Recognition of mixed-sequence DNA using double-stranded probes with interstrand zipper arrangements of O2'-triphenylene- and coronene-functionalized RNA monomers

Saswata Karmakar,[a] Dale C. Guenther,[a] Bradley C. Gibbons,[a,b] and Patrick J. Hrdlicka*[a]

[a]Department of Chemistry, University of Idaho, Moscow, ID-83844, USA

[b]Department of Chemistry, Brigham Young University-Idaho, ID-83440, USA

*Corresponding author: Ph: (+1) 208 885 0108. Fax: (+1) 208 885 6173. Email: hrdlicka@uidaho.edu

ELECTRONIC SUPPLEMENTARY INFORMATION

General experimental section.

Reagents and solvents were of analytical grade and used as received from vendors unless otherwise specified. Petroleum ether of the distillation range 60-80 °C was used. Solvents were dried over activated molecular sieves: acetonitrile and THF (3Å); CH_2Cl_2, 1,2-dichloroethane, *N,N'*-diisopropylethylamine and anhydrous DMSO (4Å). The water content of anhydrous solvents was verified on Karl-Fisher apparatus. Reactions were monitored by TLC using silica gel coated plates with a fluorescence indicator (SiO_2-60, F-254) which were visualized a) under UV light and b) by dipping in 5% conc. H_2SO_4 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. Following column chromatography, appropriate fractions were pooled, evaporated and dried at high vacuum for at least 12 h to give the obtained products in high purity (>95%) as ascertained by 1D NMR techniques. Chemical shifts of ^{1}H NMR (500 MHz), ^{13}C NMR (125.6 MHz), and ^{31}P NMR (121.5 MHz) are reported relative to deuterated solvent or other internal standards (80% phosphoric acid for ^{31}P NMR). Exchangeable (ex) protons were detected by the 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 ^{13}C NMR, but their presence is 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.

Table S1. MALDI-MS of ONs modified with monomer **X**.[a]

ONs	Sequence	Calc. *m*/*z* [M+H]	Found *m*/*z* [M+H]
X1	5'-G**X**G ATA TGC	2996	2996
X2	5'-GTG A**X**A TGC	2996	2996
X3	5'-GTG ATA **X**GC	2996	2996
X4	3'-CAC **X**AT ACG	2924	2925
X5	3'-CAC TA**X** ACG	2925	2925
X6	3'-CAC **X**A**X** ACG	3167	3167

[a] For the structure of monomer **X**, see Figure 1 in the main manuscript.

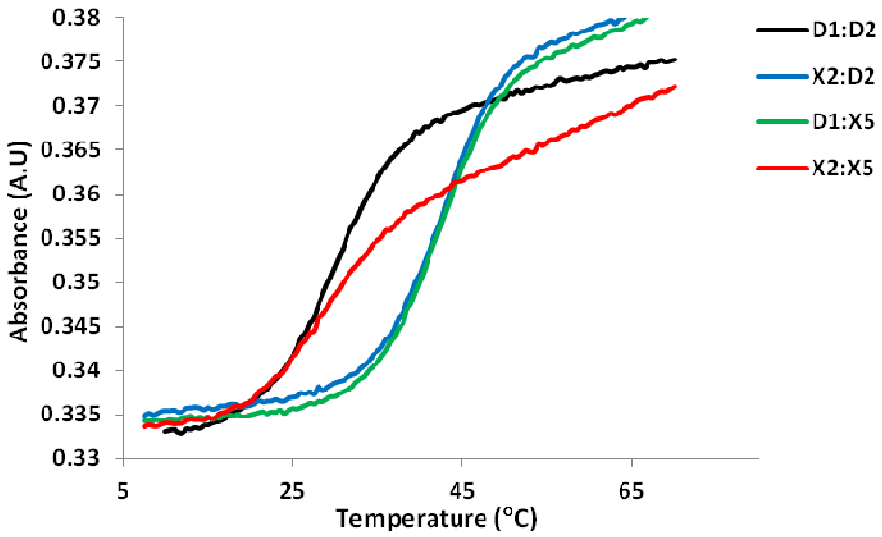


Figure S1. Representative thermal denaturation curves of **X**-modified duplexes and reference duplex **D1**:**D2**. For experimental conditions, see Table 1.

Table S2. DNA-selectivity of **B1**-**B6**.[a]

			ΔΔT_m/mod (DNA-RNA) [°C]		
ON	Sequence	**B** =	**X**	**Y**[b]	**Z**[b]
B1	5'-G**B**G ATA TGC		+5.0	+7.0	+9.5
B2	5'-GTG A**B**A TGC		+7.5	+9.0	+9.0
B3	5'-GTG ATA **B**GC		+8.0	+8.0	+7.0
B4	3'-CAC **B**AT ACG		+7.0	+8.0	+8.5
B5	3'-CAC TA**B** ACG		+6.0	+10.5	+10.0
B6	3'-CAC **B**A**B** ACG		+7.0	+7.3	+15.6

[a] DNA selectivity defined as ΔΔT_m/mod (DNA-RNA) = ΔT_m/mod (vs DNA) - ΔT_m/mod (vs RNA). [b] Data from reference S1 are included to facilitate comparison.

*RNA mismatch discrimination of **X**-modified ONs*.

Triphenylene-functionalized **X2** displays relatively efficient discrimination of mismatched RNA targets, regardless of the nature of the mismatched nucleotide opposite of monomer **X** (Table S3). The RNA mismatch discrimination characteristics of **X2** resemble those of the coronene-functionalized **Z2**, in particularly the ability to efficiently discriminate the wobble base pair with G. The pyrene-functionalized **Y2**, conversely, displays inefficient discrimination of the wobble pair.

Table S3. Discrimination of mismatched RNA targets by **X2/Y2/Z2** and reference strands.[a]

			RNA: 3′-CAC T**B**T ACG			
			T_m [°C]	ΔT_m [°C]		
ON	**Sequence**	**B** =	A	C	G	U
D1	5′-GTG ATA TGC		**26.5**	<−16.5	−4.5	<−16.5
X2	5′-GTG A**X**A TGC		**32.0**	−13.0	−10.0	−11.0
Y2[b]	5′-GTG A**Y**A TGC		**31.0**	−17.5	−3.5	−9.5
Z2[b]	5′-GTG A**Z**A TGC		**37.0**	−12.0	−9.0	−13.0

[a] For conditions of thermal denaturation experiments, see Table 1. T_m's of fully matched duplexes are shown in bold. ΔT_m = change in T_m relative to fully matched DNA:RNA duplex. [b] Data from reference S1 are included to facilitate comparison.

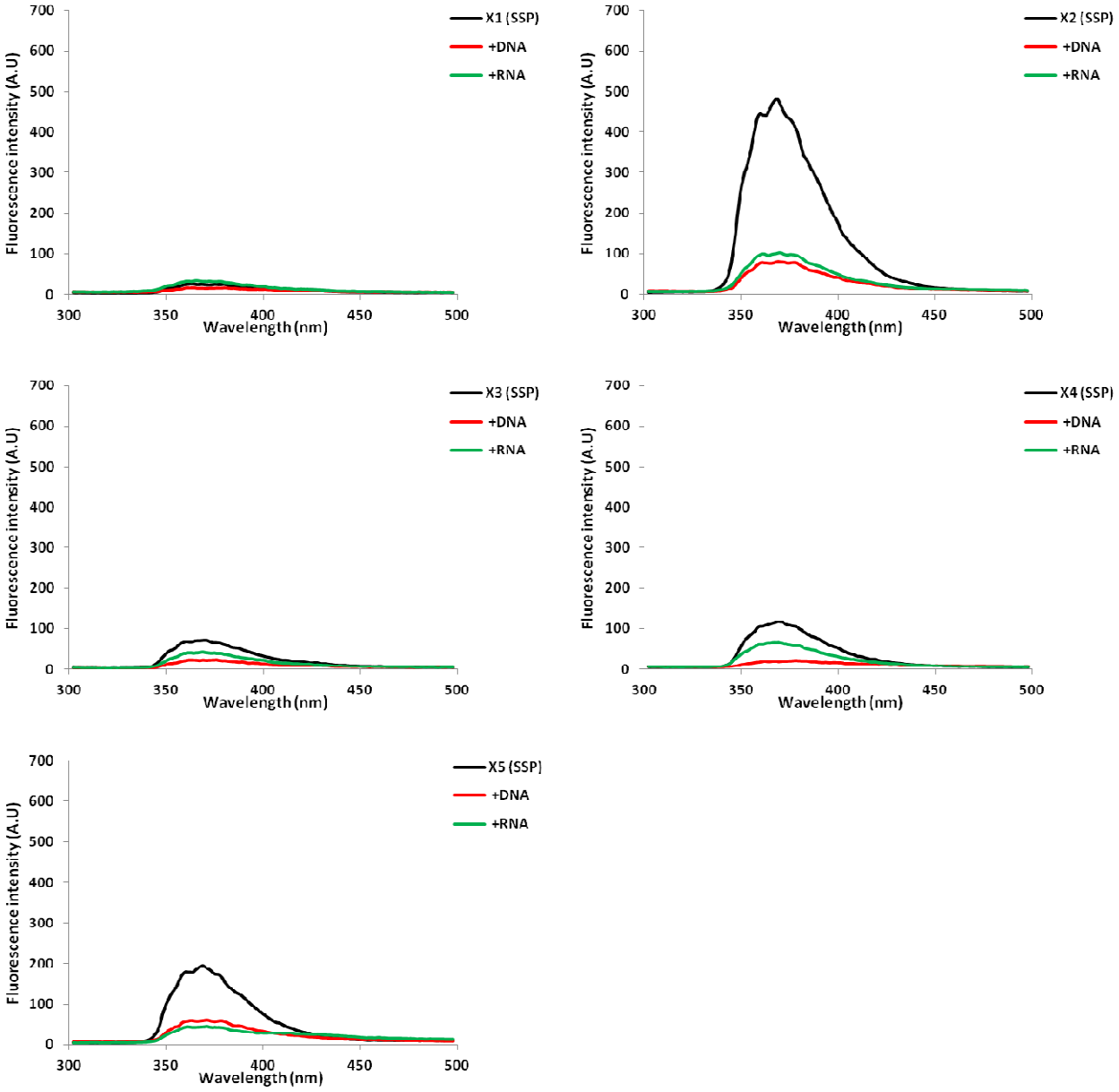


Figure S2. Steady-state fluorescence emission spectra of **X**-modified ONs and their corresponding duplexes with DNA/RNA targets. Spectra were recorded at T = 5 °C using λ_{ex} = 260 nm and each strand at 1.0 μM concentration in T_m buffer. SSP = single-stranded probe.

