Supplementary Materials for

Correlation of Local Effects of DNA Sequence and Position of Beta-Alanine Inserts with Polyamide-DNA Complex Binding Affinities and Kinetics

Shuo Wang,[†] Rupesh Nanjunda,[†] Karl Aston,[§] James K. Bashkin,[§*] W. David Wilson[†*]

[†] Department of Chemistry, Georgia State University, Atlanta, GA 30303, USA

[§] Department of Chemistry & Biochemistry, Center for Nanoscience, University of Missouri-St. Louis, St. Louis, MO 63121, USA

[*] Correspondence may be addressed to either of these authors:

Tel: +1.404.413.5503; Fax: +1.404.413.5505. Email: wdw@gsu.edu

Tel: +1.314.516.7352; Fax: +1.314.516.5342. Email: bashkinj@umsl.edu

Materials and Methods

Chemical Characterization of KA1055, dImImPyPyγPyImβPyβDp

HPLC/Mass Spectrometry (HRMS) (MS^n ESI^+) calculated for $C_{54}H_{70}N_{21}O_{10}$ MH^+ 1172.56145; found 1172.55845. Elemental analysis calculated for KA1055·3TFA·7H_2O, $C_{60}H_{86}N_{21}O_{23}F_9$: C, 43.93%; H, 5.28%; F, 10.42; N, 17.93%; found: C, 44.20%; H, 5.21%; F, 10.00; N, 17.92%.

NMR conditions and descriptions: 500 or 600 MHz referenced to residual solvent, DMSO-d^6 or MeOH-d^4, temp (500 MHz = 27 °C, 600 MHz = 25°C), ex. indicate protons exchangeable in MeOH-d^4 or D_2O.

^{1}H NMR (600 MHz, DMSO-d^6) 10.33 (s, 1H, ex.), 10.32 (s, 1H, ex.), 9.93 (s, 1H, ex.), 9.88 (s, 1H, ex.), 9.83 (s, 1H, ex.), 9.75 (s, 1H, ex.), 9.23 (brs, 1H, ex.), 8.03 (m, 3H, ex.), 7.92 (t, J=5.9Hz, 1H, ex), 7.58 (s, 1H), 7.50 (s, 1H), 7.47 (d, J=0.5Hz, 1H), 7.282 (d, J=1.8Hz, 1H), 2.276 (d, J=1.8Hz, 1H), 7.18 (d, J=1.7Hz, 1H), 7.15 (d, J=1.8Hz, 1H), 7.10 (d, J=1.7Hz, 1H), 7.09 (d, J=1.2Hz, 1H), 6.91 (d, J=1.8Hz, 2H), 6.90 (d, J=1.8Hz, 2H), 6.67 (d, J=2.3Hz, 1H**),** 4.016 (s, 3H)**,** 4.014 (s, 3H), 3.94 (s, 3H), 3.85 (s, 3H), 3.82 (s, 3H), 3.81 (s, 3H), 3.78 (s, 3H), 3.52 (m, 2H), 3.36 (m, 2H), 3.21 (m, 2H), 3.10 (m, 2H), 2.99 (m, 2H), 2.73 (d, J=4.7Hz, 6H), 2.52 (t, J=6.4Hz, 2H), 2.33 (t, J=7.0Hz, 2H), 2.28 (t, J=7.1Hz, 2H), 1.79 (m, 2H), 1.73 (m, 2H).

^{1}H NMR (500 MHz, DMSO-d^6) 10.35 (s, 1H, ex.), 10.33 (s, 1H, ex.), 9.94 (s, 1H, ex.), 9.89 (s, 1H, ex.), 9.84 (s, 1H, ex.), 9.74 (s, 1H, ex.), 9.23 (brs, 1H, ex.), 8.04 (m, 3H, ex.), 7.92 (t, J=6.0Hz, 1H, ex.), 7.58 (s, 1H), 7.50 (s, 1H), 7.47 (d, J=1.0Hz, 1H), 7.284 (d, J=1.9H, 1H), 7.281 (d, J=1.6Hz, 1H), 7.19 (d, J=1.9Hz, 1H), 7.16 (d, J=1.9Hz, 1H), 7.11 (d, J=1.9Hz, 1H), 7.09 (d,

J=1.0Hz, 1H), 6.91 (d, J=1.9Hz, 2H), 6.90 (d, J=1.9Hz, 2H), 6.67 (d, J=1.9Hz, 1H**),** 4.014 (s, 3H)**,** 4.012 (s, 3H), 3.94 (s, 3H), 3.85 (s, 3H), 3.82 (s, 3H), 3.81 (s, 3H), 3.78 (s, 3H), 3.52 (m, 2H), 3.36 (m, 2H), 3.21 (m, 2H), 3.10 (m, 2H), 2.99 (m, 2H), 2.73 (d, J=4.7Hz, 6H), 2.33 (t, J=7.3Hz, 2H), 2.28 (t, J=7.3Hz, 2H), 1.78 (m, 2H), 1.72 (m, 2H).

^{1}H NMR (500 MHz, DMSO-d^6 plus D_2O) 7.57 (s, 1H), 7.48 (s, 1H), 7.45 (d, J=0.7Hz, 1H), 7.27 (d, J=1.6Hz, 1H), 7.26 (d, J=1.9Hz, 1H), 7.17 (d, J=1.9Hz, 1H), 7.13 (d, J=1.9Hz, 1H), 7.10 (d, J=1.0Hz, 1H), 7.09 (d, J=1.9Hz, 1H), 6.90 (d, J=1.9Hz, 2H), 6.89 (d, J=1.9Hz, 2H), 6.66 (d, J=1.9Hz, 1H**),** 4.003 (s, 3H)**,** 3.999 (s, 3H), 3.93 (s, 3H), 3.84 (s, 3H), 3.81 (s, 3H), 3.80 (s, 3H), 3.76 (s, 3H), 3.35 (m, 2H), 3.20 (m, 2H), 3.09 (m, 2H), 2.98 (m, 2H), 2.72 (s, 6H), 2.33 (t, J=7.0Hz, 2H), 2.28 (t, J=7.3Hz, 2H), 1.78 (m, 2H), 1.72 (m, 2H).

^{1}H NMR (500 MHz, MeOH-d^4) 7.53 (s, 1H), 7.36 (s, 1H), 7.31 (d, J=1.0Hz, 1H), 7.27 (d, J=1.9Hz, 1H), 7.15 (d, J=1.6Hz, 1H), 7.12 (d, J=1.0Hz, 1H), 7.08 (d, J=1.9Hz, 1H), 7.04 (d, J=1.9Hz, 1H), 6.92 (d, J=1.9Hz, 1H), 6.87 (d, J=1.9Hz, 1H), 6.79 (d, J=1.9Hz, 1H), 6.72 (d, J=1.9Hz, 1H), 4.08 (s, 6H), 3.97 (s, 3H), 3.91 (s, 3H), 3.87 (s, 3H), 3.85 (s, 3H), 3.82 (s, 3H), 3.62 (t, J=6.6 Hz, 2H), 3.55 (t, J=6.6Hz, 2H), 3.40 (t, J=6.6Hz, 2H), 3.25 (t, J=6.6Hz, 2H), 3.06 (m, 2H), 2.82 (s, 6H), 2.59 (t, J=6.5Hz, 2H), 2.47 (t, J=6.6Hz, 2H), 2.41 (t, J=7.0Hz, 2H), 1.98 (m, 2H), 1.86 (m, 2H).

