

Supporting Information

Anion Mediated, Tunable Isoguanosine Self-Assemblies: Decoding the Conformation Influence and Solvent Effect

Mengjia Liu,^a Ying He,^a Chuan Shan,^a Lukasz Wojtas,^a Ion Ghiviriga,^c Yu Yan,^a Xiaopeng Li,^{a,b} and Xiaodong Shi*^a

^a Department of Chemistry, University of South Florida, Tampa, Florida 33620, United States;

^b College of Chemistry and Environmental Engineering, Shenzhen University, Shenzhen 518055, PR China;

^c Department of Chemistry, University of Florida, 125 Buckman Drive, Gainesville, Florida 32611, United States.

I. Self-assembly NMR Study	S2
II. General Methods and Materials	S8
III. Synthetic Procedures and NMR Spectra	S9
IV. Single-Crystal X-Ray Diffraction	S18
V. Reference	S21

I. Self-assembly NMR Study

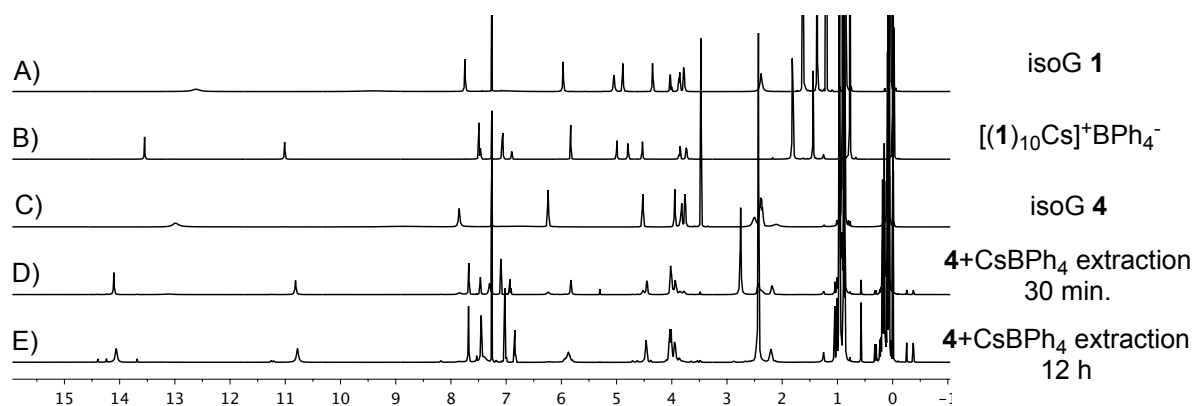


Figure S1. ^1H NMR spectra showing isoG monomers and complexes formation in CDCl_3 A) isoG **1** monomer; B) $[(1)_{10}\text{Cs}]^+(\text{BPh}_4)^-$; C) isoG **4** monomer; D) treating **4** with CsBPh_4 for 30 mins; E) treating **4** with CsBPh_4 for 12 h.

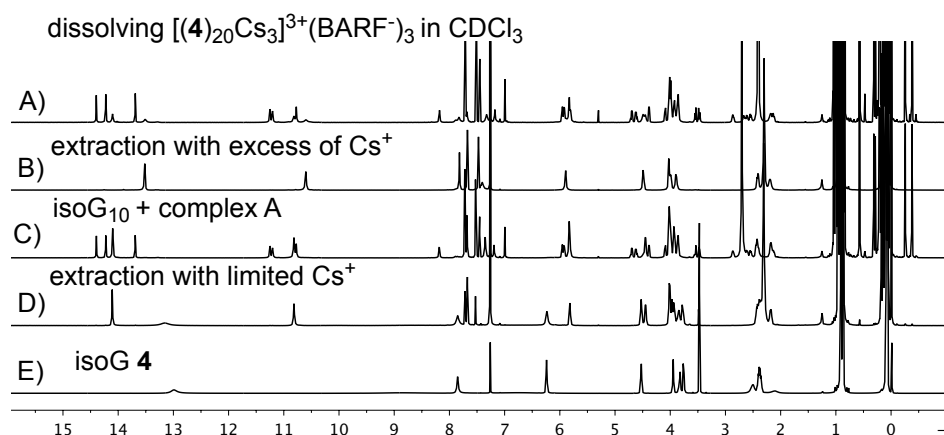


Figure S2. ^1H NMR spectra of isoG **4** assemblies in CDCl_3 A) dissolving $[(4)_{20}\text{Cs}_3]^{3+}(\text{BARF}^-)_3$ crystal; B) treating **4** with excess of CsBPh_4 ; C) addition of **4**; D) a 10:1 ratio of isoG **4**: CsCl and NaBARF in solution; E) isoG **4** monomer.

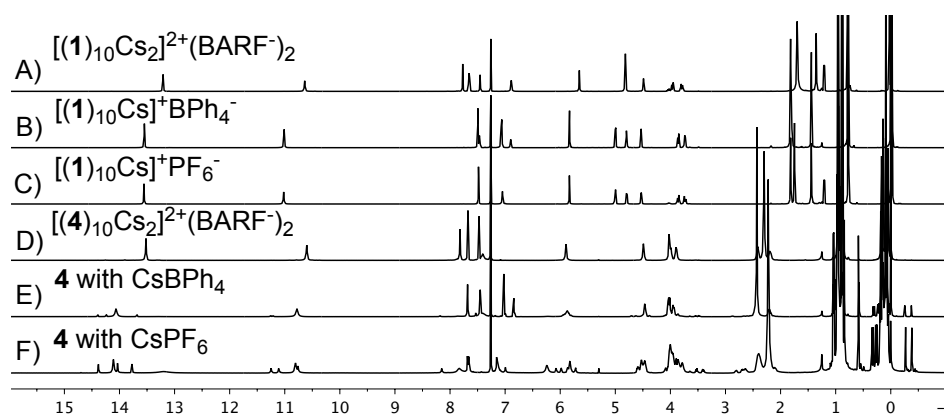


Figure S3. Critical anion-effect in isoG-star coordination with Cs^+ cation in CDCl_3 . A) $[(1)_{10}\text{Cs}_2]^{2+}(\text{BARF}^-)_2$; B) $[(1)_{10}\text{Cs}]^+\text{BPh}_4^-$; C) $[(1)_{10}\text{Cs}]^+\text{PF}_6^-$; D) $[(4)_{10}\text{Cs}_2]^{2+}(\text{BARF}^-)_2$; E) treating **4** with CsBPh_4 ; F) treating **4** with CsCl and NaPF_6 .

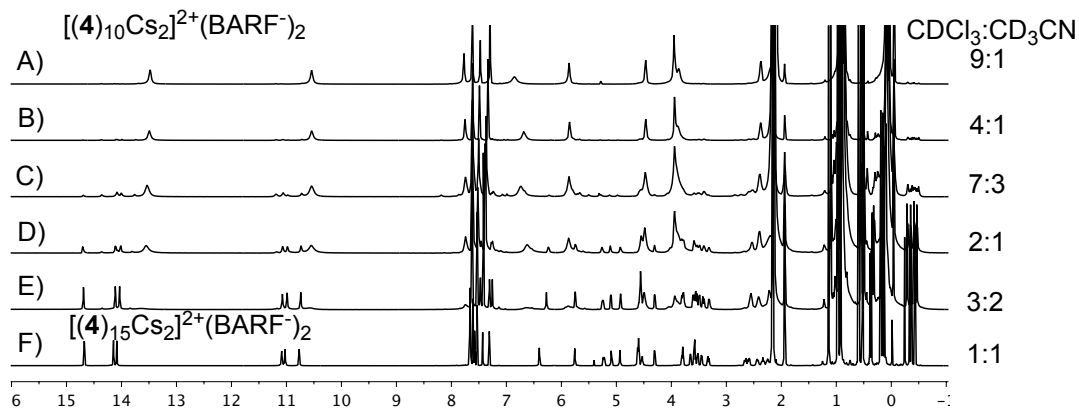


Figure S4. ^1H NMR spectra of isoG **4** in CDCl_3 and CD_3CN showing the solvent effect;

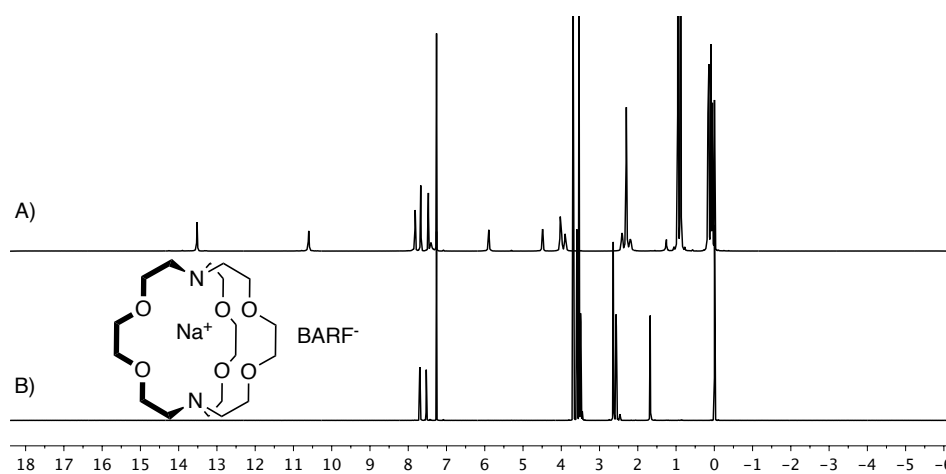


Figure S5. A) ^1H NMR spectra of $[(4)_{10}\text{Cs}_2]^{2+}(\text{BARF}^-)_2$; B) ^1H NMR spectra of $\{[2.2.2]\text{-cryptand}\}\text{Na}^+(\text{BARF}^-)$.