*Enthalpy/Entropy trends of double-stranded probes with interstrand zippers of **X/Y/Z**-monomers.* As is expected for the formation of nucleic acid duplexes, highly favorable enthalpy and highly unfavorable entropy contributions are observed for unmodified and modified duplexes. The *relative* stabilization brought about upon incorporation of triphenylene-functionalized monomer **X** into ONs is generally enthalpic in nature, although being entropic in nature for **X3** (note the negative values in the first two $\Delta\Delta H$ columns, Table S4, and the positive values in the first two $\Delta(T_{293}\Delta S)$ columns, Table S5). In contrast, the relative stabilization provided by pyrene-functionalized monomer **Y** is entropic in nature, whereas the stabilization conferred by coronene-functionalized monomer **Z** is either enthalpic or entropic in nature, depending on the sequence context. The **X/Y/Z**-monomers confer an unfavorable enthalpic contribution in essentially all of the probe duplexes (note the positive values in the third $\Delta\Delta H$ column, Table S4). As a consequence, the thermodynamic driving force for dsDNA-recognition by **B2**:**B5** probes is due to a major decrease in enthalpy (compare ΔH_{rec} and $-T^{293}\Delta S_{rec}$ values for **B2**:**B5** probes, Tables S4 and S5).

Table S4. Enthalpy change upon duplex formation (ΔH) and probe recognition of isosequential dsDNA target **D1**:**D2** (ΔH_{rec}).[a]

ON	ZP	Sequence	ΔH [$\Delta\Delta H$] (kJ/mol) upper ON vs cDNA	lower ON vs cDNA	probe duplex	ΔH_{rec} (kJ/mol)
X1 **X5**	**+4**	5'-G**X**G ATA TGC 3'-CAC TA**X** ACG	-305±2 [+1]	-359±3 [-53]	-274±2 [+32]	-84
X1 **X4**	**+2**	5'-G**X**G ATA TGC 3'-CAC **X**AT ACG	-305±2 [+1]	-319±3 [-13]	-262±1 [+44]	-56
X2 **X5**	**+1**	5'-GTG A**X**A TGC 3'-CAC TA**X** ACG	**-360±2 [-54]**	**-359±3 [-53]**	**-266±2 [+40]**	**-147**
X2 **X4**	**-1**	5'-GTG A**X**A TGC 3'-CAC **X**AT ACG	-360±2 [-54]	-319±3 [-13]	-284±4 [+22]	-89
X3 **X5**	**-1**	5'-GTG ATA **X**GC 3'-CAC TA**X** ACG	-278±3 [+28]	-359±3 [-53]	-225±7 [+81]	-106
X3 **X4**	**-3**	5'-GTG ATA **X**GC 3'-CAC **X**AT ACG	-278±3 [+28]	-319±3 [-13]	-306±5 [±0]	+15
[b]**Y1** **Y5**	**+4**	5'-G**Y**G ATA TGC 3'-CAC TA**Y** ACG	-281±4 [+25]	-299±2 [+7]	-279±1 [+27]	+5
[b]**Y1** **Y4**	**+2**	5'-G**Y**G ATA TGC 3'-CAC **Y**AT ACG	-281±4 [+25]	-300±7 [+6]	-248±8 [+58]	-27
[b]**Y2** **Y5**	**+1**	5'-GTG A**Y**A TGC 3'-CAC TA**Y** ACG	**-305±1 [+1]**	**-299±2 [+7]**	**-244±3 [+62]**	**-54**
[b]**Y2** **Y4**	**-1**	5'-GTG A**Y**A TGC 3'-CAC **Y**AT ACG	-305±1 [+1]	-300±7 [+6]	-296±6 [+10]	-3
[b]**Y3** **Y5**	**-1**	5'-GTG ATA **Y**GC 3'-CAC TA**Y** ACG	-270±7 [+36]	-299±2 [+7]	-280±5 [+26]	+17
[b]**Y3** **Y4**	**-3**	5'-GTG ATA **Y**GC 3'-CAC **Y**AT ACG	-270±7 [+36]	-300±7 [+6]	-309±4 [-3]	+45
Z1 **Z5**	**+4**	5'-G**Z**G ATA TGC 3'-CAC TA**Z** ACG	-288±6 [+18]	-317±7 [-11]	-198±5 [+108]	-101
Z1 **Z4**	**+2**	5'-G**Z**G ATA TGC 3'-CAC **Z**AT ACG	-288±6 [+18]	-326±16 [-20]	-218±5 [+88]	-90
Z2 **Z5**	**+1**	5'-GTG A**Z**A TGC 3'-CAC TA**Z** ACG	**-332±2 [-26]**	**-317±7 [-11]**	**-252±7 [+54]**	**-91**
Z2 **Z4**	**-1**	5'-GTG A**Z**A TGC 3'-CAC **Z**AT ACG	-332±2 [-26]	-326±16 [-20]	-235±8 [+71]	-117
Z3 **Z5**	**-1**	5'-GTG ATA **Z**GC 3'-CAC TA**Z** ACG	-297±4 [+9]	-317±7 [-11]	-166±8 [+140]	-142
Z3 **Z4**	**-3**	5'-GTG ATA **Z**GC 3'-CAC **Z**AT ACG	-297±4 [+9]	-326±16 [-20]	-250±6 [+56]	-67

[a] $\Delta\Delta H$ is measured relative to ΔH for **D1**:**D2 =** -306 kJ/mol. ΔH_{rec} = ΔH (upper ON vs cDNA) + ΔH (lower ON vs cDNA) - ΔH (probe duplex) - ΔH (dsDNA target). "±" denotes standard deviation. [b] Data from reference S2 shown to facilitate comparison.

Table S5. Entropy change at 293K upon duplex formation ($-T^{293}\Delta S$) and probe recognition of isosequential dsDNA target **D1**:**D2** ($-T^{293}\Delta S_{rec}$).[a]

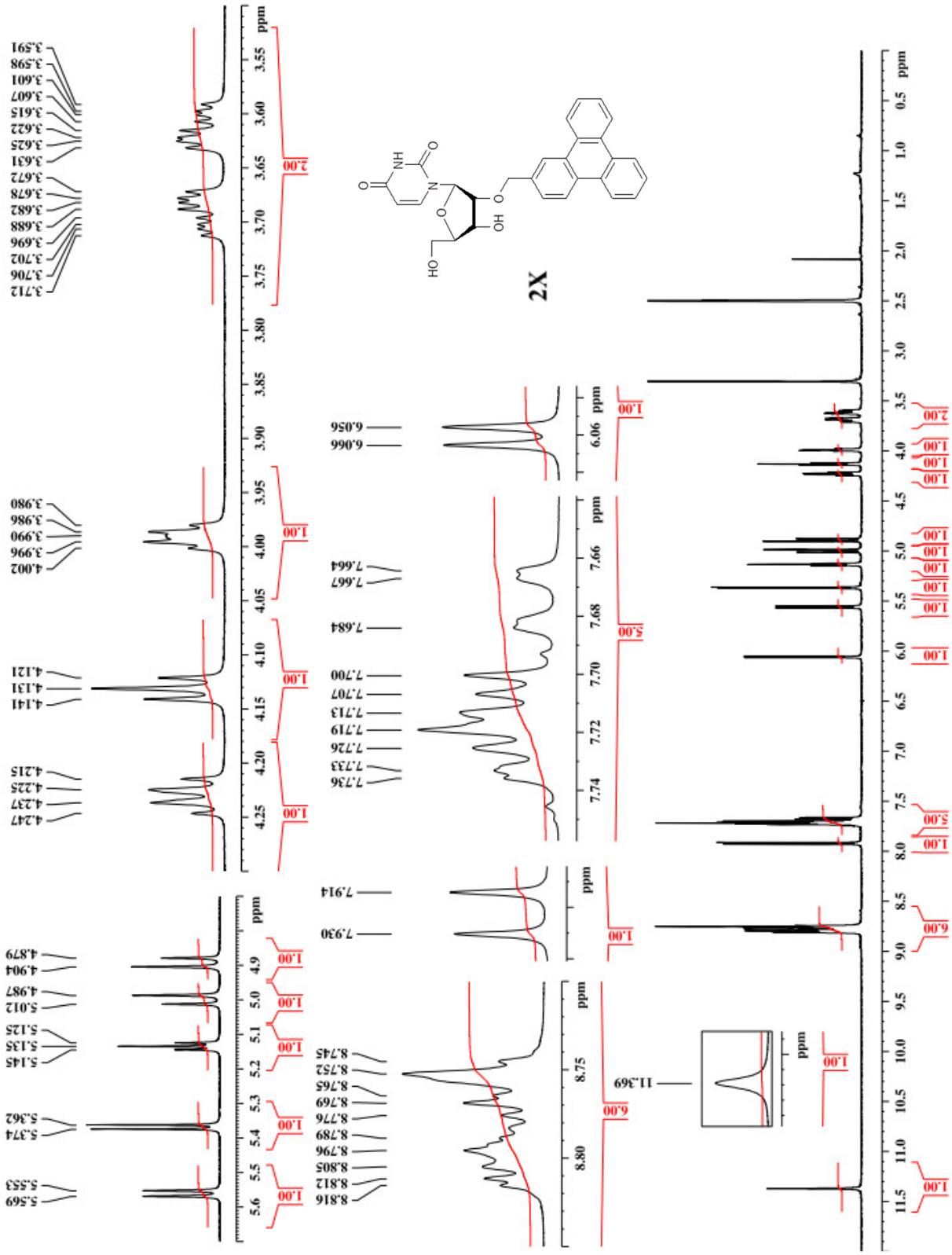
ON	ZP	Sequence	$-T^{293}\Delta S$ [$\Delta(T^{293}\Delta S)$] (kJ/mol)			$-T^{293}\Delta S_{rec}$ (kJ/mol)
			upper ON vs cDNA	lower ON vs cDNA	probe duplex	
X1 **X5**	**+4**	5'-G**X**G ATA TGC 3'-CAC TA**X** ACG	259±2 [-6]	301±3 [+36]	220±2 [-45]	+75
X1 **X4**	**+2**	5'-G**X**G ATA TGC 3'-CAC **X**AT ACG	259±2 [-6]	274±3 [+9]	217±1 [-48]	+51
X2 **X5**	**+1**	5'-GTG A**X**A TGC 3'-CAC TA**X** ACG	**301±2 [+36]**	**301±3 [+36]**	**225±2 [-40]**	**+112**
X2 **X4**	**-1**	5'-GTG A**X**A TGC 3'-CAC **X**AT ACG	301±2 [+36]	274±3 [+9]	232±3 [-33]	+78
X3 **X5**	**-1**	5'-GTG ATA **X**GC 3'-CAC TA**X** ACG	230±3 [-35]	301±3 [+36]	174±6 [-91]	+92
X3 **X4**	**-3**	5'-GTG ATA **X**GC 3'-CAC **X**AT ACG	230±3 [-35]	274±3 [+9]	254±5 [-11]	-15
[b]**Y1** **Y5**	**+4**	5'-G**Y**G ATA TGC 3'-CAC TA**Y** ACG	235±3 [-30]	247±6 [-18]	224±1 [-41]	-7
[b]**Y1** **Y4**	**+2**	5'-G**Y**G ATA TGC 3'-CAC **Y**AT ACG	235±3 [-30]	254±6 [-11]	203±8 [-62]	+21
[b]**Y2** **Y5**	**+1**	5'-GTG A**Y**A TGC 3'-CAC TA**Y** ACG	**251±1 [-14]**	**247±6 [-18]**	**204±3 [-61]**	**+29**
[b]**Y2** **Y4**	**-1**	5'-GTG A**Y**A TGC 3'-CAC **Y**AT ACG	251±1 [-14]	254±6 [-11]	245±5 [-20]	-5
[b]**Y3** **Y5**	**-1**	5'-GTG ATA **Y**GC 3'-CAC TA**Y** ACG	220±7 [-45]	247±6 [-18]	225±5 [-40]	-23
[b]**Y3** **Y4**	**-3**	5'-GTG ATA **Y**GC 3'-CAC **Y**AT ACG	220±7 [-45]	254±6 [-11]	256±3 [-9]	-47
Z1 **Z5**	**+4**	5'-G**Z**G ATA TGC 3'-CAC TA**Z** ACG	243±5 [-22]	258±6 [-7]	146±4 [-119]	+90
Z1 **Z4**	**+2**	5'-G**Z**G ATA TGC 3'-CAC **Z**AT ACG	243±5 [-22]	275±15 [+10]	167±4 [-98]	+86
Z2 **Z5**	**+1**	5'-GTG A**Z**A TGC 3'-CAC TA**Z** ACG	**271±2 [+6]**	**258±6 [-7]**	**201±7 [-64]**	**+63**
Z2 **Z4**	**-1**	5'-GTG A**Z**A TGC 3'-CAC **Z**AT ACG	271±2 [+6]	275±15 [+10]	178±7 [-87]	+103
Z3 **Z5**	**-1**	5'-GTG ATA **Z**GC 3'-CAC TA**Z** ACG	244±4 [-21]	258±6 [-7]	116±8 [-149]	+121
Z3 **Z4**	**-3**	5'-GTG ATA **Z**GC 3'-CAC **Z**AT ACG	244±4 [-21]	275±15 [+10]	196±6 [-69]	+58

