mol.  $-T^{293}\Delta S_{rec} = T^{293}\Delta S$  (upper strand vs cDNA) +  $T^{293}\Delta S$  (lower strand vs cDNA) -  $T^{293}\Delta S$  (probe duplex) -  $T^{293}\Delta S$  (dsDNA target). "±" denotes standard deviation. N/A = the lack of a clear lower base line prevented determination of this value.



Figure S8. Absorbance spectra of representative Invaders, duplexes between individual probe strands and cDNA, and single-stranded probes (SSP). Spectra were recorded at T = 5 °C using each strand at 1.0  $\mu$ M concentration in  $T_m$  buffer.

Р	D1:D4		X2:D4		Y2:D4			Z2:D4		
A <sub>4</sub> :T <sub>15</sub>	93	112	96	116	105	110		102	106	
B <sub>5</sub> :A <sub>14</sub>	129	111	142	98	148	103		149	175	
A <sub>6</sub> :B <sub>13</sub>	102	109	124	97	155	113		171	141	
T <sub>7</sub> :A <sub>12</sub>	121	123	90	100	106	102		111	98	

 Table S7. Pseudorotational phase angles P of selected nucleotides in B2:D4 duplexes.

 Table S8. Pseudorotational phase angles P of selected nucleotides in B2:B5 duplexes.

Р	D1	:D4	X2:X5				Y2:Y5			Z2:Z5				
			Тур	e A	Тур	be B	Type A			Тур	e A	Тур	e B	
A <sub>4</sub> :T <sub>15</sub>	93	112	70	103	100	104	101	106		98	105	100	106	
B <sub>5</sub> :A <sub>14</sub>	129	111	161 <sup>a</sup>	122	149	168	148	104		147	107	150	180	
A <sub>6</sub> :B <sub>13</sub>	102	109	123	160	176	148	104	147		107	146	173	$148^{a}$	
T <sub>7</sub> :A <sub>12</sub>	121	123	104	71	106	116	109	117		105	110	106	111	

<sup>a</sup> Not intercalated.



Figure S9. Structural base pair parameters for duplexes D1:D4 and B2:D4.



Figure S10. Structural dinucleotide step parameters for duplexes D1:D4 and B2:D4.



Figure S11. Structural base pair parameters for duplexes D1:D4 and B2:B5 (Type A).



Figure S12. Structural dinucleotide step parameters for duplexes D1:D4 and B2:B5 (Type A).



Figure S13. Structural base pair parameters for duplexes D1:D4 and B2:B5 (Type B).



Figure S14. Structural dinucleotide step parameters for duplexes D1:D4 and B2:B5 (Type B).



**Figure S15**. Lowest energy structures of **X2**:**X5** (left) and **Z2**:**Z5** (right) in Type B conformation (**Y2**:**Y5** maintained a Type A conformation throughout the simulation protocol). Top: side view of duplex; bottom: top view of the central duplex region. Color code: sugar phosphate backbone (red); pyrene moieties (blue) and nucleobases (green). Hydrogen atoms, sodium ions and bond orders are omitted for clarity.

## References.

(S1) S. Karmakar, B. A. Anderson, R. L. Rathje, S. Andersen, T. Jensen, P. Nielsen and P. J. Hrdlicka, J. Org. Chem., 2011, 76, 7119-7131.



S26




































S42















S46































S57


