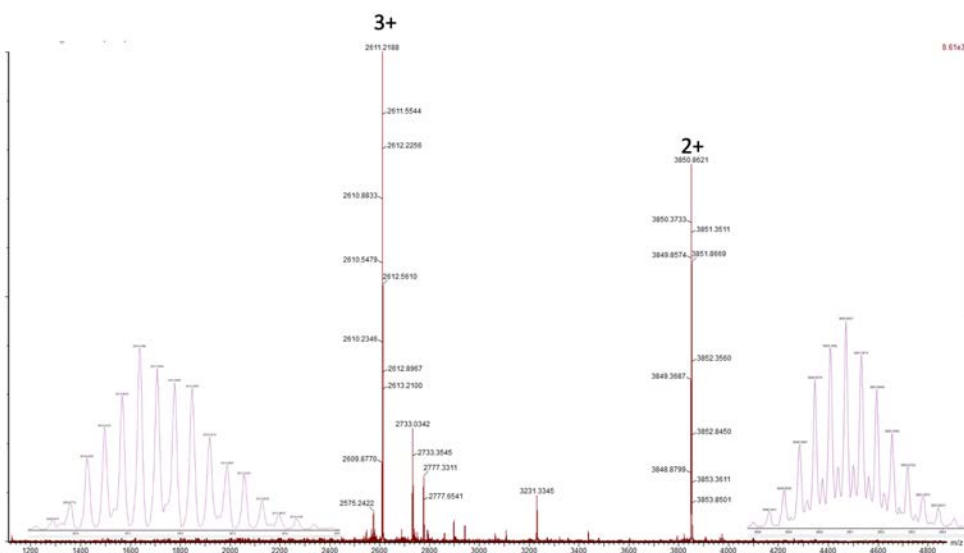


Figure S6. ESI-MS of $[(4)_{15}\text{Cs}_2]^{3+}(\text{BARF}^-)_3$.

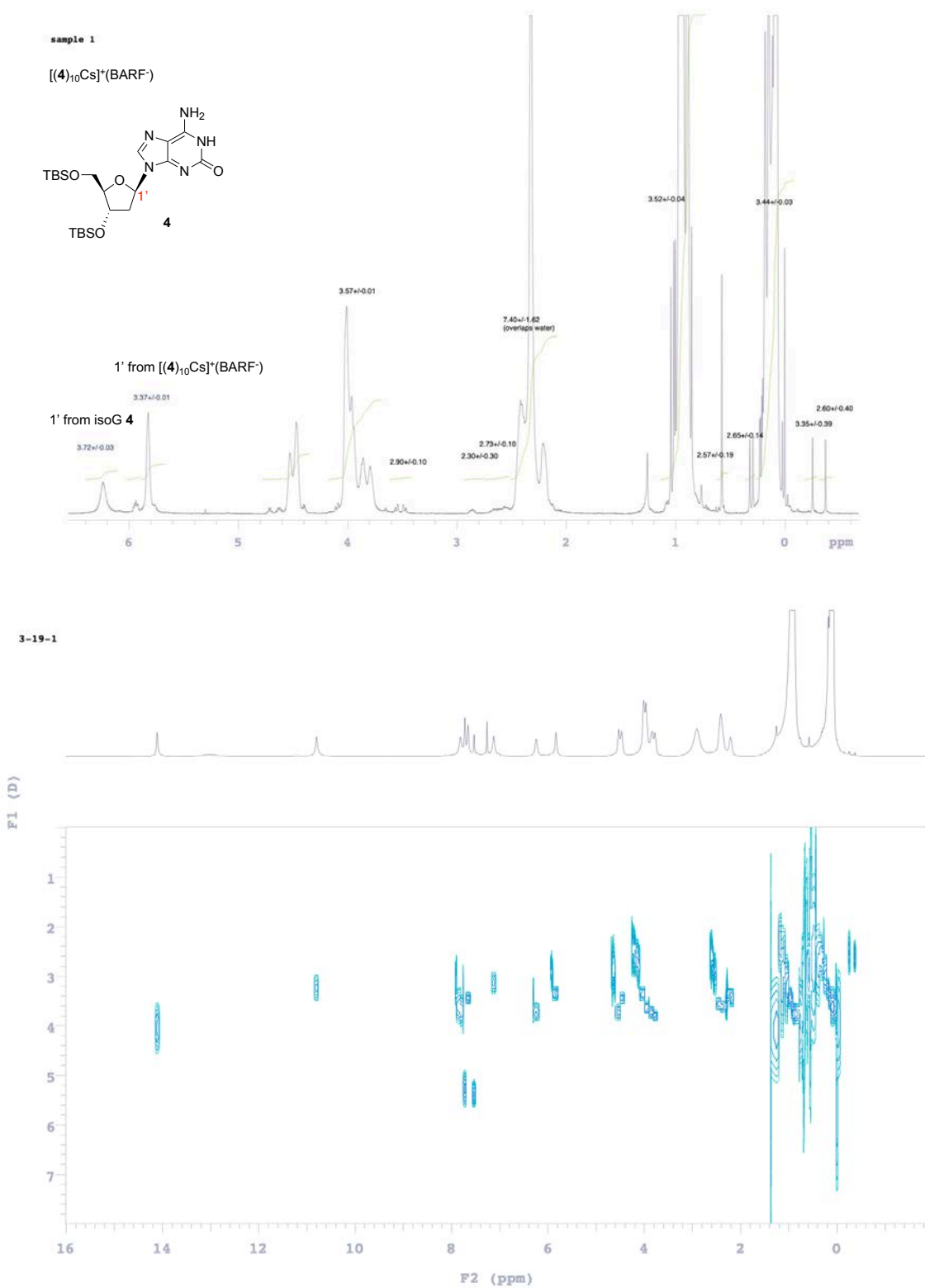


Figure S7. Diffusion coefficients (D_s) of the $[(4)_{10}Cs]^+(BARF^-)$ determined by PFG-NMR in $CDCl_3$.

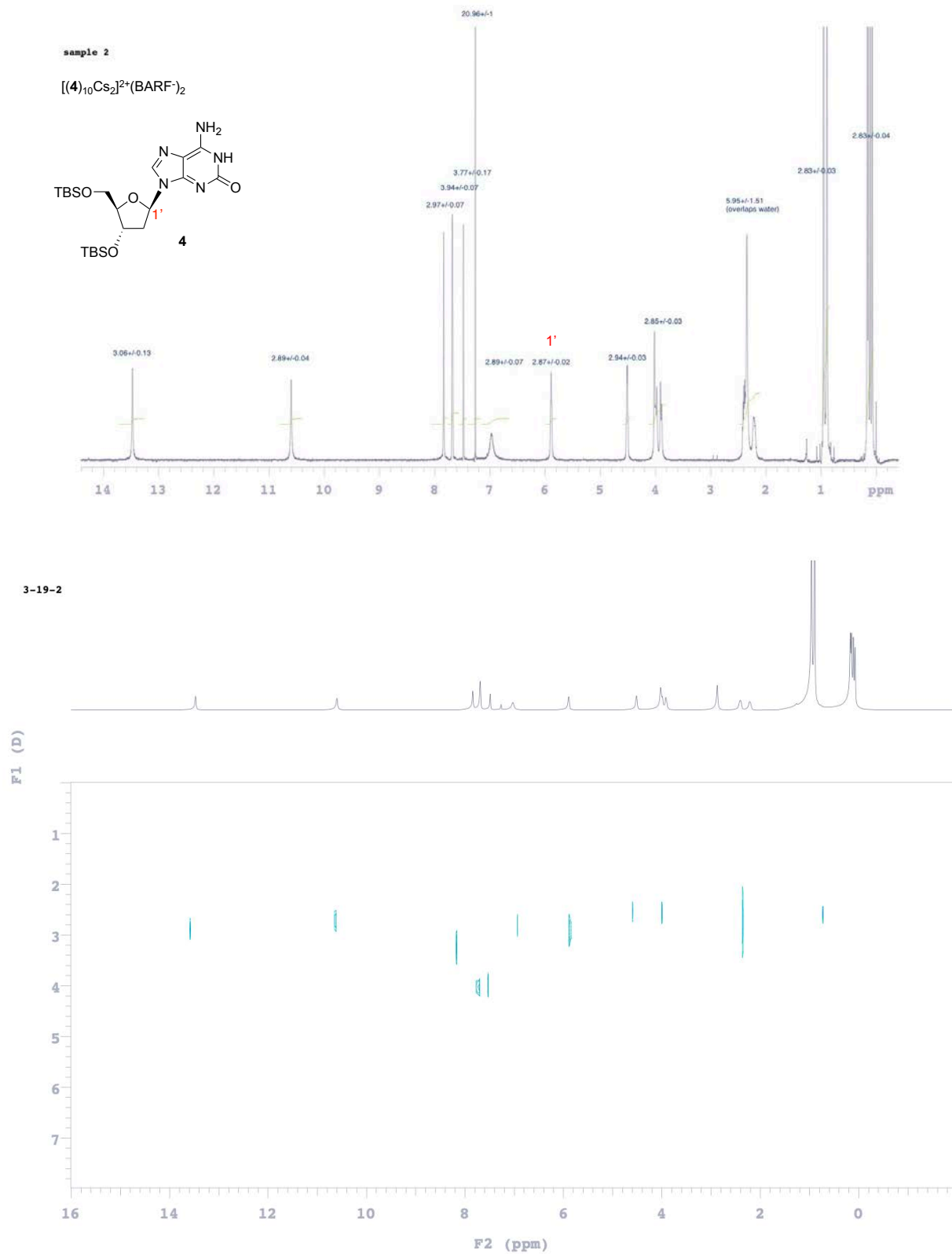


Figure S8. Diffusion coefficients (Ds) of the $[(4)_{10}Cs_2]^{2+}(BARF^-)_2$ determined by PFG-NMR in $CDCl_3$.