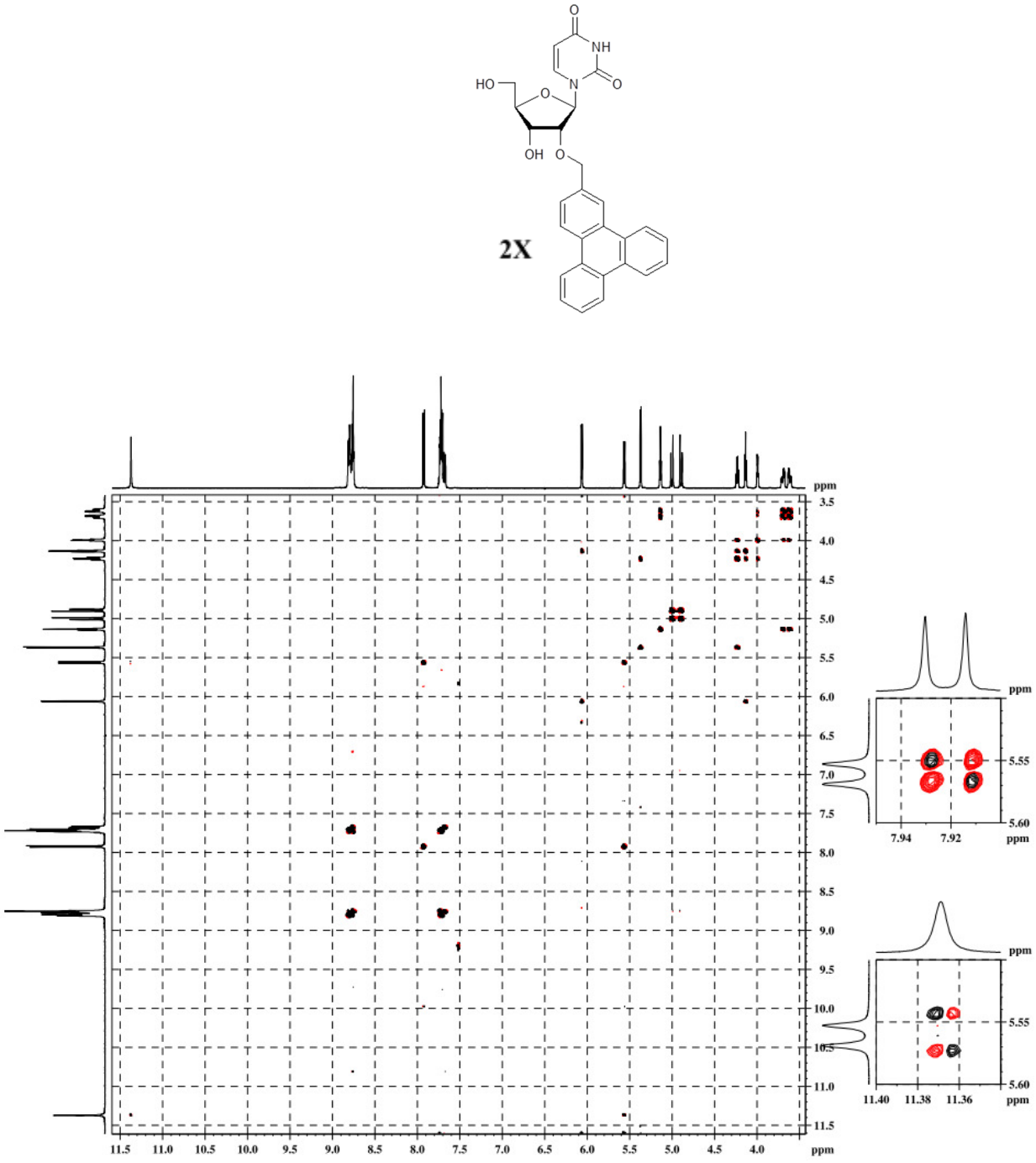
[a] $\Delta(T^{293}\Delta S)$ is measured relative to $-T^{293}\Delta S$ for **D1**:**D2 =** 265 kJ/mol. $-T^{293}\Delta S_{rec} = T^{293}\Delta S$ (upper strand vs cDNA) + $T^{293}\Delta S$ (lower strand vs cDNA) - $T^{293}\Delta S$ (probe duplex) - $T^{293}\Delta S$ (dsDNA target). "±" denotes standard deviation. [b] Data from reference S2 shown to facilitate comparison.