Reverse Phase HPLC/mass spectrometry and NMR data are shown for KA1055 (Figures S1-9): Agilent 1100 HPLC; MSD ESI^+ and diode array detection; column: 70 × 4.6 mm Macherey-

Nagel Nucleodur C18 column; 5µ particles; gradient: 5-60% MeCN/water with 0.1% formic acid at 0.5 mL/min; retention time 11.3 min (300 nm).

Figure Legends

Figure S1. HPLC/MS of KA1055: UV/Vis traces at 254, 300 and 220 nm, Total Ion Current (TIC).

Figure S2. HPLC/MS of KA1055: Area % at 300 nm

Figure S3. Mass spectrum of KA1055 from ESI^+ HPLC/MS showing mono- and di-protonated species, $[M+H]^+$ and $[M+2H]^{2+}$.

Figure S4. Expansion of the mass spectrum of KA1055 for $[M+H]^+$.

Figure S5. Further expansion of the mass spectrum of KA1055 for $[M+H]^+$.

Figure S6. 1H NMR (600 MHz, DMSO-d^6) of KA1055.

Figure S7. 1H NMR (500 MHz, DMSO-d^6) of KA1055.

Figure S8. 1H NMR (500 MHz, DMSO-d^6 plus D_20) of KA1055.

Figure S9. 1H NMR (500 MHz, MeOH-d^4) of KA1055.

Figure S10. Thermal melting curves of the DNA hairpin duplex containing the TGGCTT site in the absence and presence of KA1002, KA1007, KA1055 and KA1039 at 0.5:1, 1:1, 1.5:1 and 2:1 molar ratios of compound to DNA.

Figure S11. Mutant DNA sequences of TGGCTT with the hairpin loop underlined.

Figure S12. Models of PA binding induced DNA bending. (A) X-ray crystal structure of cyclic eight-ring PA bound to 5'-CCAGTACTGG-3' (PDB 3OMJ, ref. 39.) (B) Docking models of B-

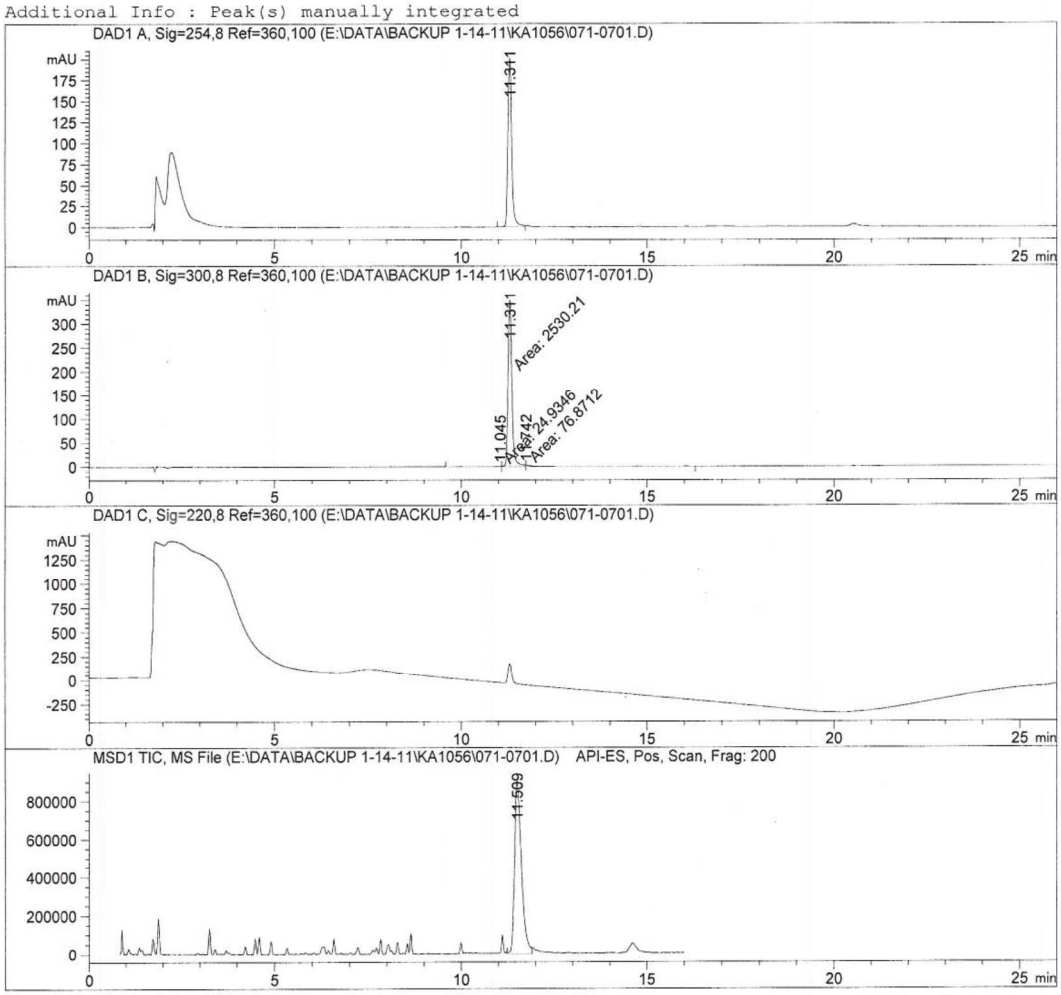
DNA, Glucocorticoid receptor DNA-binding domain (GRDBD)+ DNA, PA+ DNA and PA+ DNA+ GRDBD based on the X-ray crystal structures PDB 3G6P and PDB 3OMJ (ref. 41).