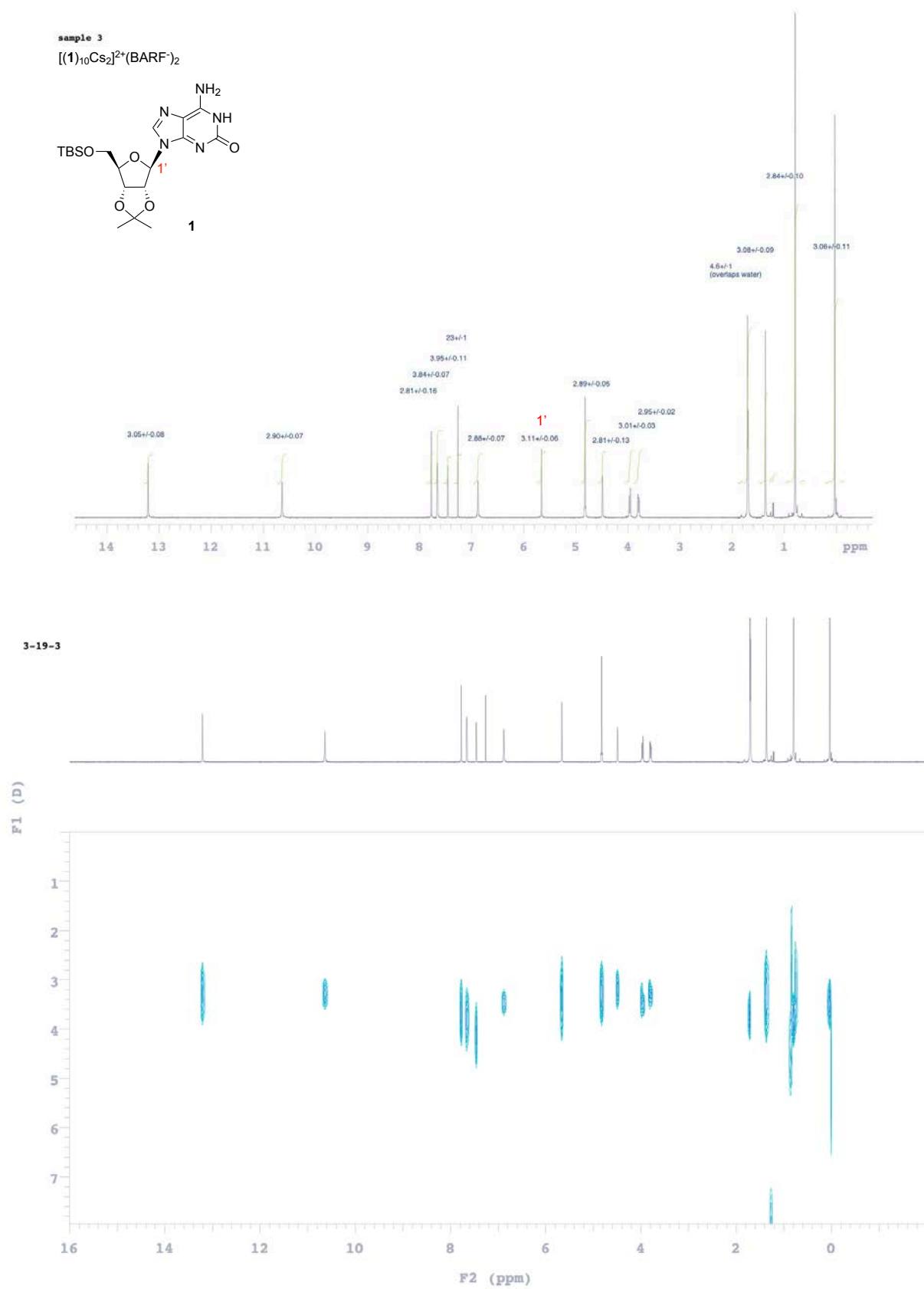


Figure S9. Diffusion coefficients (Ds) of the $[(1)_{10}Cs_2]^{2+}(BARF^-)_2$ determined by PFG-NMR in $CDCl_3$.

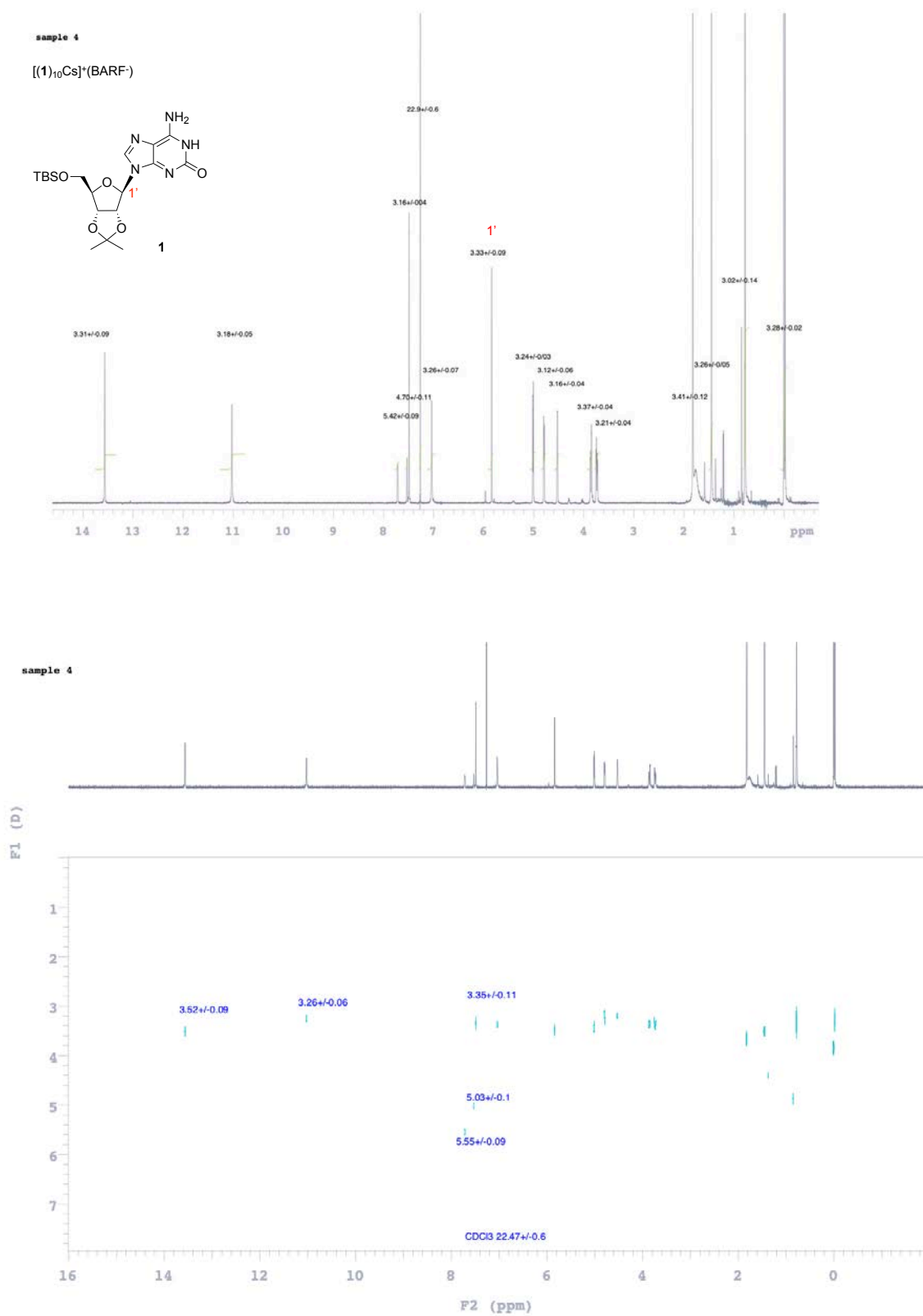


Figure S10. Diffusion coefficients (D_s) of the $[(1)_{10}Cs]^+(BARF^-)$ determined by PFG-NMR in $CDCl_3$

Table S1. Diffusion coefficients for isoG **4** complex and isoG **1** complex. All diffusion coefficients [m^2s^{-1}] have been multiplied by 10^{-10} .

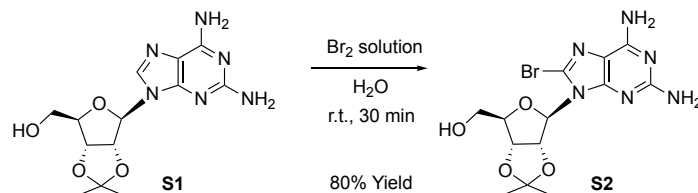
Complexes	isoG 4		isoG 1	
	4 :Cs ⁺ =5:1	4 :Cs ⁺ =10:1	1 :Cs ⁺ =5:1	1 :Cs ⁺ =10:1
ribose H-1'	2.87±0.02	3.37±0.01	3.11±0.06	3.33±0.09
Anion: BARF	3.94±0.01	4.22±0.03	3.84±0.07	5.42±0.09

II. General Methods and Materials

All of the reactions dealing with air and/or moisture-sensitive compounds were carried out under an atmosphere of argon using oven/flame-dried glassware and standard syringe/septa techniques. Unless otherwise noted, all commercial reagents and solvents were obtained from the commercial provider and used without further purification. ^1H NMR, ^{13}C NMR, spectra were recorded on Bruker Avance NEO-600 MHz spectrometers and Bruker Avance NEO-400 MHz spectrometers. Chemical shifts were reported relative to internal tetramethylsilane (δ 0.00 ppm) or CDCl_3 (δ 7.26 ppm) for ^1H ; CD_3CN (δ 1.94 ppm) for ^1H ; DMSO (δ 2.50 ppm) for ^1H and DMSO (δ 39.52 ppm) for ^{13}C . Flash column chromatography was performed on 230-430 mesh silica gel. Analytical thin layer chromatography was performed with precoated glass baked plates (250 μ) and visualized by fluorescence and by charring after treatment with potassium permanganate stain. HRMS were recorded on Agilent 6320 TOF MS/Agilent 1200 HPLC spectrometer.