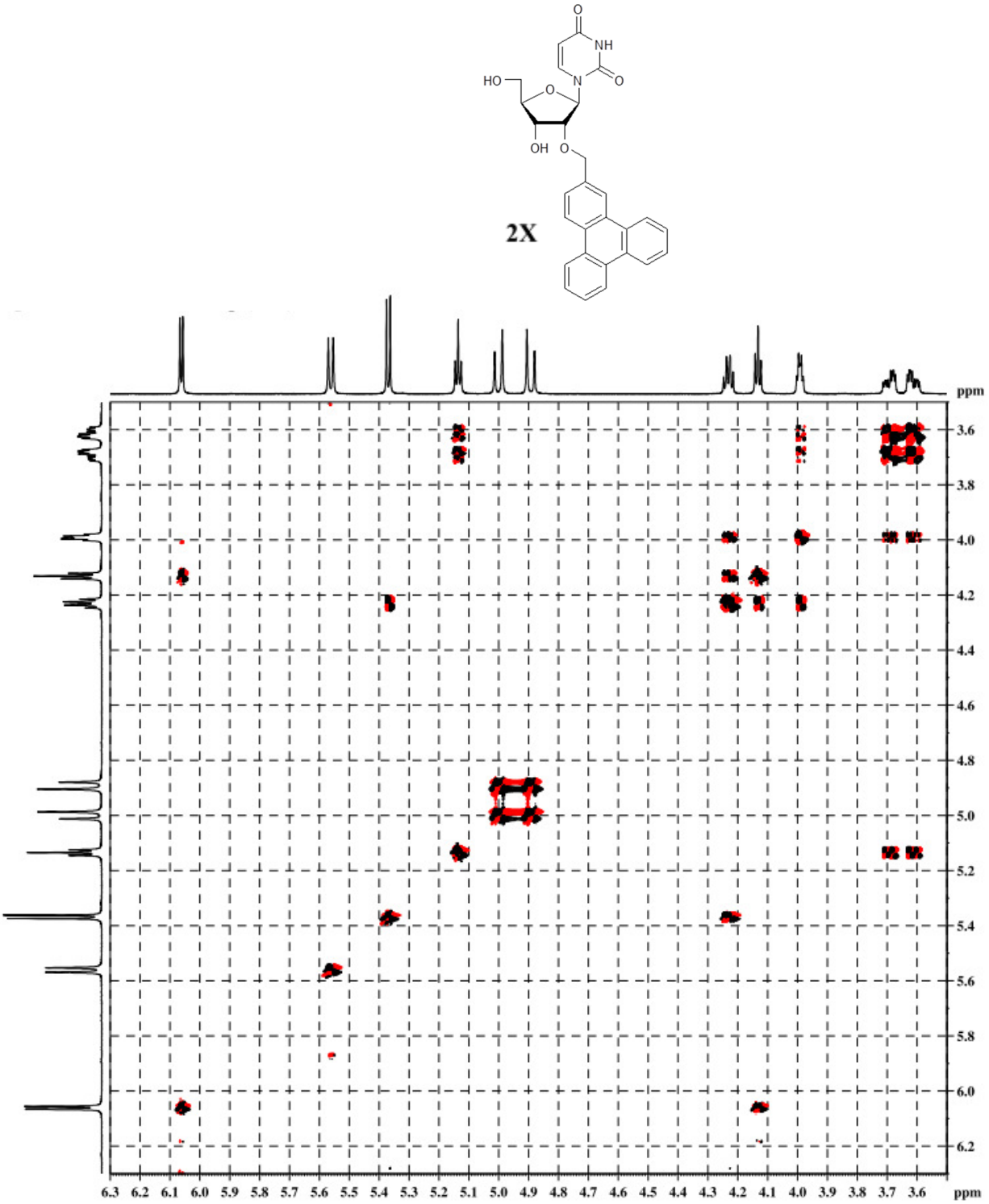
References

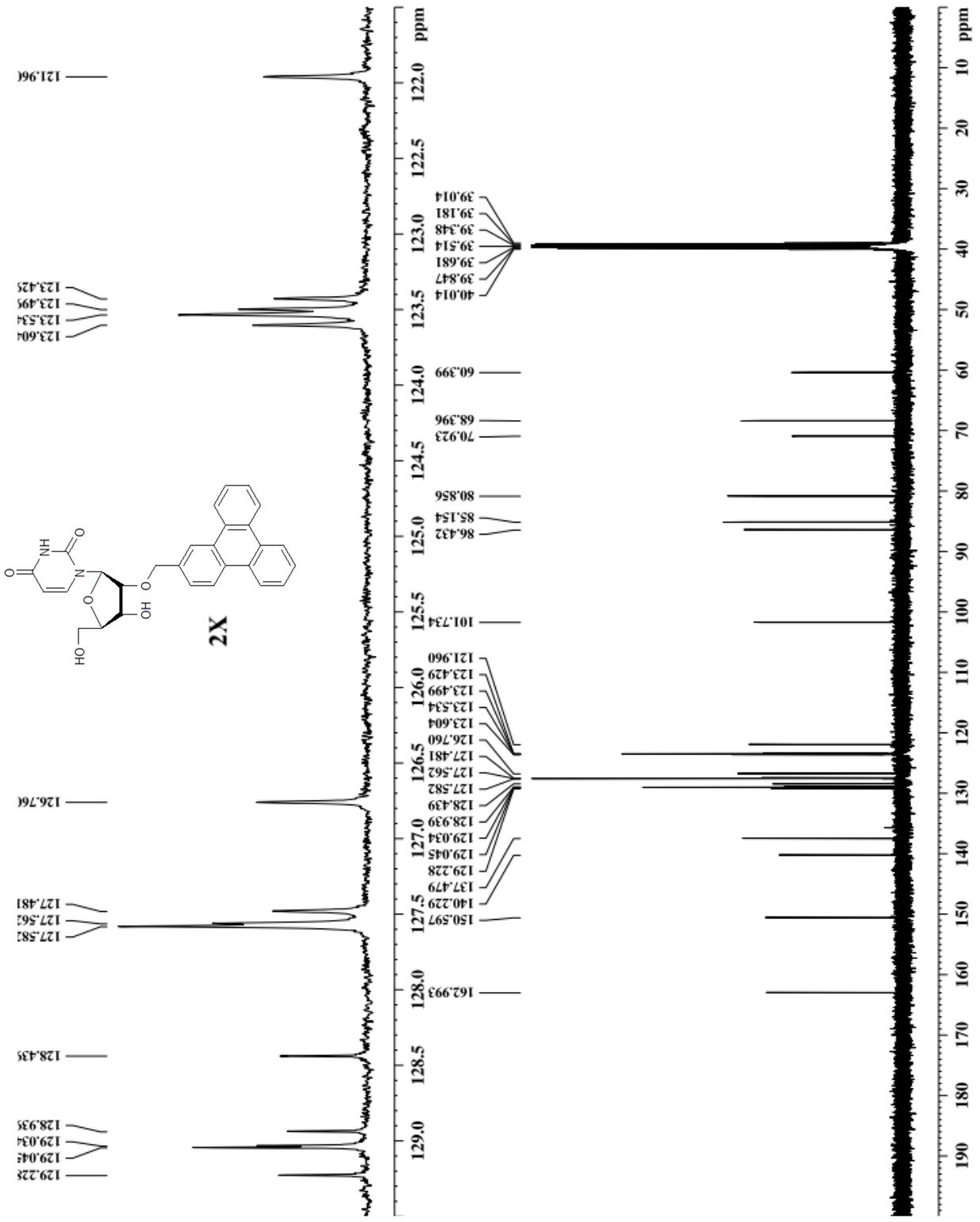
S1) S. Karmakar, B. A. Anderson, R. L. Rathje, S. Andersen, T. B. Jensen, P. E. Nielsen and P. J. Hrdlicka, *J. Org. Chem.*, 2011, **76**, 7119-7131.

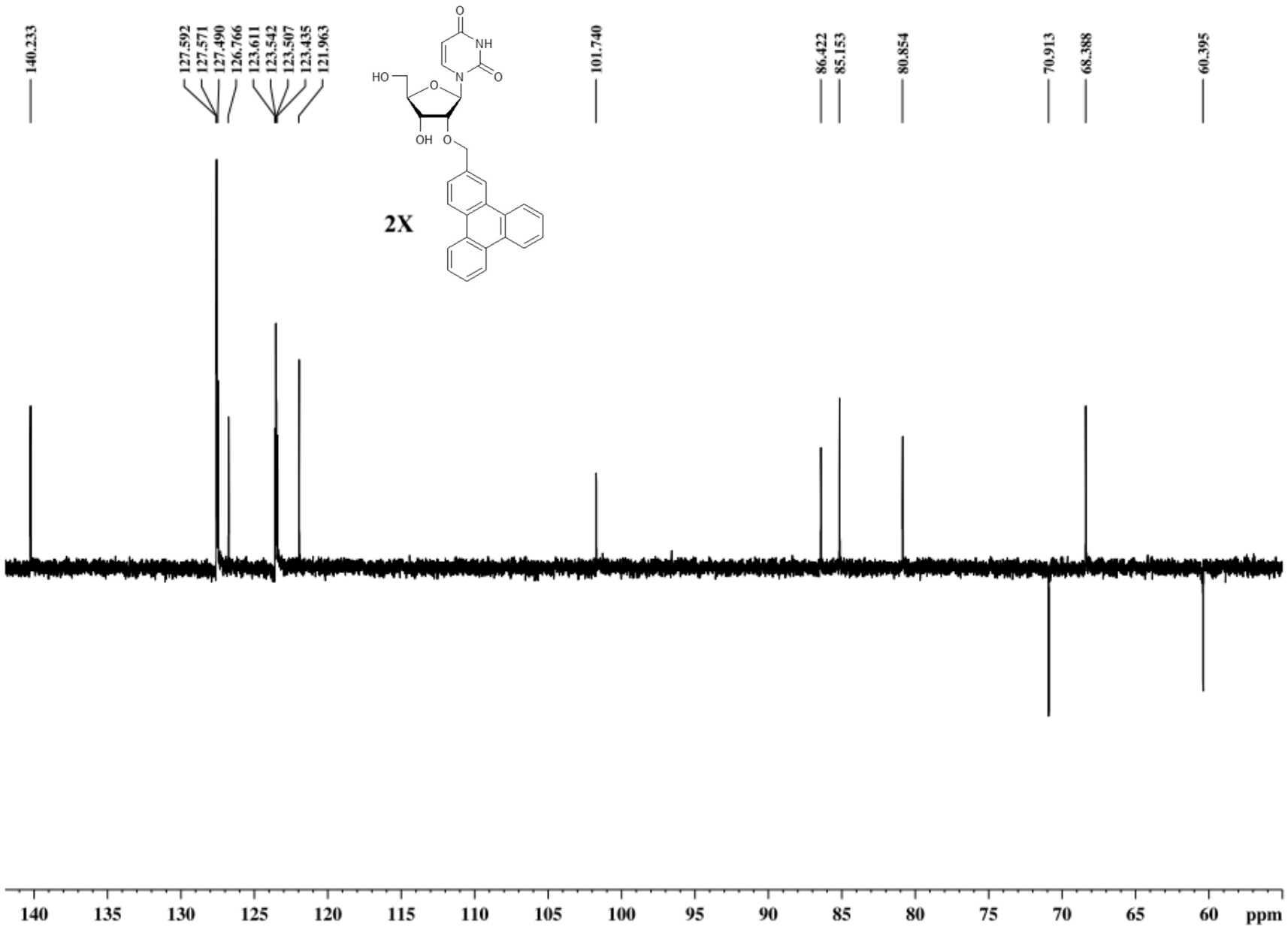
S2) S. Karmakar, A. S. Madsen, D. C. Guenther, B. C. Gibbons and P. J. Hrdlicka, *Org. Biomol. Chem.*, 2014, **12**, 7758-7773.

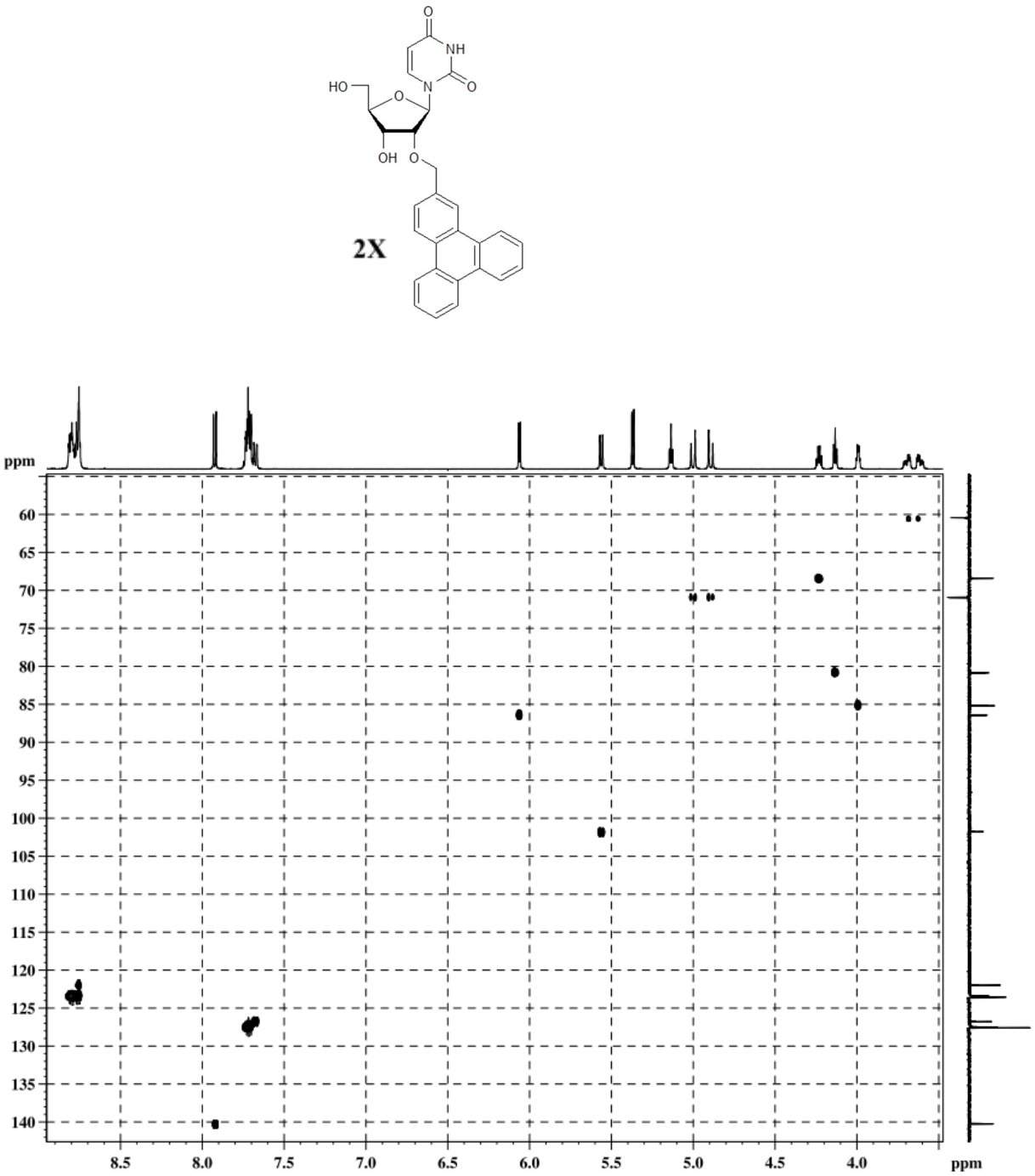

2X

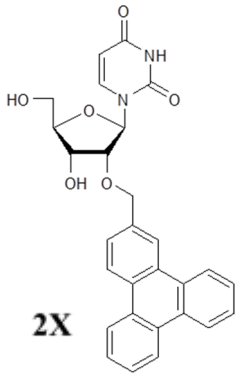

O
NH
HO
N
O
O
OH
O
2X
ppm
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5
9.0
9.5
10.0
10.5
11.0
11.5
11.5
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
ppm
ppm
5.55
5.60
7.94
7.92
ppm
ppm
5.55
5.60
11.40
11.38
11.36
ppm