Data File E:\DATA\BACKUP 1-14-11\KA1056\071-0701.D
Sample Name: KA1055A

```
==========================================================================
Acq. Operator   : Aston                          Seq. Line :   7
Acq. Instrument : Instrument 1                    Location : Vial 71
Injection Date  : 10/29/2010 1:49:53 PM              Inj :   1
                                                 Inj Volume : 10.000 µl
Different Inj Volume from Sequence !     Actual Inj Volume : 5.000 µl
Acq. Method     : C:\CHEM32\1\METHODS\5_60ACN_26_500-2200
Last changed    : 10/28/2010 6:10:56 PM by Kevin
Analysis Method : C:\CHEM32\1\METHODS\5_60ACN_6_5_500-2200_ESI+.M
Last changed    : 6/12/2012 1:18:33 PM
                  (modified after loading)
Method Info     : 5 to 60% ACN/0.1% aq. formic acid over 4 min, 2.0 mL/min, 500-2200 MW,
                  ESI+, Analytical

Sample Info     : lyophilized combined lots of PA#3


Additional Info : Peak(s) manually integrated
```




```
LC1 6/14/2012 1:11:39 PM                                    Page   1 of 3
```

Figure S1

Area Percent Report

Sorted By : Signal
Multiplier: : 1.0000
Dilution: : 1.0000
Use Multiplier & Dilution Factor with ISTDs

Signal 1: DAD1 A, Sig=254,8 Ref=360,100

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	11.311	BB	0.1101	1426.93359	200.44853	100.0000
Totals :				1426.93359	200.44853	

Signal 2: DAD1 B, Sig=300,8 Ref=360,100

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	11.045	MF	0.5239	24.93456	7.93263e-1	0.9474
2	11.311	MF	0.1203	2530.20776	350.51517	96.1320
3	11.742	FM	0.4155	76.87117	3.08343	2.9206
Totals :				2632.01349	354.39186	

Figure S2

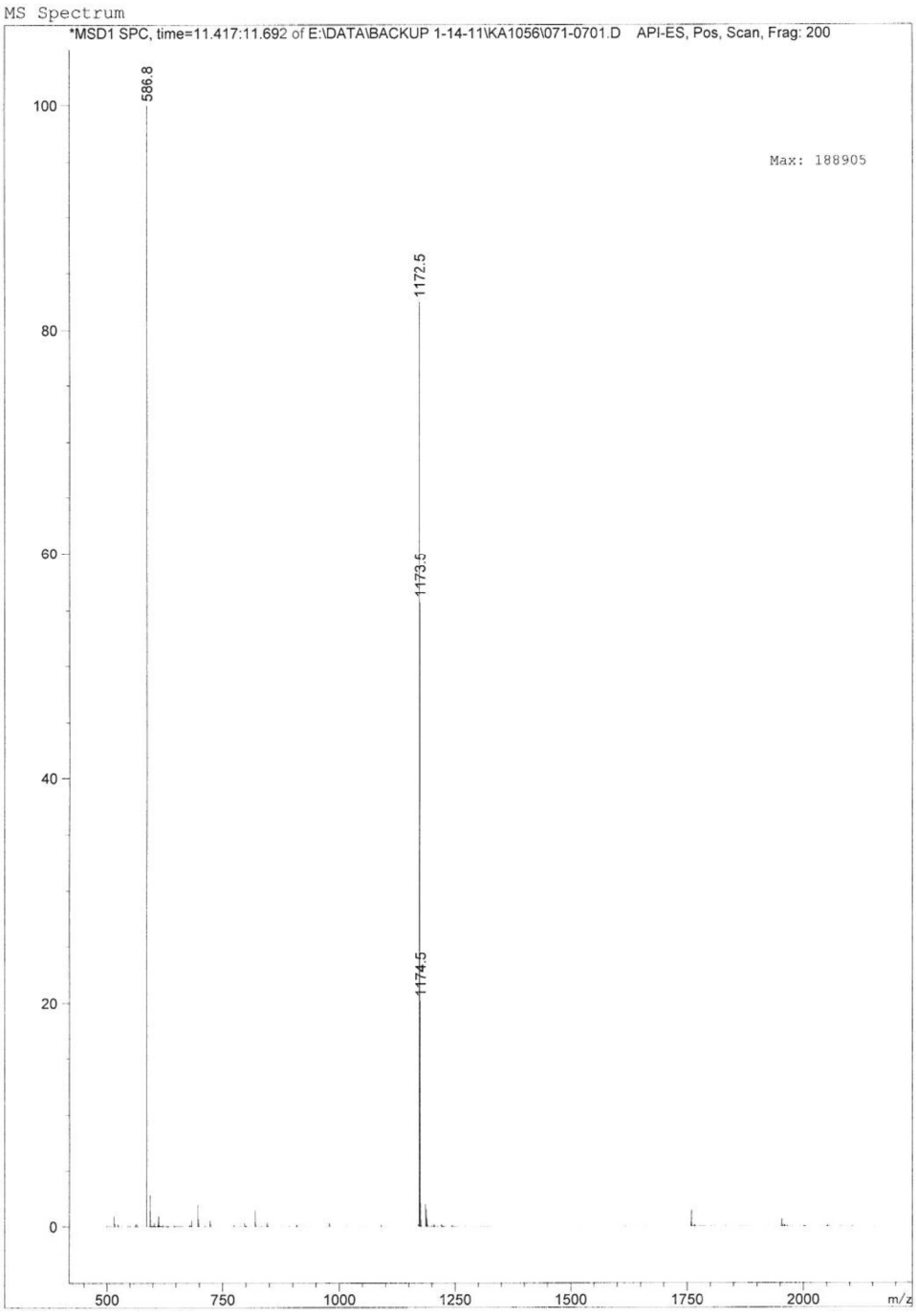

MS Spectrum
*MSD1 SPC, time=11.417:11.692 of E:\DATA\BACKUP 1-14-11\KA1056\071-0701.D API-ES, Pos, Scan, Frag: 200
Max: 188905
586.8
1172.5
1173.5
1174.5
100
80
60
40
20
0
500
750
1000
1250
1500
1750
2000
m/z