III. Synthetic Procedures and NMR Spectra

3.1 Synthesis of S2



S1 was synthesized according to the literature procedure.

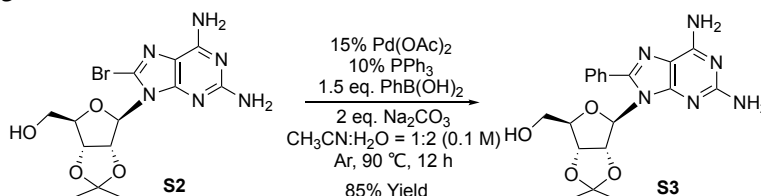
S1 (5.0 g, 15.5 mmol) was suspended in distilled water (300 mL). Saturated Br_2 aqueous solution (300 mL) was added successively to the suspension under vigorous stirring until the yellow color of Br_2 maintained in the solution. After being stirred for an additional 20 min, $\text{Na}_2\text{S}_2\text{O}_3$ was added and dried with Lyophilizer. The crude product was purified by recrystallization (DCM: MeOH=10:1) to afford **S2** as a white crystal (4.9 g, 80%).

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 6.97 (s, 2H), 6.02 (s, 2H), 5.89 (d, $J = 2.1$ Hz, 1H), 5.51 (dd, $J = 6.3, 2.1$ Hz, 1H), 5.14 (dd, $J = 6.3, 3.3$ Hz, 1H), 5.05 (t, $J = 5.8$ Hz, 1H), 4.10 (td, $J = 6.1, 5.5, 2.7$ Hz, 1H), 3.56 (dt, $J = 11.3, 5.6$ Hz, 1H), 1.52 (s, 3H), 1.32 (s, 3H).

^{13}C NMR (101 MHz, $\text{DMSO}-d_6$) δ 160.02, 155.28, 151.77, 120.81, 113.49, 112.94, 90.10, 88.04, 82.54, 81.78, 61.85, 27.17, 25.40.

HRMS m/z (ESI) calcd. for $\text{C}_{13}\text{H}_{18}\text{BrN}_6\text{O}_4^+$ ($\text{M}+\text{H}$) $^+$ 401.0567, found 401.0573.

3.2 Synthesis of S3



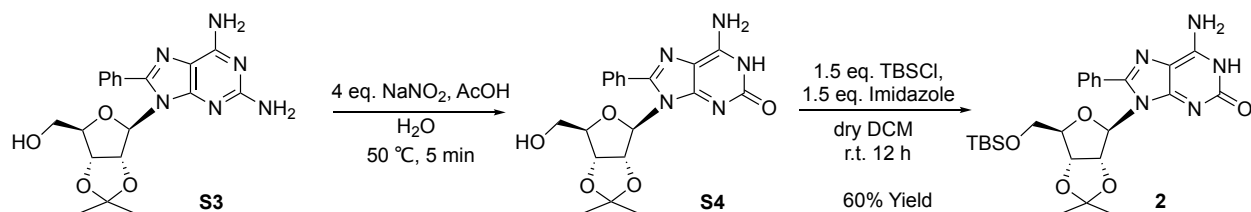
S2 (2.0 g, 2.5 mmol), arylboronic acid (928 mg, 7.6 mmol), $\text{Pd}(\text{OAc})_2$ (168 mg, 0.75 mmol), PPh_3 (140 mg, 0.5 mmol), Na_2CO_3 (1.06 g, 10.0 mmol), acetonitrile (6 mL) and water (12 mL) were added to a Schlenk tube under argon. The mixture was degassed by three freeze-pump-thaw cycles. Then the reaction was stirred at 90°C overnight till the completion monitored by LC-MS. The solution was dried with Lyophilizer. The crude product was purified by column chromatography (DCM:MeOH=10:1) to afford **S3** as white solid (1.7 g, 85%).²

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 7.72 – 7.63 (m, 2H), 7.56 (dd, $J = 9.5, 3.9$ Hz, 3H), 6.95 (s, 2H), 5.92 (s, 2H), 5.78 (d, $J = 2.5$ Hz, 1H), 5.44 (dd, $J = 6.1, 2.5$ Hz, 1H), 5.35 (dd, $J = 7.0, 4.9$ Hz, 1H), 5.17 (dd, $J = 6.1, 2.8$ Hz, 1H), 4.14 (td, $J = 5.5, 2.7$ Hz, 1H), 3.69 (dt, $J = 11.6, 4.9$ Hz, 1H), 3.55 (dt, $J = 12.0, 6.3$ Hz, 1H), 1.42 (s, 3H), 1.29 (s, 3H).

^{13}C NMR (101 MHz, $\text{DMSO}-d_6$) δ 159.87, 156.34, 151.82, 146.46, 129.80, 129.60, 129.21, 128.77, 113.09, 112.68, 89.76, 87.61, 82.49, 82.02, 62.07, 27.10, 25.32.

HRMS m/z (ESI) calcd. for $\text{C}_{19}\text{H}_{23}\text{N}_6\text{O}_4^+$ ($\text{M}+\text{H}$) $^+$ 399.1775, found 399.1778.

3.3 Synthesis of 2



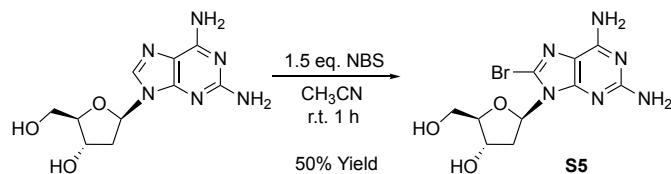
S3 (1 g 2.5 mmol) was suspended in H₂O (20 mL) at 50 °C, and NaNO₂ (690 mg, 10 mmol) in H₂O (5 mL) was added. Then AcOH (2 mL) was added at 50 °C over 1.5 min. After being stirred for 3.5 min till the completion monitored by LC-MS, the solution was diluted with H₂O (15 mL) and NH₄OH was added to PH 8. The solution was dried with Lyophilizer. Crude product **S4** was obtained without further purification. Suspension of **S4** (1.0 g, 2.5 mmol) in dry dichloromethane (40 mL) was added imidazole (0.3 g, 3.8 mmol) and TBDMSCl (0.6 g, 3.8 mmol) subsequently. The reaction was left overnight. Upon the reaction completed, 1 M HCl solution was added and extracted by DCM (3*40 mL). The combined organic layer was washed with saturated Na₂CO₃ solution then dried over MgSO₄. Removing the solvent under reduced pressure and recrystallization the pasty solid with methanol yielded **2** as white solid (773 mg, 60%).³

¹H NMR (400 MHz, DMSO-*d*₆) δ 7.71-7.63 (m, 2H), 7.61-7.49 (m, 3H), 5.73 (d, *J* = 1.6 Hz, 1H), 5.58 (dd, *J* = 6.1, 1.7 Hz, 1H), 5.03 (dd, *J* = 6.2, 2.9 Hz, 1H), 4.11 (ddt, *J* = 7.4, 5.9, 2.9 Hz, 1H), 3.82 (qd, *J* = 10.8, 6.8 Hz, 2H), 1.40 (s, 3H), 1.26 (s, 3H), 0.78 (s, 9H), -0.09 (s, 6H).

¹³C NMR (101 MHz, DMSO-*d*₆) δ 156.03, 152.21, 148.81, 130.34, 129.87, 129.48, 129.29, 113.08, 109.58, 90.09, 88.67, 82.82, 82.69, 64.00, 27.37, 26.21, 25.64, 18.44, -4.96.

HRMS *m/z* (ESI) calcd. for C₂₅H₃₆N₅O₅Si⁺ (M+H)⁺ 514.2480, found 514.2488.

3.4 Synthesis of S5



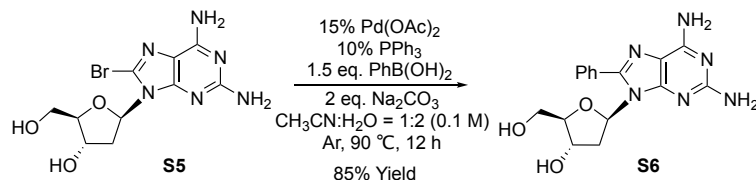
2-Amino-2'-deoxyadenosine (5 g, 18.8 mmol) was suspended in CH₃CN (50 mL), and N-Bromosuccinimide (5 g 28.2 mmol) was added. After being stirred for 1h, the precipitate was filtered. The crude product was purified by column chromatography (DCM:MeOH=10:1) to afford **S5** as white solid (3.2 g, 50%).

¹H NMR (400 MHz, DMSO-*d*₆) δ 7.01 (s, 2H), 6.18 (dd, *J* = 8.5, 6.3 Hz), 5.83 (s 2H), 5.53 (s, 1H), 5.29 (d, *J* = 4.0 Hz, 1H), 4.42 (dq, *J* = 5.7, 2.5 Hz, 1H), 3.87 (td, *J* = 4.4, 2.0 Hz, 1H), 3.65 (dd, *J* = 11.8, 4.4 Hz, 1H), 3.51 (dt, *J* = 11.4, 4.9 Hz, 1H), 3.18 (ddd, *J* = 13.8, 8.6, 5.7 Hz, 1H), 2.09 (ddd, *J* = 13.1, 6.3, 2.2 Hz, 1H).

¹³C NMR (101 MHz, DMSO-*d*₆) δ 159.58, 155.32, 151.99, 121.24, 114.02, 88.31, 85.86, 71.50, 62.40, 36.85.