O
NH
HO
N
O
O
OH
O
2X
ppm
3.6
3.8
4.0
4.2
4.4
4.6
4.8
5.0
5.2
5.4
5.6
5.8
6.0
6.2
6.3 6.2 6.1 6.0 5.9 5.8 5.7 5.6 5.5 5.4 5.3 5.2 5.1 5.0 4.9 4.8 4.7 4.6 4.5 4.4 4.3 4.2 4.1 4.0 3.9 3.8 3.7 3.6
ppm

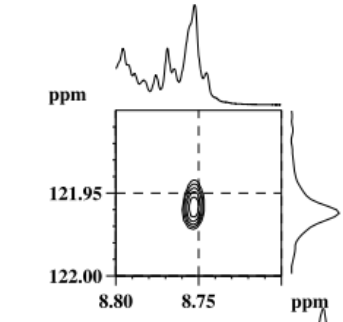

2X
ppm

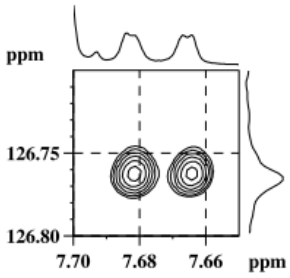

140.233
127.592
127.571
127.490
126.766
123.611
123.542
123.507
123.435
121.963
101.740
86.422
85.153
80.854
70.913
68.388
60.395
HO
O
NH
N
O
O
OH
O
2X
140
135
130
125
120
115
110
105
100
95
90
85
80
75
70
65
60
ppm

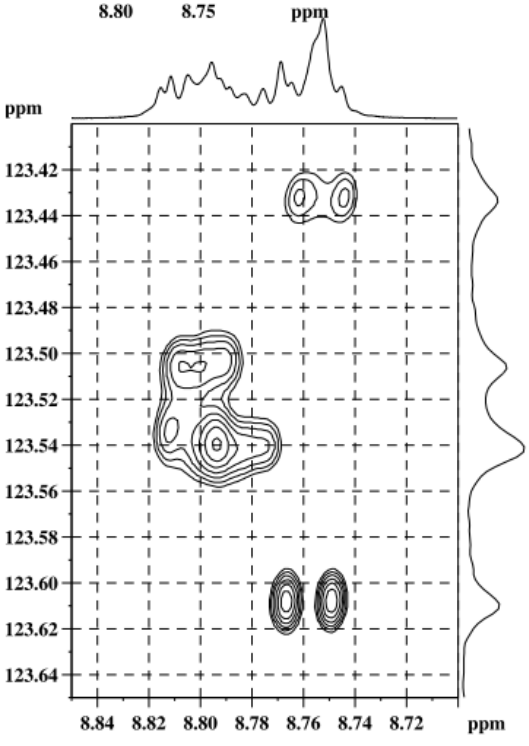

O
NH
O
HO
N
O
OH
O
2X
ppm
60
65
70
75
80
85
90
95
100
105
110
115
120
125
130
135
140
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
ppm