Figure S3

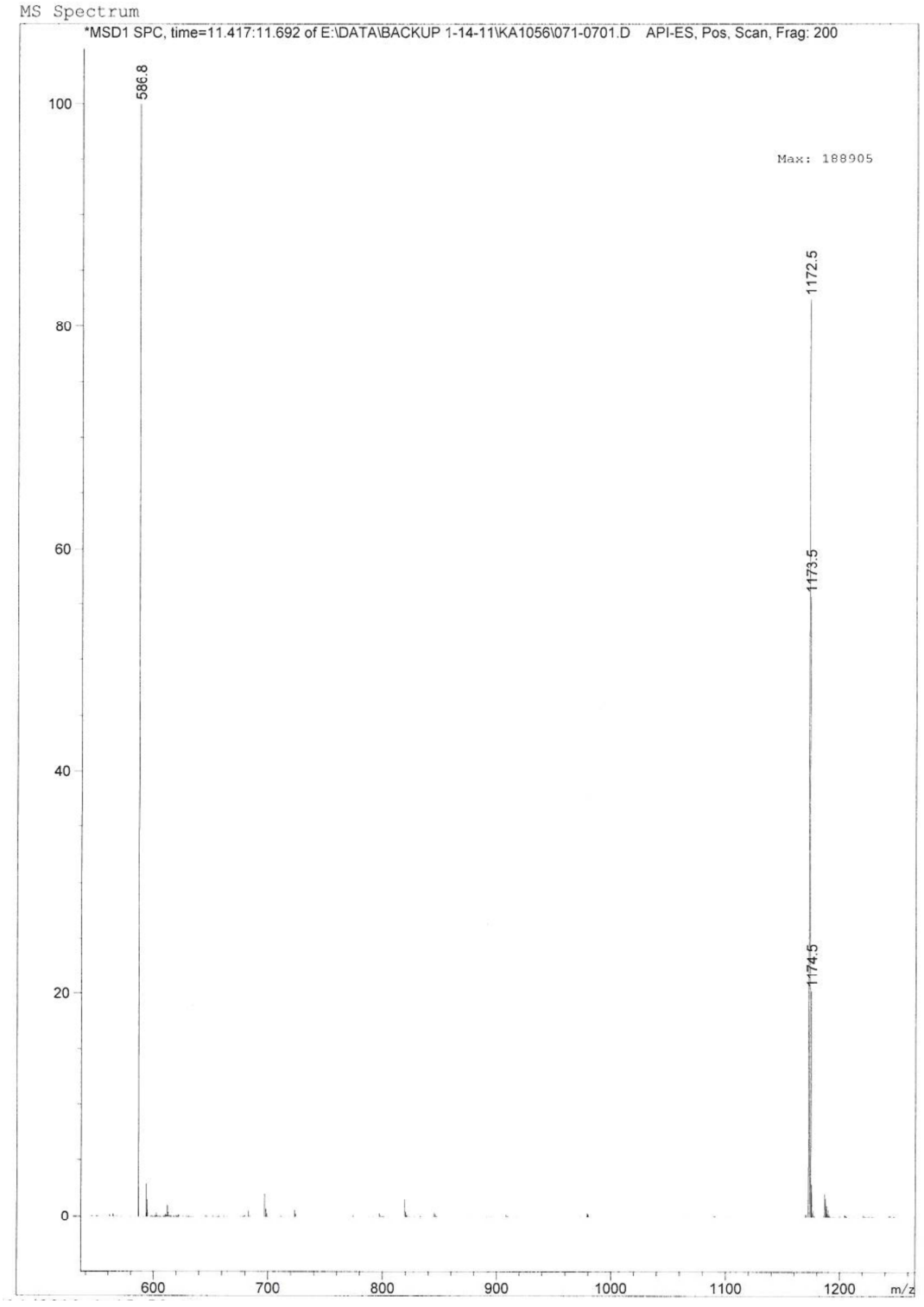

MS Spectrum
*MSD1 SPC, time=11.417:11.692 of E:\DATA\BACKUP 1-14-11\KA1056\071-0701.D API-ES, Pos, Scan, Frag: 200
586.8
Max: 188905
1172.5
1173.5
1174.5
100
80
60
40
20
0
600
700
800
900
1000
1100
1200
m/z

Figure S4

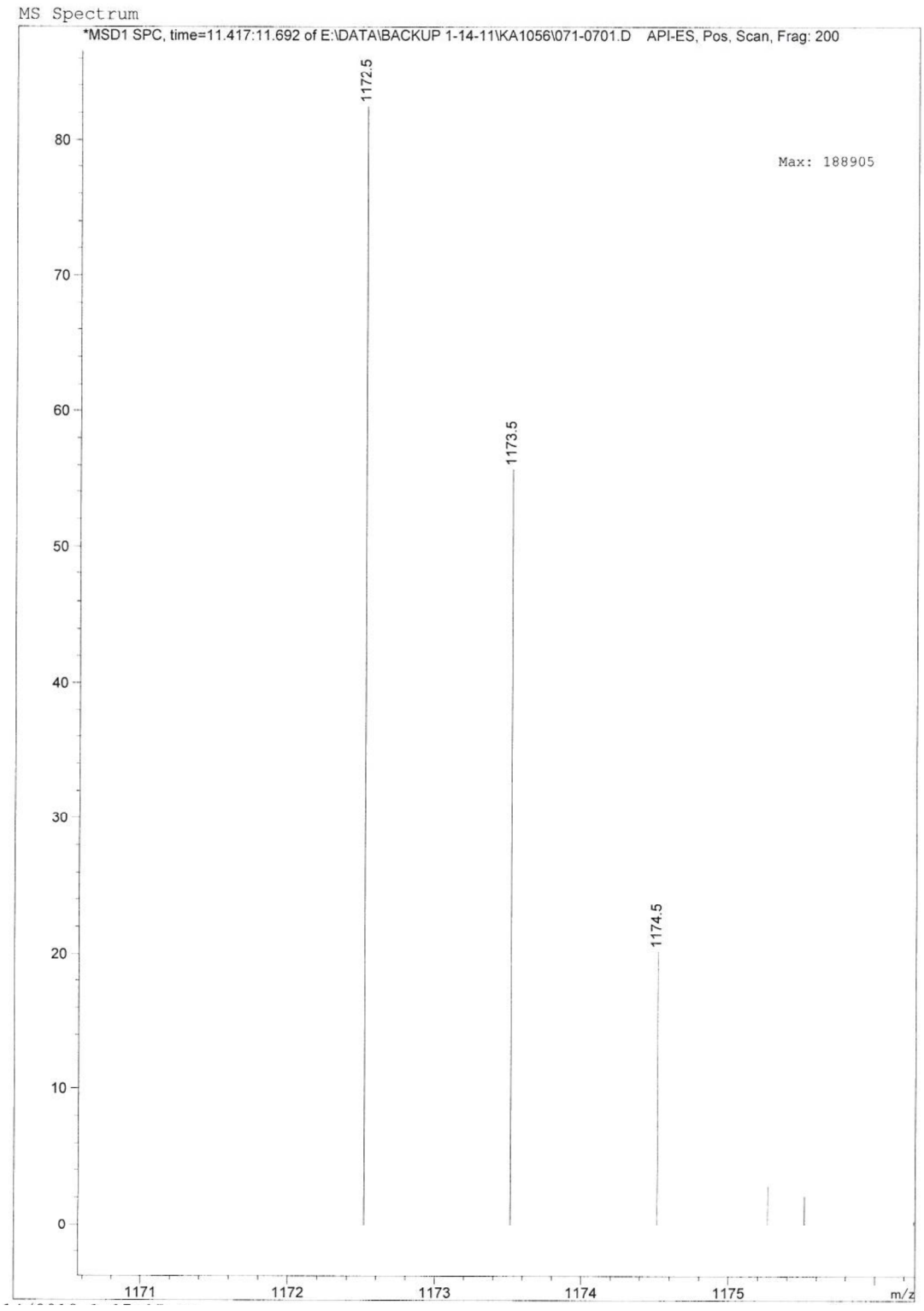

MS Spectrum
*MSD1 SPC, time=11.417:11.692 of E:\DATA\BACKUP 1-14-11\KA1056\071-0701.D API-ES, Pos, Scan, Frag: 200
Max: 188905
1172.5
1173.5
1174.5
80
70
60
50
40
30
20
10
0
1171
1172
1173
1174
1175
m/z