HRMS *m/z* (ESI) calcd. for C₁₀H₁₄BrN₆O₃⁺ (M+H)⁺ 345.0305, found 345.0307.

3.5 Synthesis of S6



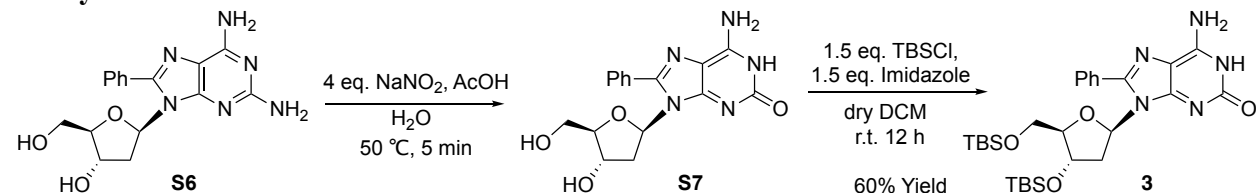
S5 (3.0 g, 8.7 mmol), arylboronic acid (1.6 g, 13.1 mmol), Pd(OAc)₂ (292 mg, 1.3 mmol), PPh₃ (228 mg, 0.9 mmol), Na₂CO₃ (1.8 g, 17.4 mmol), acetonitrile (30 mL) and water (60 mL) were added to a Schlenk tube under argon. The mixture was degassed by three freeze-pump-thaw cycles. Then the reaction was stirred at 90 °C overnight till the completion monitored by LC-MS. The solution was dried with Lyophilizer. The crude product was purified by column chromatography (DCM:MeOH=10:1) to afford **S6** as white solid (2.5 g, 85%).

¹H NMR (400 MHz, DMSO-*d*₆) δ 7.70 – 7.59 (m, 2H), 7.58 – 7.46 (m, 3H), 7.01 (s, 2H), 6.08 (dd, *J* = 9.0, 6.0 Hz, 1H), 6.02 – 5.86 (m, 1H), 5.75 (s, 2H), 5.22 (d, *J* = 3.9 Hz, 1H), 4.38 (t, *J* = 4.6 Hz, 1H), 3.86 (d, *J* = 2.2 Hz, 1H), 3.69 (dd, *J* = 11.9, 3.9 Hz, 1H), 3.55 (d, *J* = 11.8 Hz, 1H), 3.21 (ddd, *J* = 12.9, 9.0, 5.7 Hz, 1H), 2.13 – 1.96 (m, 1H).

¹³C NMR (101 MHz, DMSO-*d*₆) δ 159.78, 156.76, 152.48, 147.59, 130.74, 129.95, 129.67, 129.14, 113.99, 88.79, 85.89, 72.15, 63.00, 37.53.

HRMS *m/z* (ESI) calcd. for C₁₆H₁₉N₆O₃⁺ (M+H)⁺ 343.1513, found 343.1523.

3.6 Synthesis of **3**



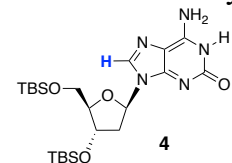
S6 (1 g, 2.9 mmol) was suspended in H₂O (20 mL) at 50 °C, and NaNO₂ (810 mg, 11.7 mmol) in H₂O (5 mL) was added. Then AcOH (2 mL) was added at 50 °C over 1.5 min. After being stirred for 3.5 min till the completion monitored by LC-MS, the solution was diluted with H₂O (15 mL) and NH₄OH was added to PH=8. The solution was dried with Lyophilizer. Crude product **S6** was obtained. Suspension of **S7** (1.0 g, 2.9 mmol) in dry dichloromethane (40 mL) was added imidazole (0.3 g, 4.4 mmol) and TBDMSCl (0.7 g, 4.4 mmol) subsequently. The reaction was left overnight. Upon the reaction completed, 1 M HCl solution was added and extracted by DCM (3*40 mL). The combined organic layer was washed with saturated Na₂CO₃ solution then dried over MgSO₄. Removing the solvent under reduced pressure and recrystallization the pasty solid with methanol yielded **3** as white solid (1 g, 60%).

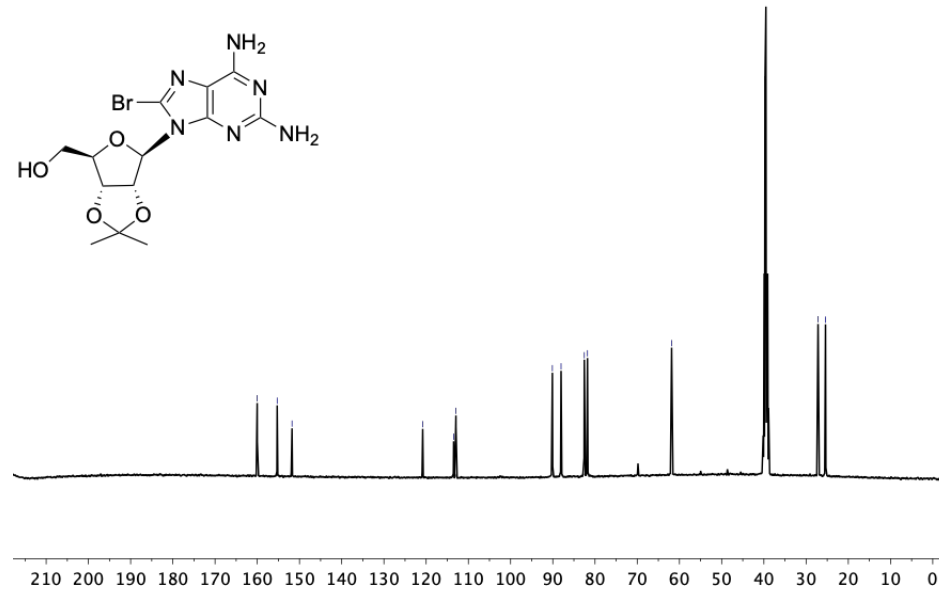
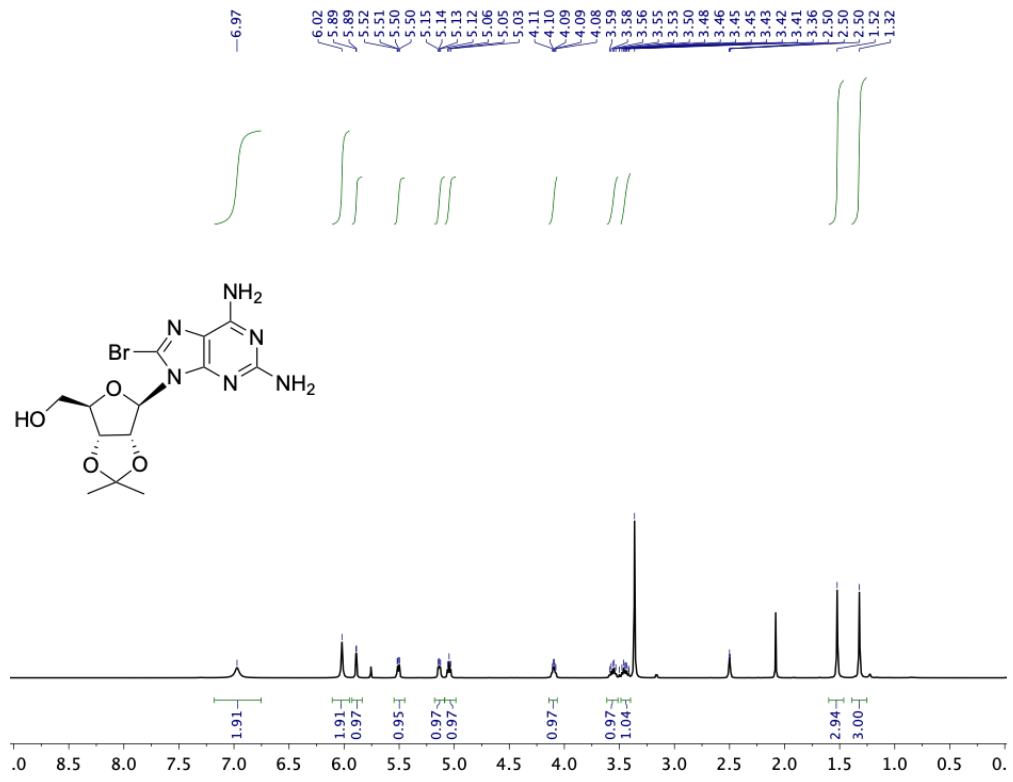
¹H NMR (400 MHz, DMSO-*d*₆) δ 10.52 (s, 1H), 7.67 (dt, *J* = 6.5, 2.9 Hz, 2H), 7.52 (q, *J* = 3.0 Hz, 3H), 6.01 (t, *J* = 6.9 Hz, 1H), 4.72 – 4.37 (m, 1H), 3.87 (dd, *J* = 10.8, 7.0 Hz, 1H), 3.79 – 3.66 (m, 1H), 3.63 (dd, *J* = 11.1, 5.7 Hz, 1H), 3.50 – 3.41 (m, 1H), 2.02 (ddd, *J* = 13.8, 6.8, 3.4 Hz, 1H), 0.97 – 0.66 (m, 18H), 0.03 (dt, *J* = 17.8, 3.4 Hz, 12H).