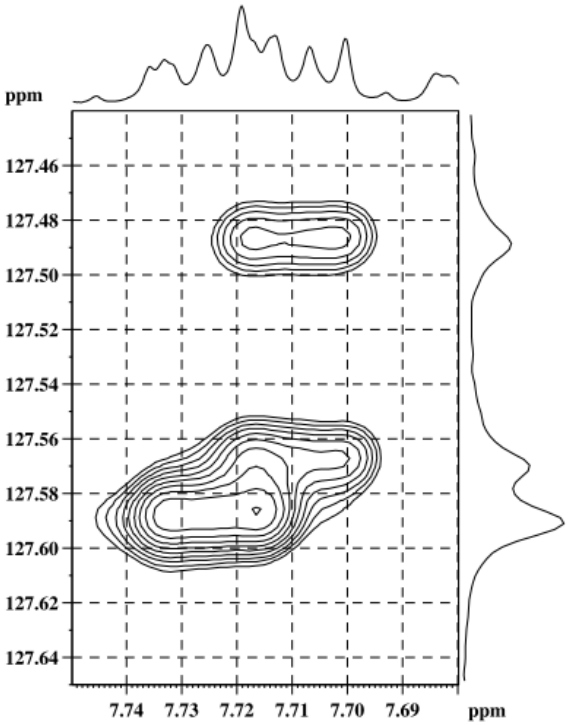

O
NH
O
N
HO
O
OH
O

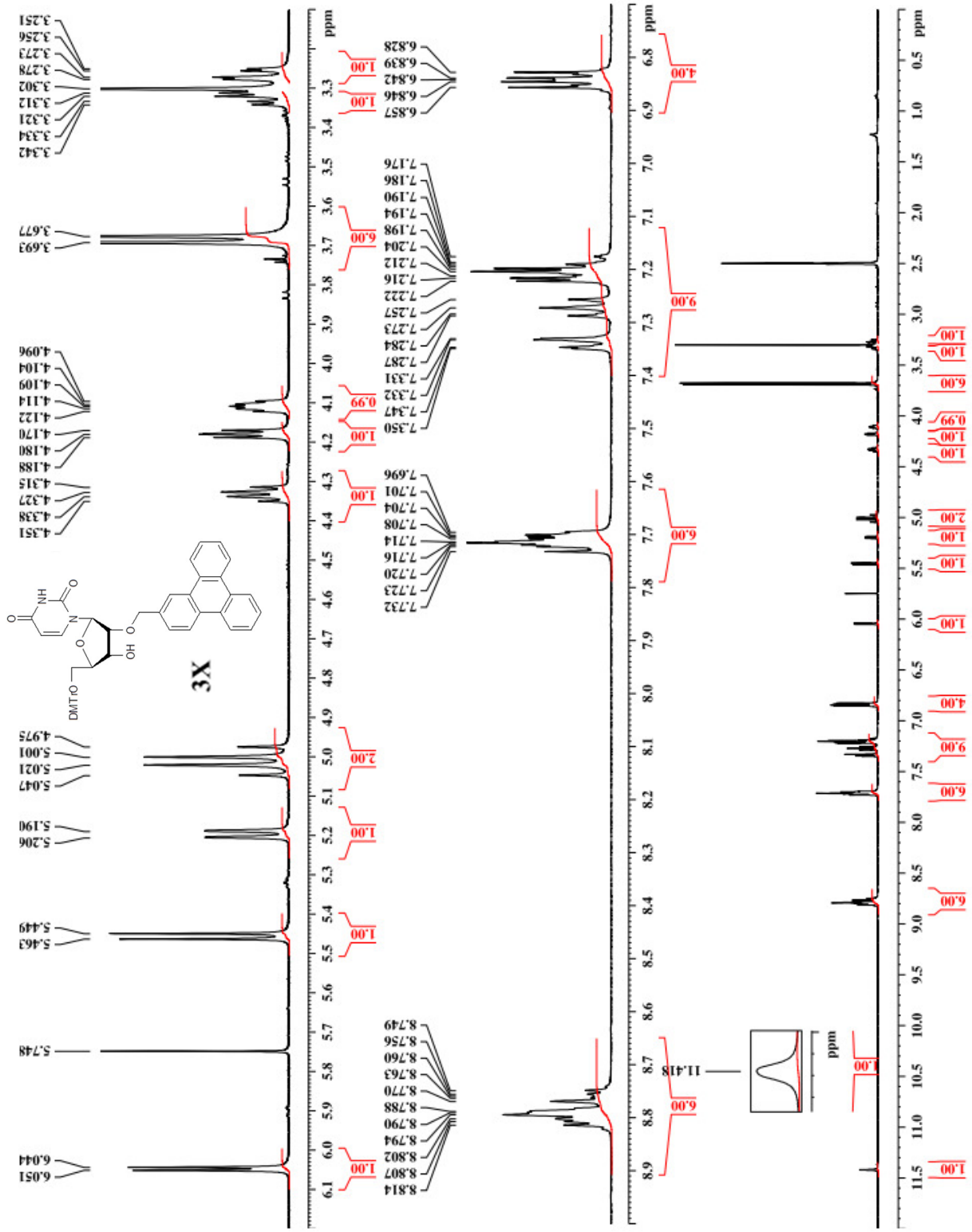
2X

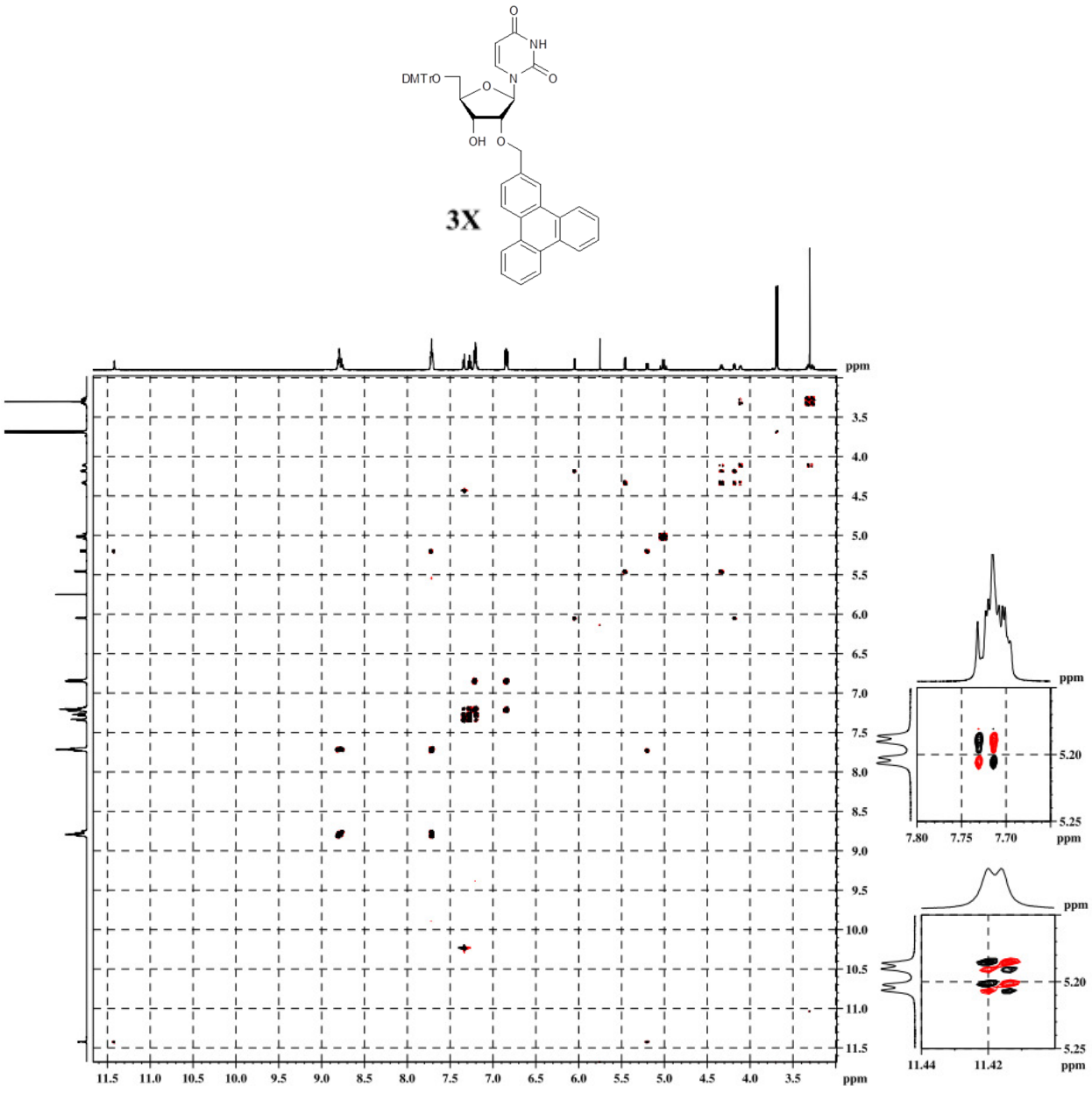

ppm
121.95
122.00
8.80
8.75
ppm

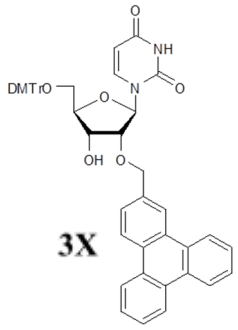

ppm
126.75
126.80
7.70
7.68
7.66
ppm


ppm
123.42
123.44
123.46
123.48
123.50
123.52
123.54
123.56
123.58
123.60
123.62
123.64
8.84
8.82
8.80
8.78
8.76
8.74
8.72
ppm

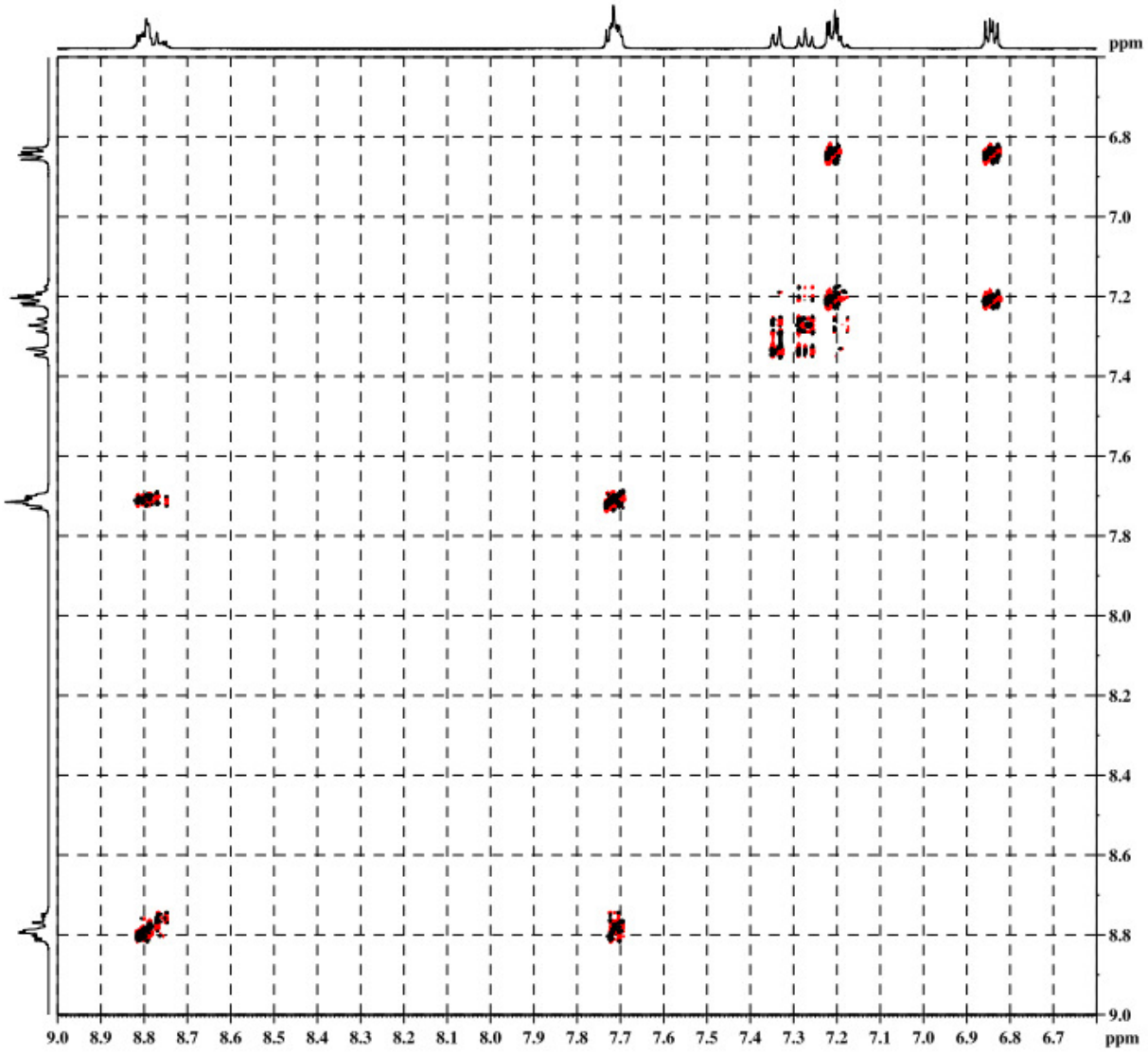

ppm
127.46
127.48
127.50
127.52
127.54
127.56
127.58
127.60
127.62
127.64
7.74
7.73
7.72
7.71
7.70
7.69
ppm