Figure S5

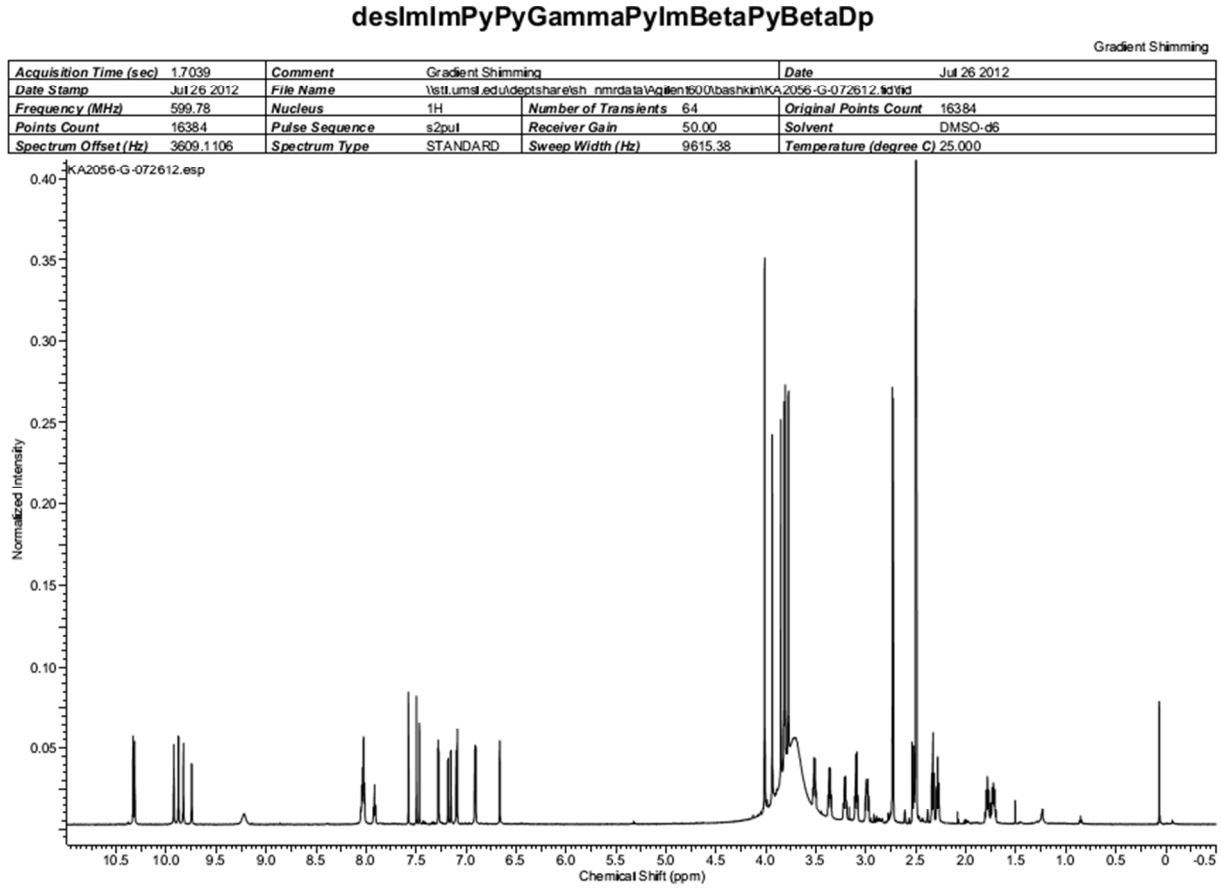

desImImPyPyGammaPyImBetaPyBetaDp
Gradient Shimming
Acquisition Time (sec) 1.7039
Comment Gradient Shimming
Date Jul 26 2012
Date Stamp Jul 26 2012
File Name \\stf.umsl.edu\deptshare\sh_nmrdata\Agilent500\bashkin\KA2056-G-072612.fid\fid
Frequency (MHz) 599.78
Nucleus 1H
Number of Transients 64
Original Points Count 16384
Points Count 16384
Pulse Sequence s2pul
Receiver Gain 50.00
Solvent DMSO-d6
Spectrum Offset (Hz) 3609.1106
Spectrum Type STANDARD
Sweep Width (Hz) 9615.38
Temperature (degree C) 25.000
KA2056-G-072612.esp
Normalized Intensity
Chemical Shift (ppm)