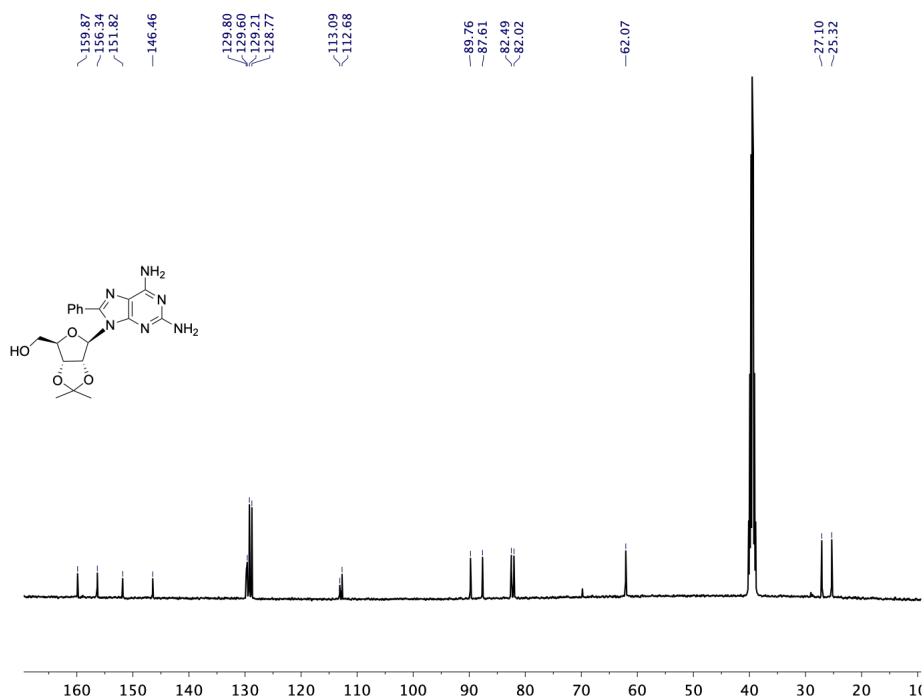
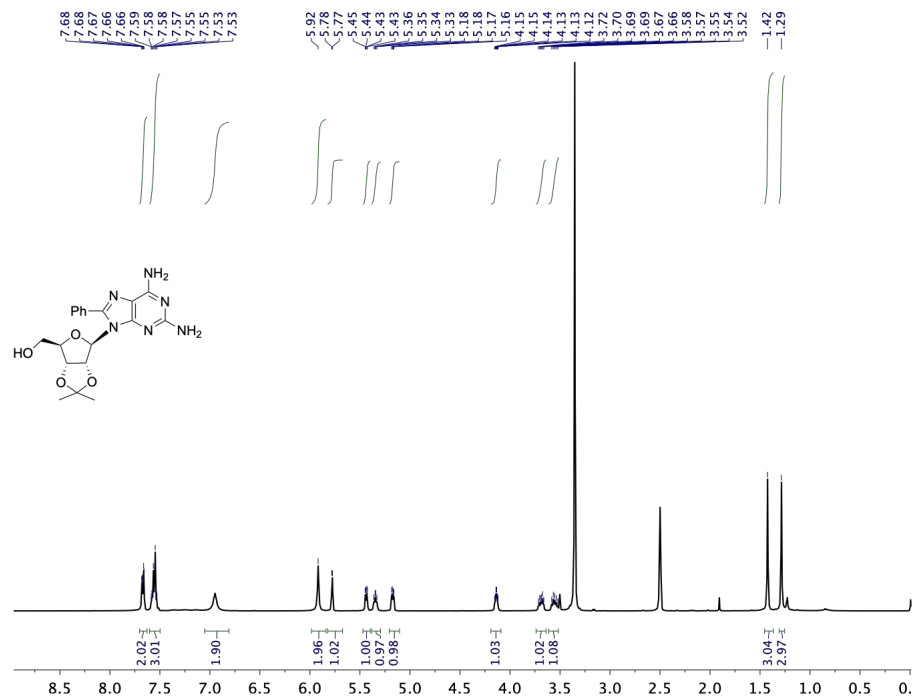
¹³C NMR (101 MHz, DMSO-*d*₆) δ 156.47, 155.56, 151.66, 149.02, 130.12, 129.73, 129.20, 128.59, 109.37, 86.91, 84.23, 72.79, 62.90, 35.36, 25.84, 25.73, 18.02, 17.71, -4.68, -4.88, -5.29, -5.36.

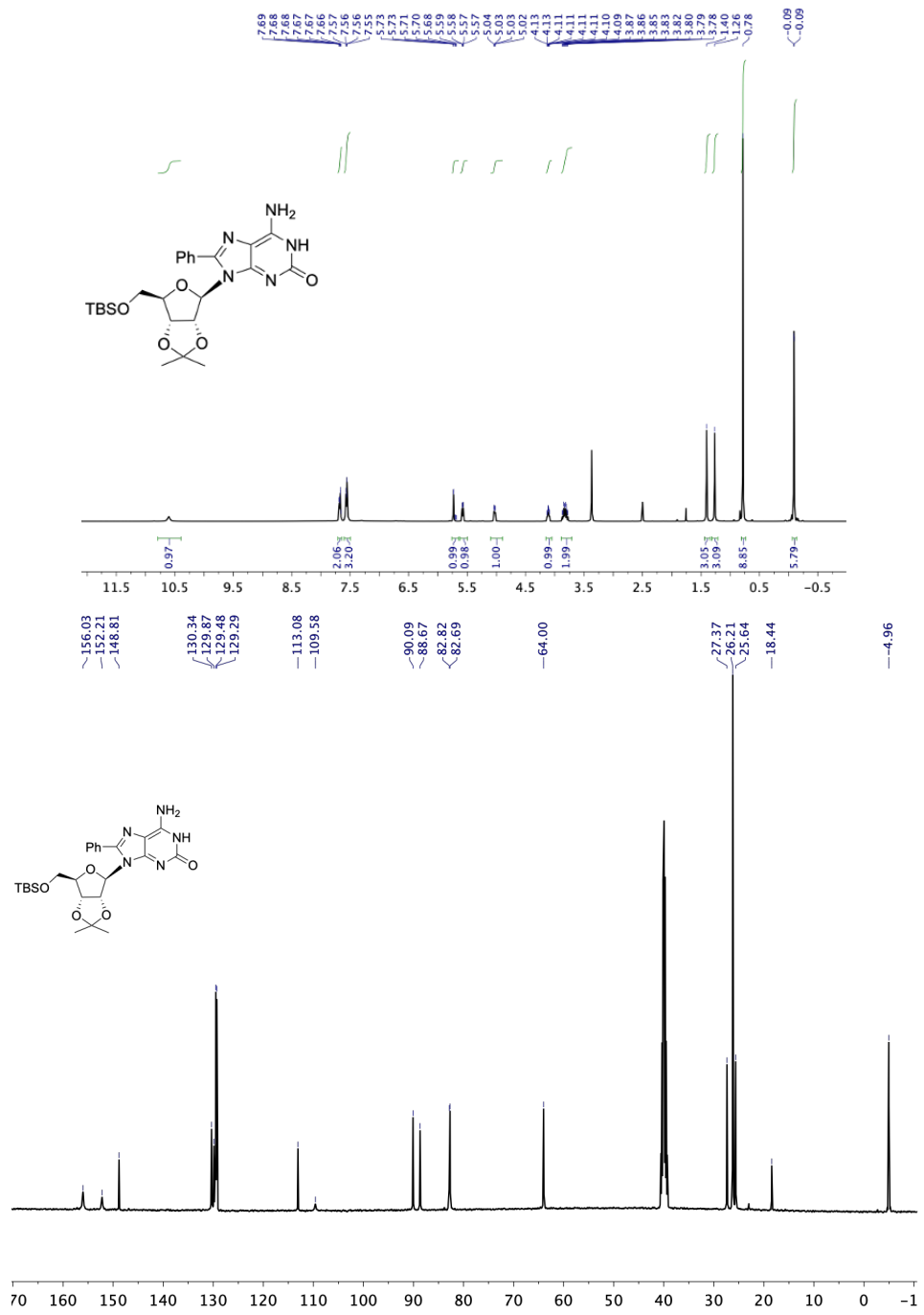
HRMS *m/z* (ESI) calcd. for C₂₈H₄₆N₅O₄Si₂⁺ (M+H)⁺ 572.3083, found 572.3088