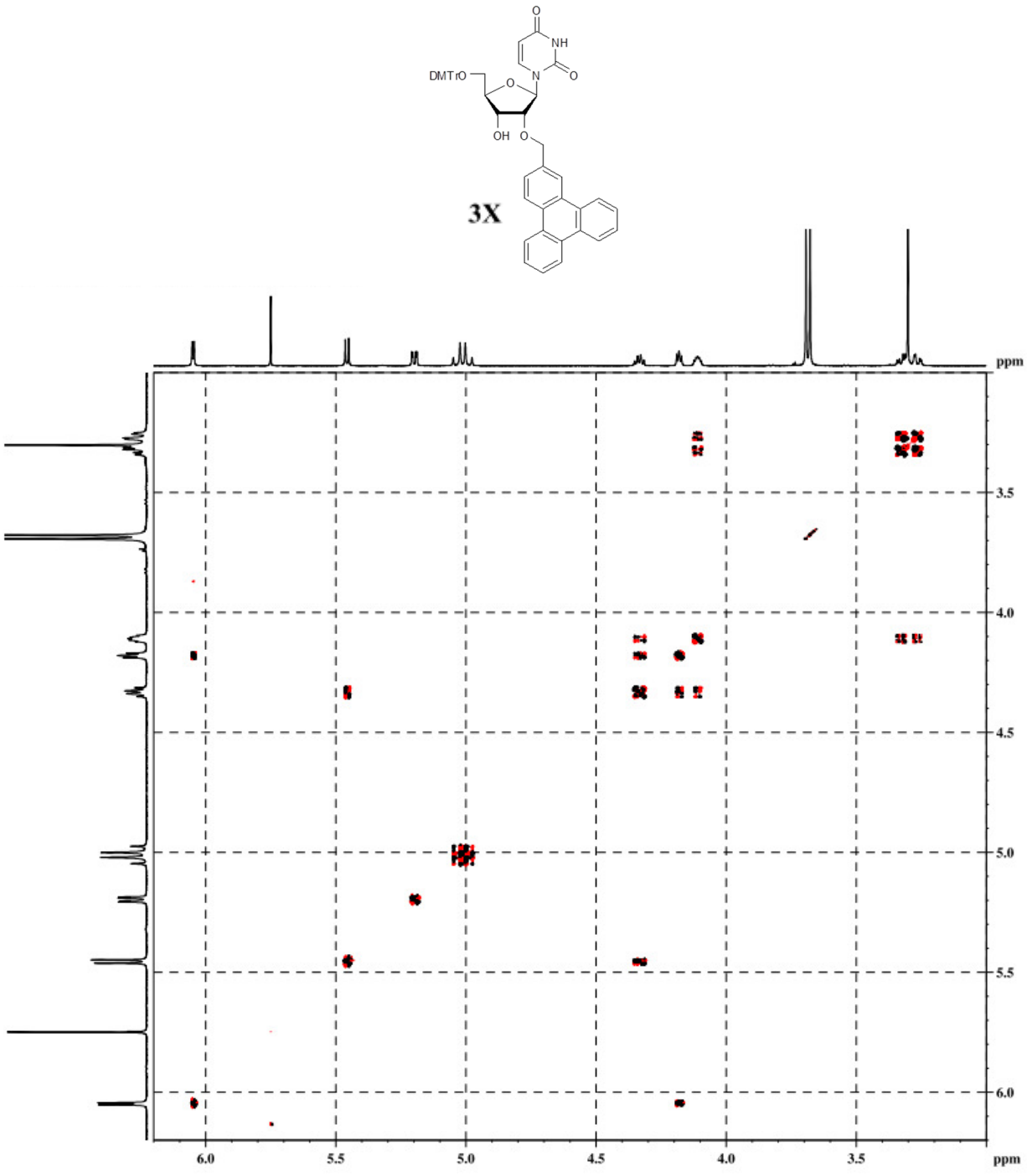

3X
DMTrO
HO
NH
O
N
ppm

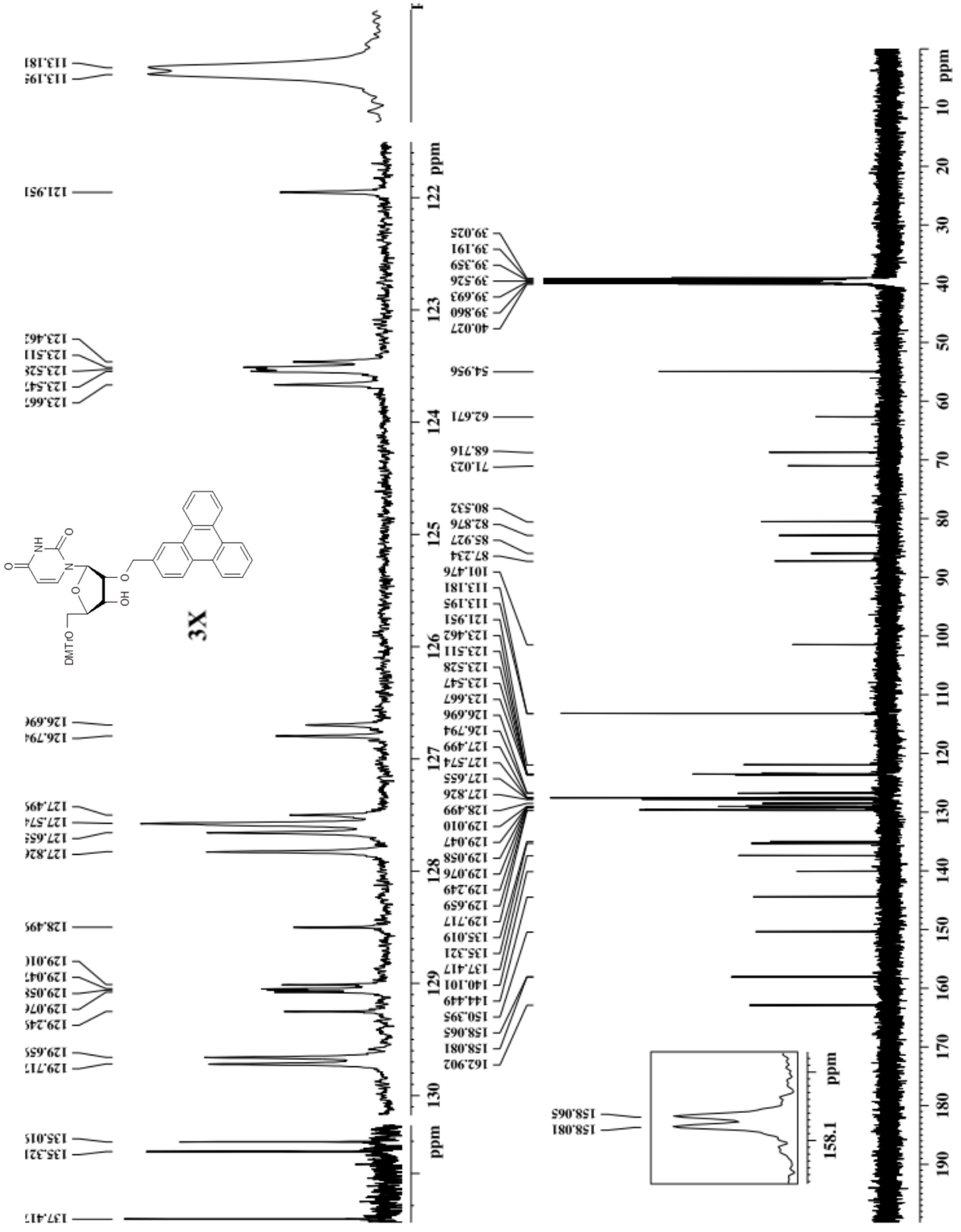

O
NH
DMTrO
N
O
O
OH
O
3X
ppm
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5
9.0
9.5
10.0
10.5
11.0
11.5
11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5
ppm
ppm
5.20
5.25
7.80 7.75 7.70
ppm
ppm
5.20
5.25
11.44 11.42
ppm