Figure S6

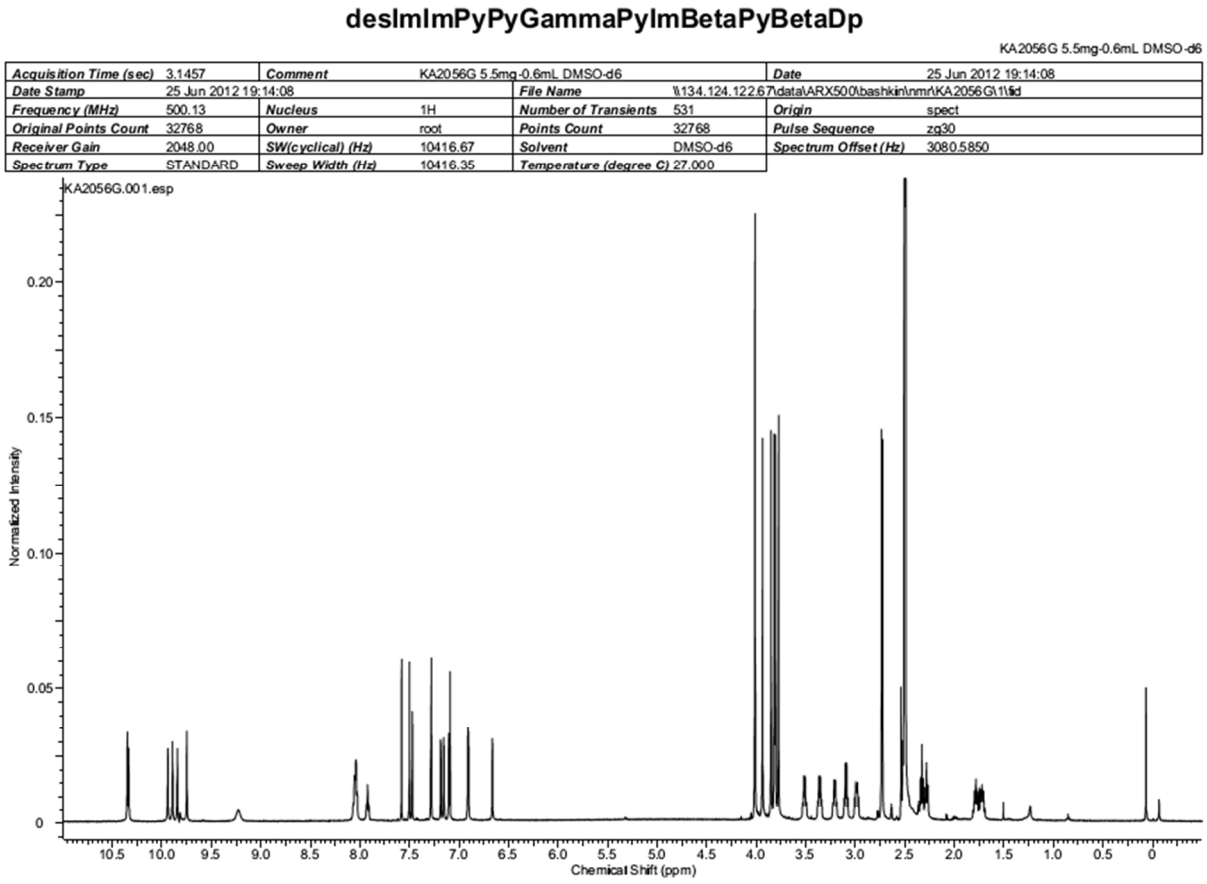

desImImPyPyGammaPyImBetaPyBetaDp
KA2056G 5.5mg-0.6mL DMSO-d6
Acquisition Time (sec) 3.1457
Comment KA2056G 5.5mg-0.6mL DMSO-d6
Date 25 Jun 2012 19:14:08
Date Stamp 25 Jun 2012 19:14:08
File Name \\134.124.122.67\data\ARX500\bashkin\nmr\KA2056G\1\fid
Frequency (MHz) 500.13
Nucleus 1H
Number of Transients 531
Origin spect
Original Points Count 32768
Owner root
Points Count 32768
Pulse Sequence zg30
Receiver Gain 2048.00
SW(cyclical) (Hz) 10416.67
Solvent DMSO-d6
Spectrum Offset (Hz) 3080.5850
Spectrum Type STANDARD
Sweep Width (Hz) 10416.35
Temperature (degree C) 27.000
KA2056G.001.esp
Normalized Intensity
0.20
0.15
0.10
0.05
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 3.0 2.5 2.0 1.5 1.0 0.5 0
Chemical Shift (ppm)

Figure S7

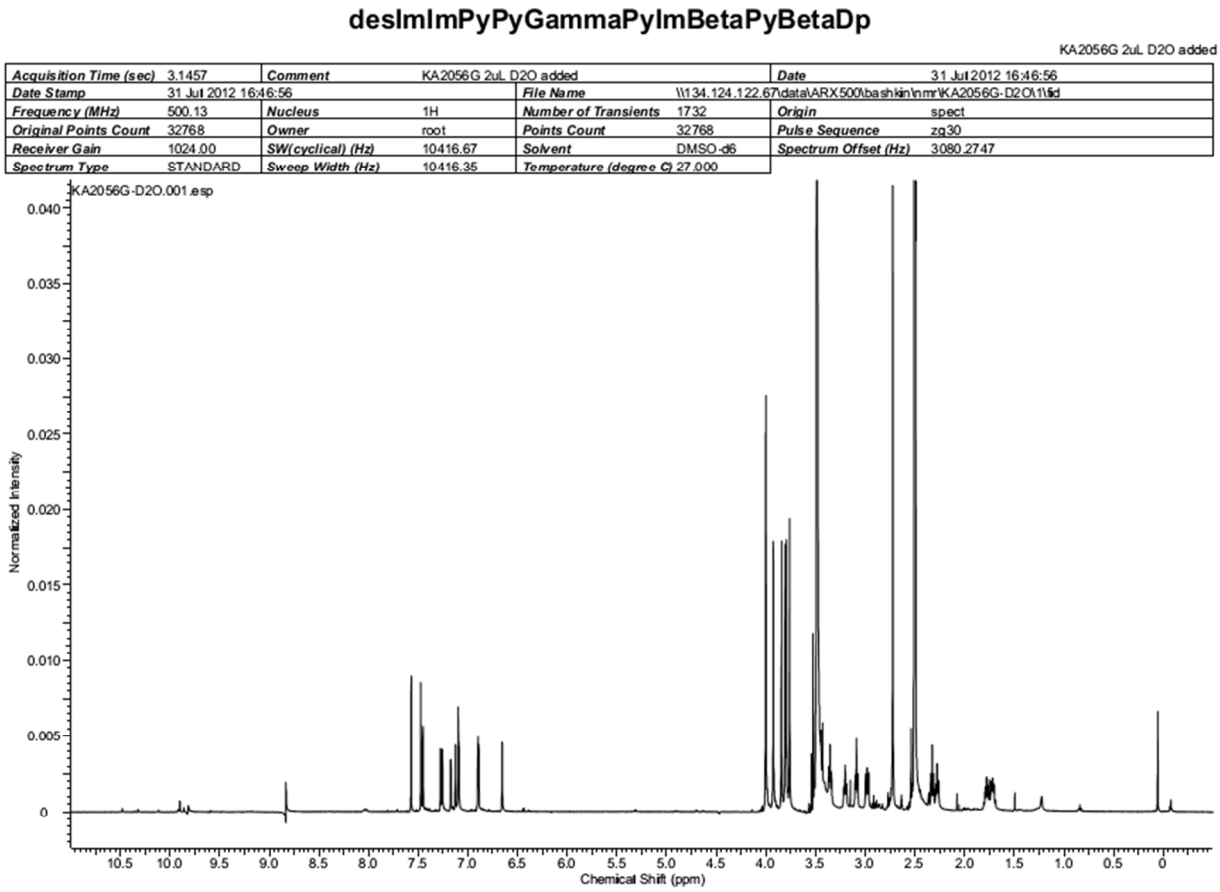

desImImPyPyGammaPyImBetaPyBetaDp
KA2056G 2uL D2O added
Acquisition Time (sec) 3.1457
Comment KA2056G 2uL D2O added
Date 31 Jul 2012 16:46:56
Date Stamp 31 Jul 2012 16:46:56
File Name \\134.124.122.67\data\ARX500\bashkin\nmr\KA2056G-D2O\1\fid
Frequency (MHz) 500.13
Nucleus 1H
Number of Transients 1732
Origin spect
Original Points Count 32768
Owner root
Points Count 32768
Pulse Sequence zg30
Receiver Gain 1024.00
SW(cyclical) (Hz) 10416.67
Solvent DMSO-d6
Spectrum Offset (Hz) 3080.2747
Spectrum Type STANDARD
Sweep Width (Hz) 10416.35
Temperature (degree C) 27.000
KA2056G-D2O.001.esp
Normalized Intensity
0.040
0.035
0.030
0.025
0.020
0.015
0.010
0.005
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 3.0 2.5 2.0 1.5 1.0 0.5 0
Chemical Shift (ppm)