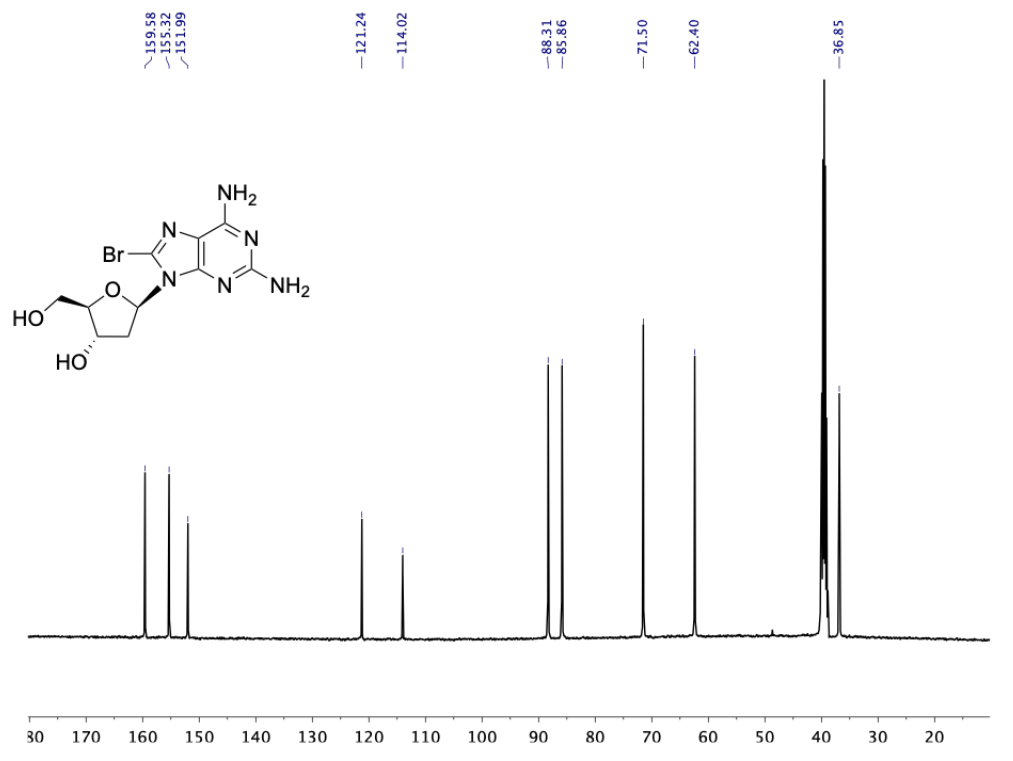
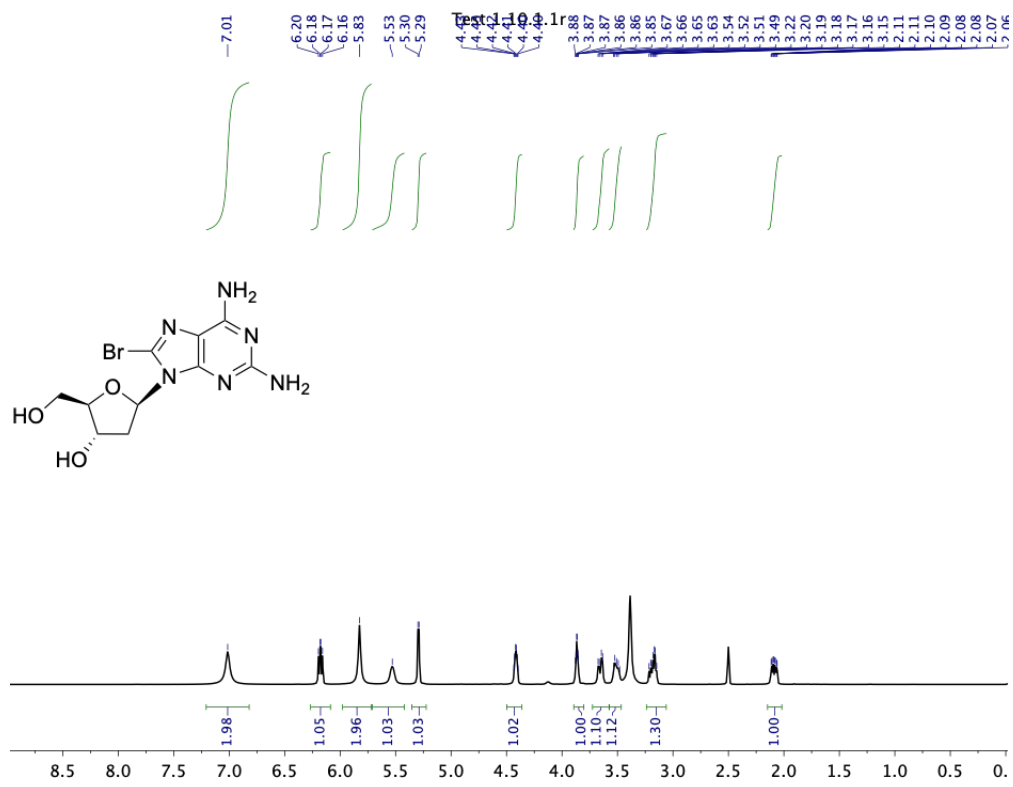
3.7 isoG 4 was synthesized according to the literature procedure.⁴

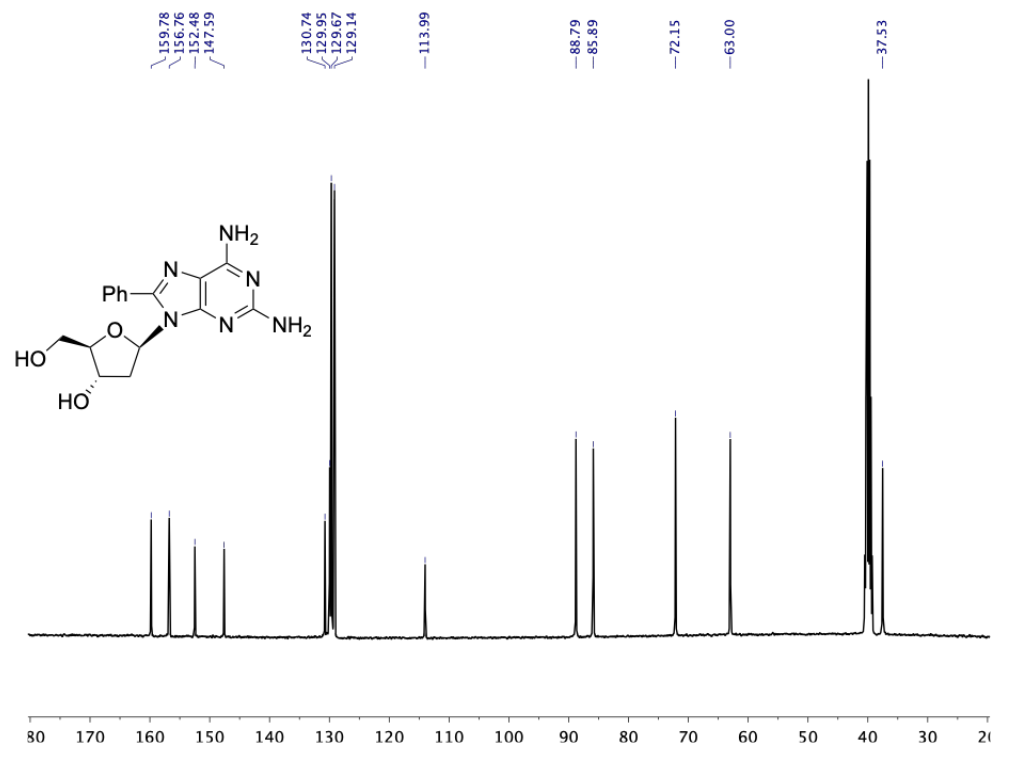
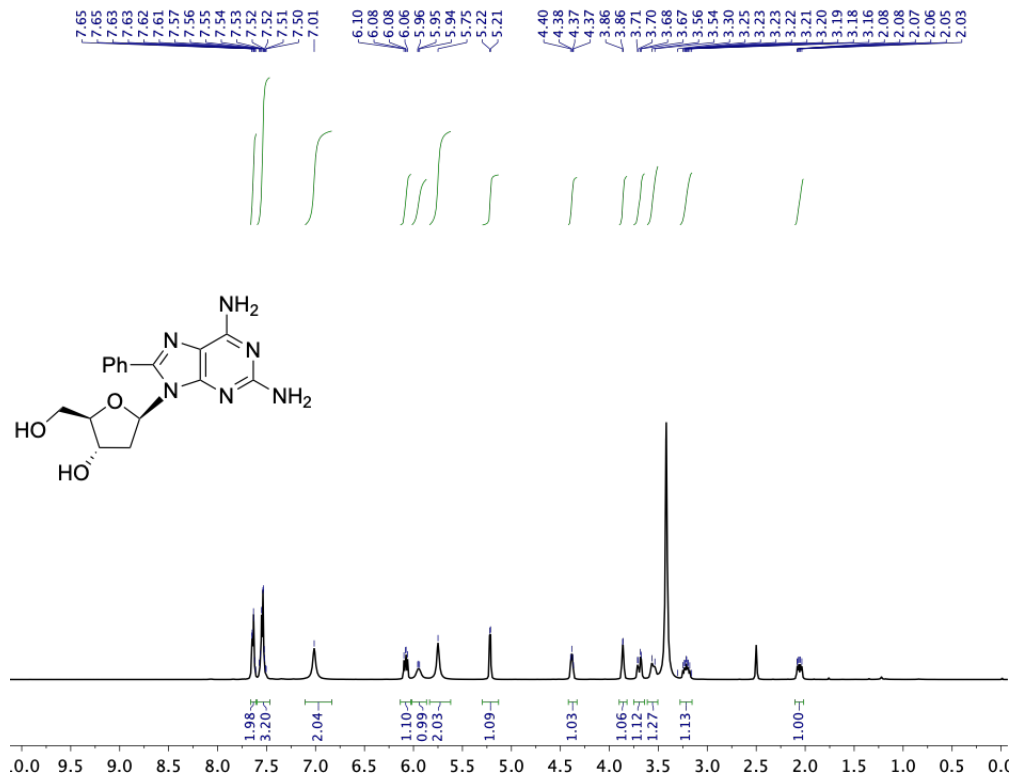