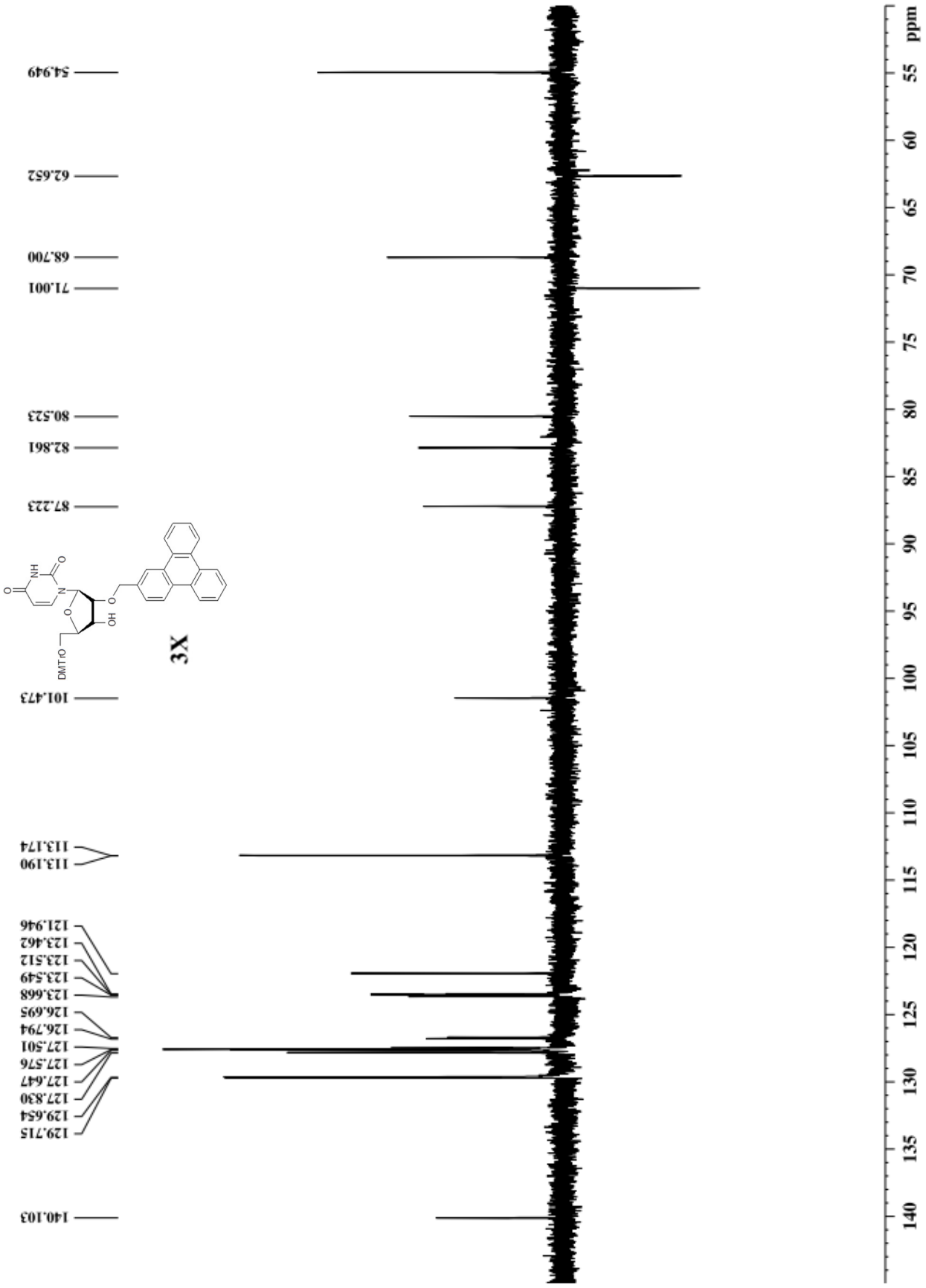

O
NH
DMTrO
N
O
O
OH
O

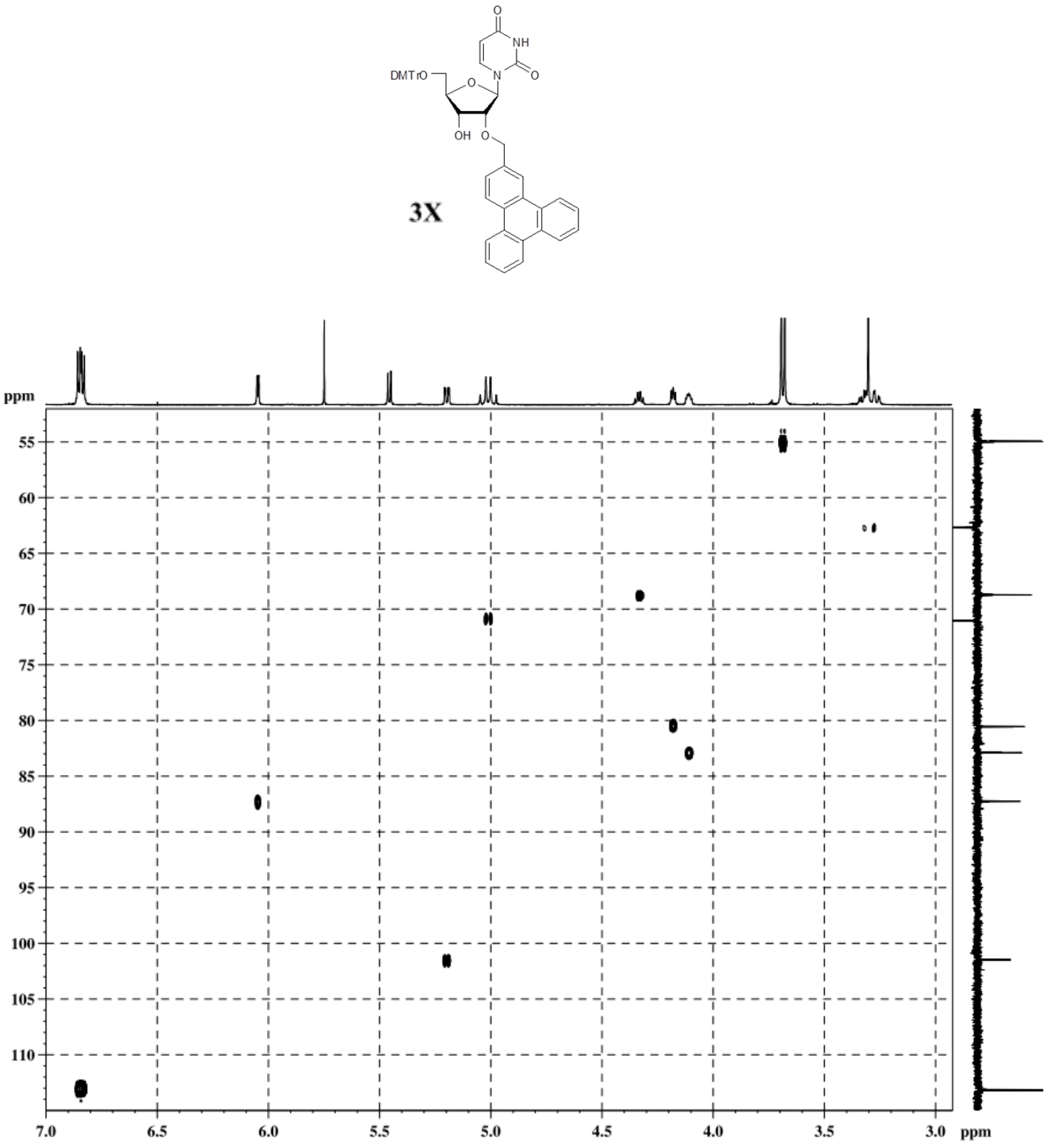
3X

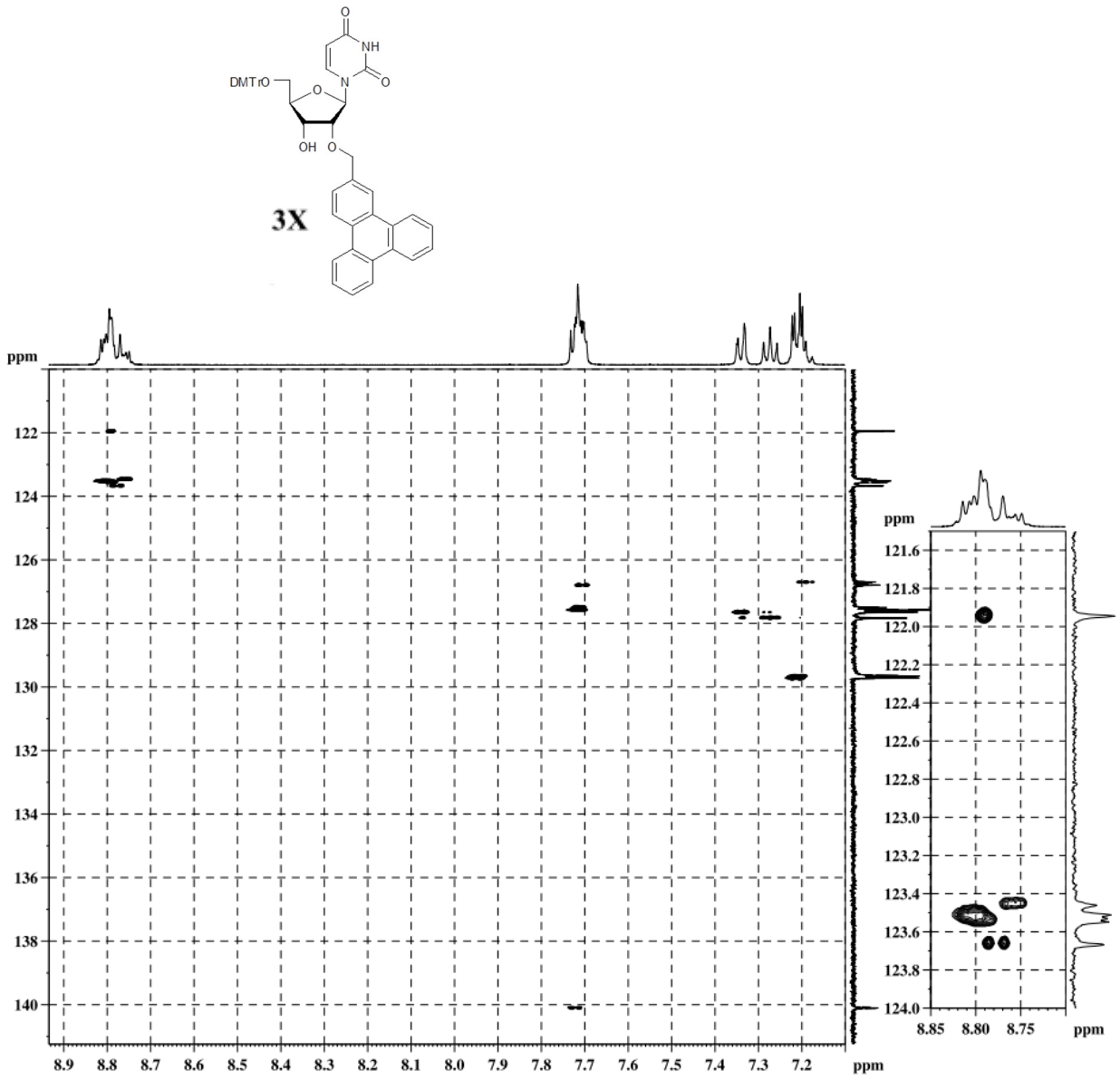

ppm
6.8
7.0
7.2
7.4
7.6
7.8
8.0
8.2
8.4
8.6
8.8
9.0
9.0 8.9 8.8 8.7 8.6 8.5 8.4 8.3 8.2 8.1 8.0 7.9 7.8 7.7 7.6 7.5 7.4 7.3 7.2 7.1 7.0 6.9 6.8 6.7
ppm

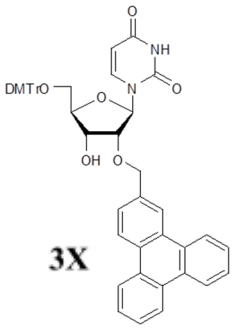

DMTrO
O
NH
N
O
O
OH
O
3X
ppm
3.5
4.0
4.5
5.0
5.5
6.0
6.0
5.5
5.0
4.5
4.0
3.5
ppm


3X
DMTrO
OH
ppm
158.1

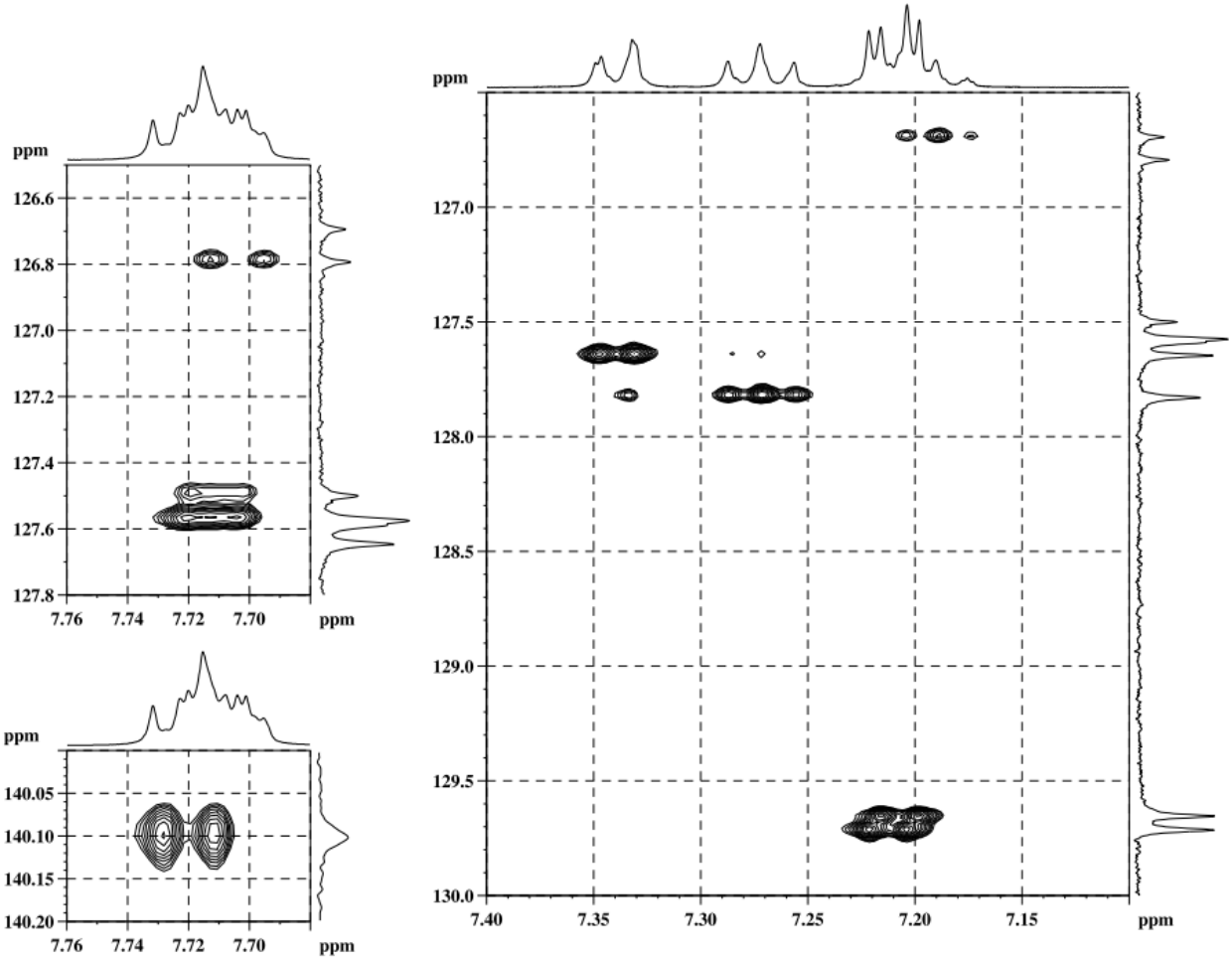

54.949
62.652
68.700
71.001
80.523
82.861
87.223
101.473
113.174
113.190
121.946
123.462
123.512
123.549
123.668
126.695
126.794
127.501
127.576
127.647
127.830
129.654
129.715
140.103
3X
DMTrO
OH
NH
O
N
ppm
55
60
65
70
75
80
85
90
95
100
105
110
115
120
125
130
135
140


O
NH
DMTrO
N
O
O
OH
O
3X
ppm
55
60
65
70
75
80
85
90
95
100
105
110
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
ppm


DMTrO
O
NH
N
O
OH
O
3X
ppm
122
124
126
128
130
132
134
136
138
140
8.9 8.8 8.7 8.6 8.5 8.4 8.3 8.2 8.1 8.0 7.9 7.8 7.7 7.6 7.5 7.4 7.3 7.2
ppm
ppm
121.6
121.8
122.0
122.2
122.4
122.6
122.8
123.0
123.2
123.4
123.6
123.8
124.0
8.85 8.80 8.75
ppm


O
NH
DMTrO
N
O
O
OH
O

3X


ppm
126.6
126.8
127.0
127.2
127.4
127.6
127.8
7.76
7.74
7.72
7.70
ppm
ppm
140.05
140.10
140.15
140.20
7.76
7.74
7.72
7.70
ppm
ppm
127.0
127.5
128.0
128.5
129.0
129.5
130.0
7.40
7.35
7.30
7.25
7.20
7.15
ppm

4X
149.460
148.910
149.460
148.910
149.5
149.0
148.5 ppm
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
ppm