Figure S8

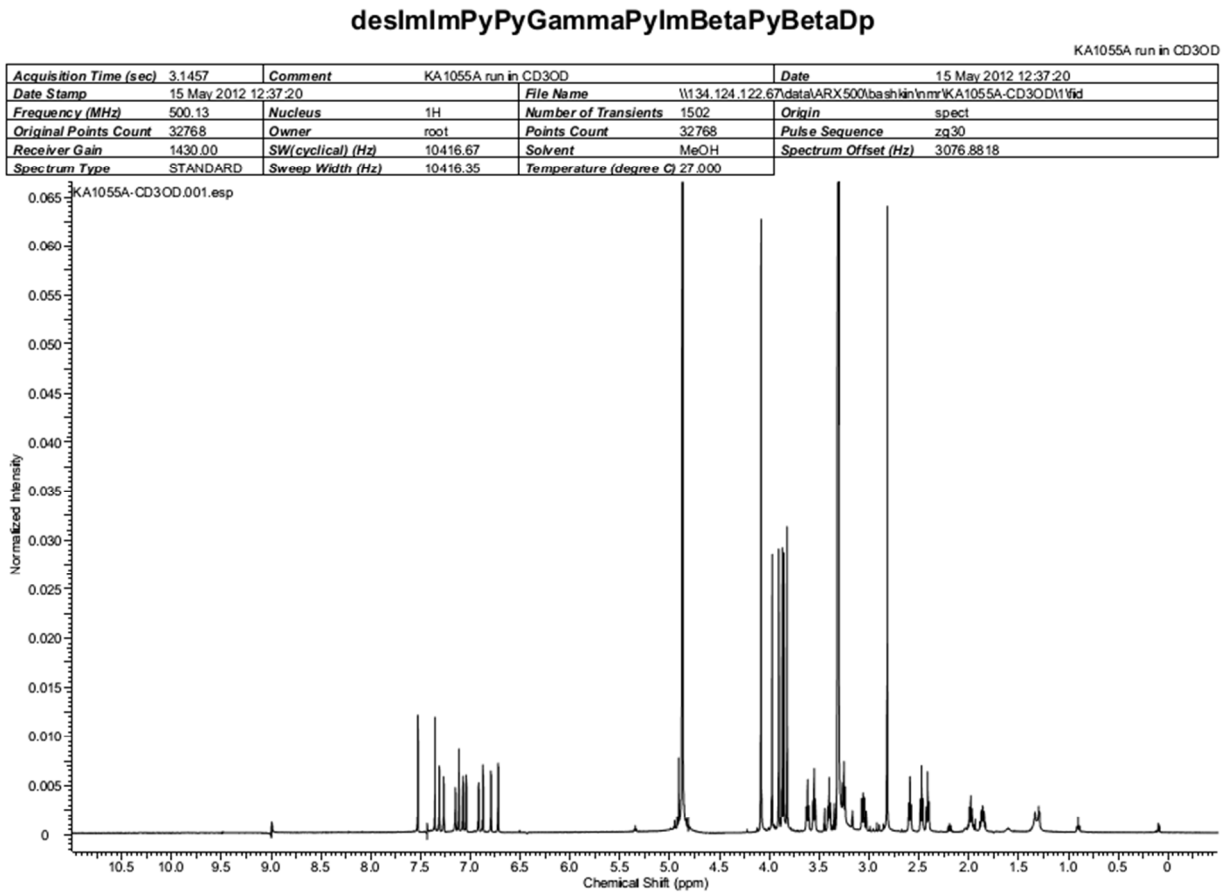

desImImPyPyGammaPyImBetaPyBetaDp
KA1055A run in CD3OD
Acquisition Time (sec) 3.1457
Comment KA1055A run in CD3OD
Date 15 May 2012 12:37:20
Date Stamp 15 May 2012 12:37:20
File Name \\134.124.122.67\data\ARX500\bashkin\nmr\KA1055A-CD3OD\1\fid
Frequency (MHz) 500.13
Nucleus 1H
Number of Transients 1502
Origin spect
Original Points Count 32768
Owner root
Points Count 32768
Pulse Sequence zg30
Receiver Gain 1430.00
SW(cyclical) (Hz) 10416.67
Solvent MeOH
Spectrum Offset (Hz) 3076.8818
Spectrum Type STANDARD
Sweep Width (Hz) 10416.35
Temperature (degree C) 27.000
KA1055A-CD3OD.001.esp
Normalized Intensity
Chemical Shift (ppm)