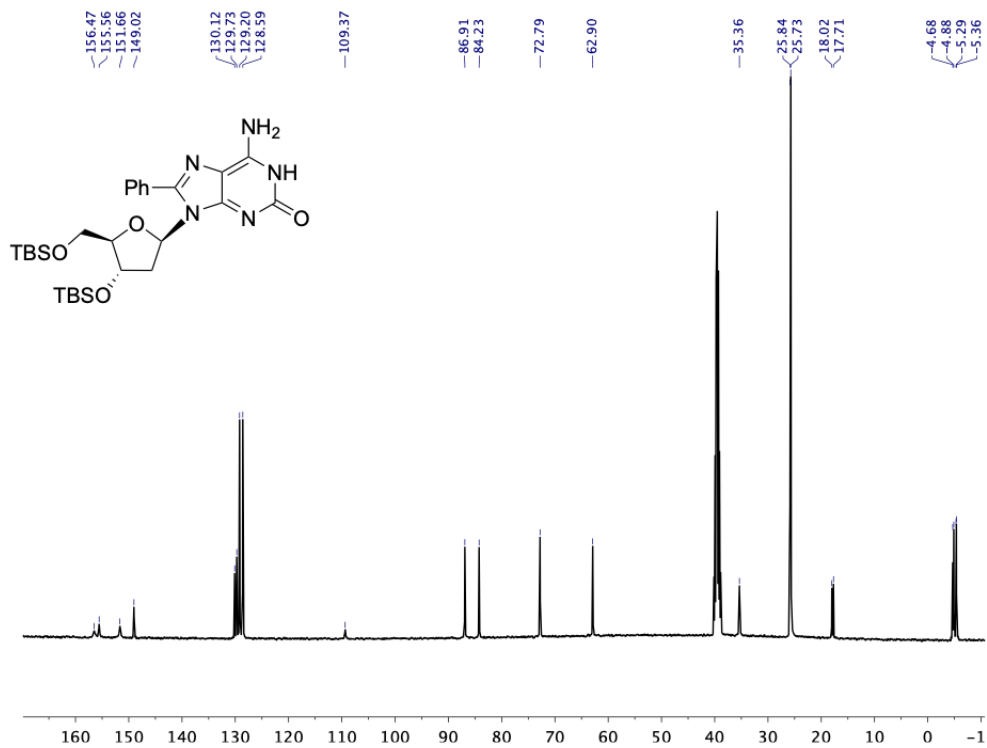
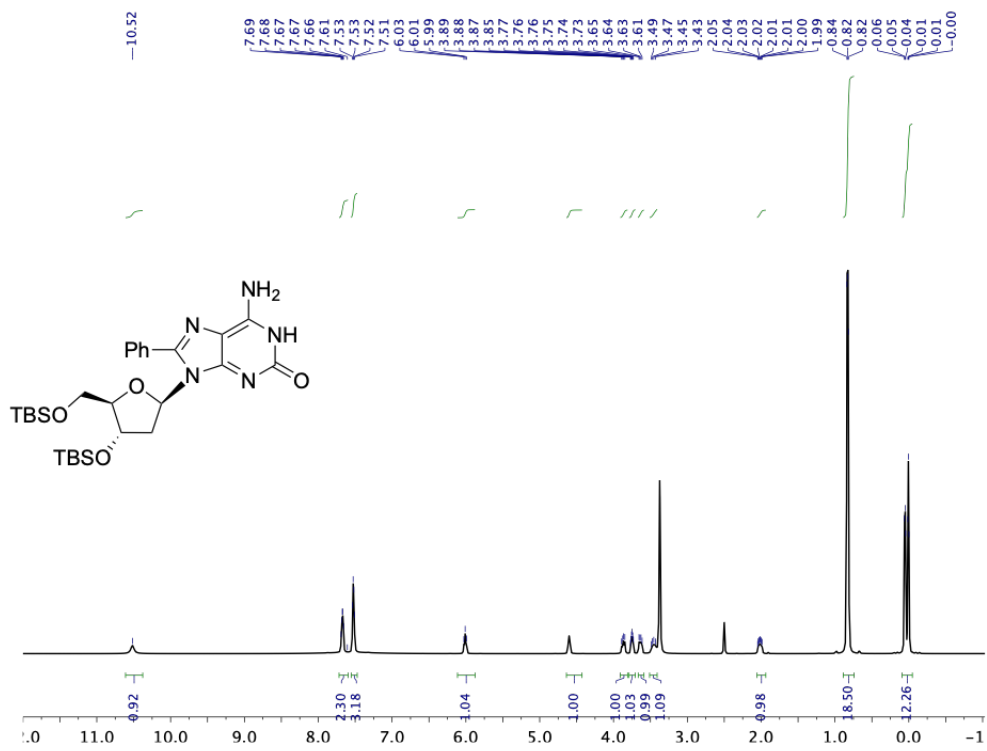












IV. Single-Crystal X-Ray Diffraction

X-ray diffraction data were measured on Bruker D8 Venture PHOTON II CPAD diffractometer equipped with a Cu K α INCOATEC ImuS micro-focus source ($\lambda = 1.54178 \text{ \AA}$). Indexing was performed using APEX3 [1] (Difference Vectors method). Data integration and reduction were performed using SaintPlus [2]. Absorption correction was performed by multi-scan method implemented in SADABS [3]. Space groups were determined using XPREP implemented in APEX3 [1]. Structures were solved using SHELXT [4] and refined using SHELXL-2018/3 [5] (full-matrix least-squares on F2) through OLEX2 interface program [6]. Ellipsoid plots were drawn with Platon [7]. **DPOT**: Disordered parts of ligand were refined with restraints. **DOT_CSBARF**: Disordered parts of ligands, counterions and solvent molecules were refined with restraints. It was not possible to reliably model diffuse electron density in structural voids. Crystal data and refinement conditions are shown in Tables 1 – 2.

[1] Bruker (2019). APEX3. Bruker AXS LLC, Madison, Wisconsin, USA.

[2] Bruker (2019) SAINT. Bruker AXS LLC, Madison, Wisconsin, USA.

[3] Krause, L., Herbst-Irmer, R., Sheldrick, G. M., Stalke, D. (2015).

"Comparison of silver and molybdenum microfocus X-ray sources for single-crystal structure determination" J. Appl. Cryst. 48, 3-10.

[4] Sheldrick, G. M. (2015). "SHELXT - Integrated space-group and crystal-structure determination", Acta Cryst. A71, 3-8.

[5] Sheldrick, G. M. (2015) "Crystal structure refinement with SHELXL", Acta Cryst., C71, 3-8

[6] Dolomanov, O.V.; Bourhis, L.J.; Gildea, R.J.; Howard, J.A.K.; Puschmann, H., OLEX2: A complete structure solution, refinement and analysis program (2009). J. Appl. Cryst., 42, 339-341

[7] Spek, A. L. (2009). "Structure validation in chemical crystallography", Acta Cryst. D65, 148-155.

[8] A.L.Spek, Acta Cryst. 2009, D65, 148-155.

[9] R. W. W. Hooft, L. H. Straver, A. L. Spek J. Appl. Cryst. (2008), 41, 96-103

Identification code	DPOT
Empirical formula	$C_{28.69}H_{47.76}N_5O_{4.94}Si_2$
Moiety formula	' $C_{28}H_{45}N_5O_4Si_2, 0.69(CH_4O), 0.25(O)$ '
Formula weight	598.00
Temperature/K	115(2)
Crystal system	monoclinic
Space group	C2
a/Å	64.413(2)
b/Å	7.1861(2)
c/Å	22.2867(7)
$\alpha/^\circ$	90
$\beta/^\circ$	99.2210(10)
$\gamma/^\circ$	90
Volume/Å ³	10182.7(5)
Z	12
ρ_{calc}/cm^3	1.170
μ/mm^{-1}	1.287
F(000)	3869.0
Crystal size/mm ³	0.200 × 0.150 × 0.120
Radiation	CuK α ($\lambda = 1.54178$)
2 Θ range for data collection/ $^\circ$	5.238 to 160.7
Index ranges	-75 ≤ h ≤ 77, -8 ≤ k ≤ 9, -28 ≤ l ≤ 28
Reflections collected	78792
Independent reflections	21131 [$R_{int} = 0.0606$, $R_{sigma} = 0.0606$]
Data/restraints/parameters	21131/1931/1519
Goodness-of-fit on F ²	1.024
Final R indexes [$I \geq 2\sigma(I)$]	$R_1 = 0.0646$, $wR_2 = 0.1792$
Final R indexes [all data]	$R_1 = 0.0746$, $wR_2 = 0.1909$
Largest diff. peak/hole / e Å ⁻³	0.44/-0.36
Flack parameter	0.007(13)

Table 2 Crystal data and structure refinement for DOT_CSBARF.

Identification code	DOT_CSBARF
Empirical formula	C ₅₅₄ H ₈₅₆ B ₃ CS ₃ F ₇₂ N ₁₀₉ O _{90.25} Si ₄₀
Moiety formula	C ₄₄₀ H ₈₂₀ CS ₃ N ₁₀₀ O ₈₀ Si ₄₀ , 3(C ₃₂ H ₁₂ BF ₂₄), 9(C ₂ N), 10.25(O)
Formula weight	13410.19
Temperature/K	100.0
Crystal system	monoclinic
Space group	P2 ₁
a/Å	26.3005(16)
b/Å	46.071(3)
c/Å	30.4842(18)
α/°	90
β/°	91.290(4)
γ/°	90
Volume/Å ³	36928(4)
Z	2
ρ _{calc} /cm ³	1.206
μ/mm ⁻¹	2.499
F(000)	14106.0
Crystal size/mm ³	0.8 × 0.4 × 0.15
Radiation	CuKα (λ = 1.54178)
2θ range for data collection/°	4.79 to 134.048
Index ranges	-31 ≤ h ≤ 31, -54 ≤ k ≤ 54, -36 ≤ l ≤ 36
Reflections collected	762378
Independent reflections	128052 [R _{int} = 0.0817, R _{sigma} = 0.0673]
Data/restraints/parameters	128052/14596/9173
Goodness-of-fit on F ²	1.040
Final R indexes [I ≥ 2σ (I)]	R ₁ = 0.1013, wR ₂ = 0.2578
Final R indexes [all data]	R ₁ = 0.1252, wR ₂ = 0.2823
Largest diff. peak/hole / e Å ⁻³	0.92/-1.77
Flack parameter	0.1340(14)

V. Reference

- [1] L. Troxler, G. Wipff, *J. Am. Chem. Soc.* **1994**, *116*, 1468-1480.
- [2] E. C. Western, J. R. Daft, E. M. Johnson, P. M. Gannett, K. H. Shaughnessy, *J. Org. Chem.* **2003**, *68*, 6767-6774.
- [3] S. C. Jurczyk, J. T. Kodra, J. H. Park, S. A. Benner, T. R. Battersby, *Helv. Chim. Acta.* **1999**, *82*, 1005-1015.
- [4] Z. Kazimierczuk, R. Mertens, W. Kawczynski, F. Seela, *Helv. Chim. Acta.* **1991**, *74*, 1742-1748