Figure S9

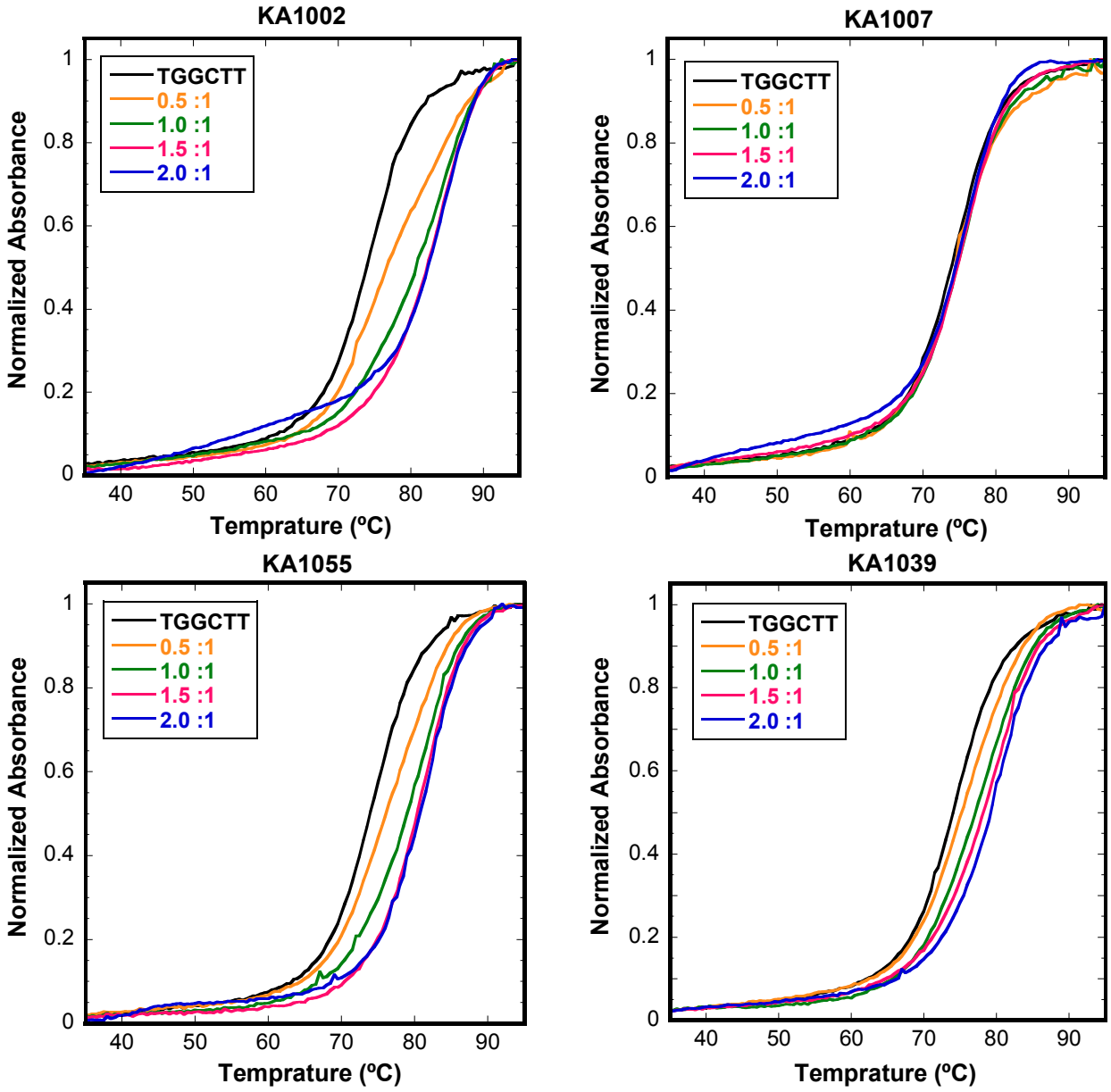

KA1002
KA1007
KA1055
KA1039
TGGCTT
0.5 :1
1.0 :1
1.5 :1
2.0 :1
Normalized Absorbance
Temprature (ºC)
0
0.2
0.4
0.6
0.8
1
40
50
60
70
80
90

Figure S10

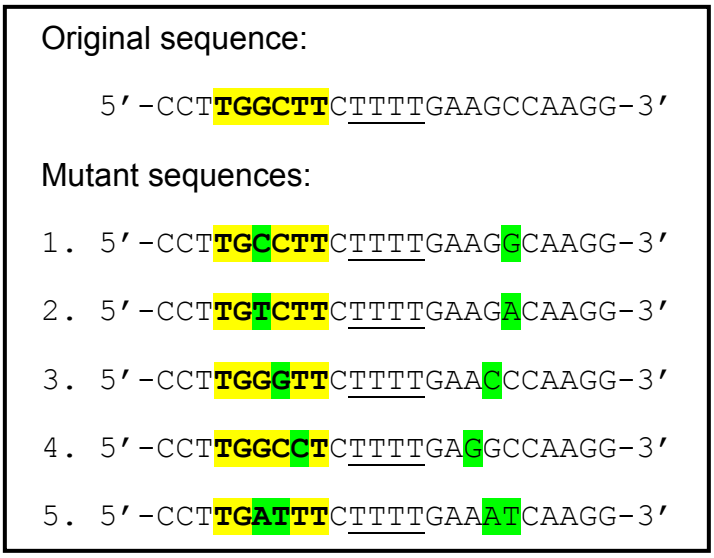

Original sequence:
5′-CCTTGGCTTCTTTTGAAGCCAAGG-3′
Mutant sequences:
1. 5′-CCTTGCCTTCTTTTGAAGGCAAGG-3′
2. 5′-CCTTGTCTTCTTTTGAAGACAAGG-3′
3. 5′-CCTTGGGTTCTTTTGAACCCAAGG-3′
4. 5′-CCTTGGCCTCTTTTGAGGCCAAGG-3′
5. 5′-CCTTGATTTCTTTTGAAATCAAGG-3′

Figure S11

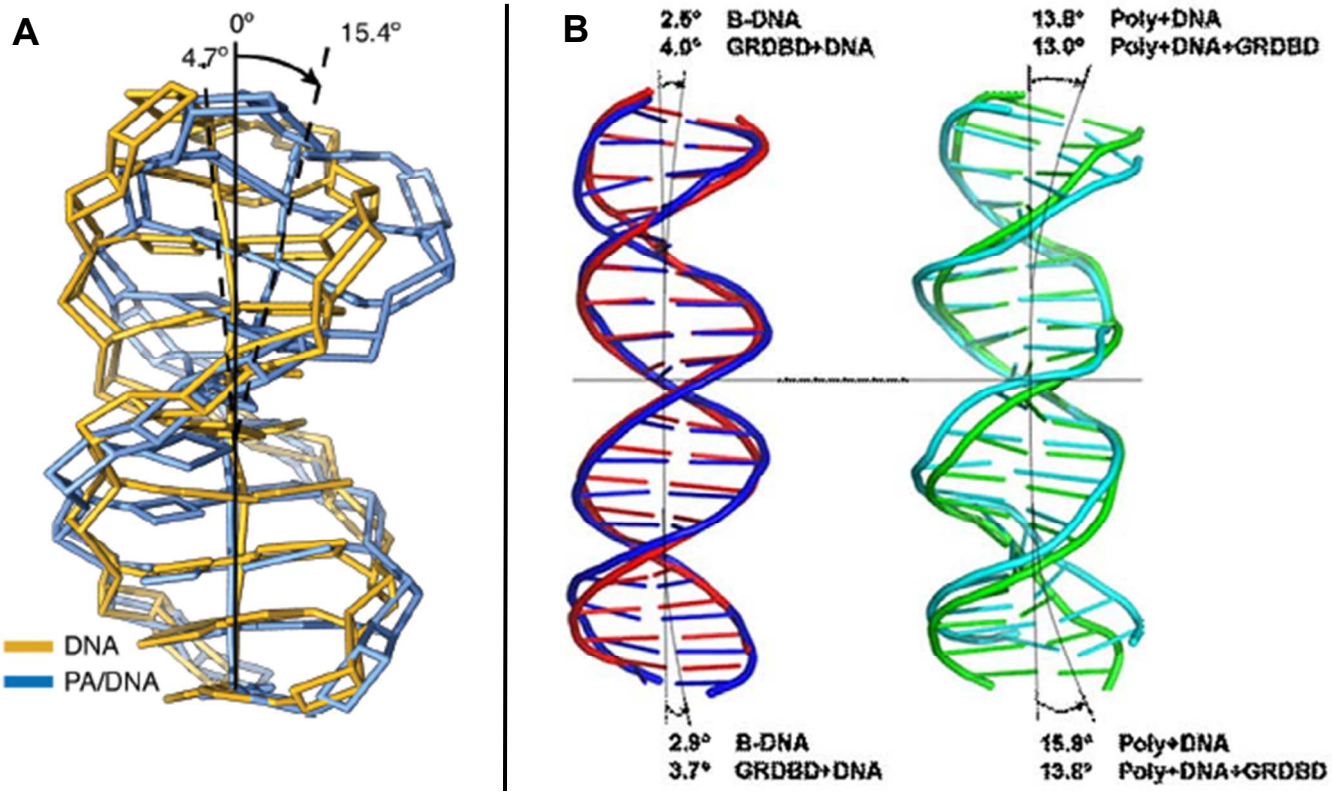

A
0°
15.4°
4.7°
DNA
PA/DNA
B
2.5° B-DNA
4.0° GRDBD+DNA
13.8° Poly+DNA
13.0° Poly+DNA+GRDBD
2.9° B-DNA
3.7° GRDBD+DNA
15.8° Poly+DNA
13.8° Poly+DNA+GRDBD

Figure S12