Mechanism of Action of KL-50, a Novel Imidazotetrazine for the Treatment of Drug-Resistant Brain Cancers.

Eric D. Huseman,[1] Anna Lo,[1] Olga Fedorova,[2,3] James L. Elia,[4] Susan E. Gueble,[5] Kingson Lin,[1,4,5] Ranjini Sundaram,[5] Danielle M. Burgenske,[6] Joonseok Oh,[1,7] Jinchan Liu,[1] Fabian Menges,[1,8] Matthew G. Rees,[9] Melissa M. Ronan,[9] Jennifer A. Roth,[9] Victor S. Batista,[1] Jason M. Crawford,[1,7,10] Jann N. Sarkaria,[6] Anna M. Pyle,[1,2,3] Ranjit S. Bindra,[4,5]* and Seth B. Herzon[1,5,11,]*

[1]Department of Chemistry, Yale University; New Haven, Connecticut 06520, United States.
[2]Department of Molecular, Cellular, and Developmental Biology, Yale University; New Haven, Connecticut 06520, United States
[3]Howard Hughes Medical Institute; New Haven, Connecticut 06520, United States
[4]Department of Pathology, Yale School of Medicine; New Haven, Connecticut 06520, United States.
[5]Department of Therapeutic Radiology, Yale School of Medicine; New Haven, Connecticut 06520, United States.
[6]Department of Radiation Oncology, Mayo Clinic, Rochester, Minnesota.
[7]Institute of Biomolecular Design & Discovery, Yale University; West Haven, CT, United States
[8]Department of Chemistry, Chemical and Biophysical Instrumentation Center, Yale University; New Haven, Connecticut 06520, United States.
[9]Broad Institute of MIT and Harvard; Cambridge, MA 02142, Unites States
[10]Department of Microbial Pathogenesis, Yale School of Medicine; New Haven, Connecticut, United States
[11]Department of Pharmacology, Yale School of Medicine; New Haven, Connecticut 06520, United States.

Email: *ranjit.bindra@yale.edu; *seth.herzon@yale.edu

ChemRxiv

Supporting Information

Index:

Supplementary Figures. S3
Supplementary Tables. S34
General chemical experimental procedures S36
Chemical materials S36
Chemistry instrumentation S36
Chemicals for molecular biology experiments S38
Biological reagents for molecular biology experiments S38
Cell culture S39
Long-term clonogenic survival assays S39

DepMap comparison of PRISM screens S39
Modified alkaline comet assay S39
Denaturing gel electrophoresis S39
Animal studies. S40
Statistical Analysis S40
Alkylation of calf thymus DNA with KL-50 (19). S40
5′-End labeling of deoxyoligonucleotides. S41
Kinetic analysis of the formation of the ICL and DNA–MGMT crosslinks. S42
Chemical characterization of the ICL 14c. S42
Synthetic Procedures S45
Catalog of Nuclear Magnetic Resonance Spectra S73
Bibliography S152

Supplementary Figures.

		Comparison #1[a]	Comparison #2[b]	Comparison #3[c]
X-Axis (Input)	Variable Type	Gene	Gene	Gene
	Gene Symbol	BCL2 (PPP1R50, Bcl-2)	EGFR (ERBB, ERBB1)	TOP2A (TOP2)
	Dataset	Expression Public 23Q2[d]	Expression Public 23Q2[d]	Expression Public 23Q2[d]
Y-Axis (Input)	Variable Type	Compound	Compound	Compound
	Identity	VENETOCLAX	ERLOTINIB	ETOPOSIDE
	Dataset	Drug sensitivity AUC (Sanger GDSC2)[e]	Drug sensitivity AUC (Sanger GDSC2)[e]	Drug sensitivity AUC (Sanger GDSC1)[f]
Linear Regression (Output)	Number of Points	467	488	610
	Pearson (r)	−0.590	−0.358	−0.185
	Spearman	−0.363	−0.345	−0.156
	Slope	-4.11×10^{-2}	-7.12×10^{-3}	-4.02×10^{-2}
	Intercept	0.987	0.957	1.02
	p-Value[g]	3.82×10^{-45}	3.66×10^{-16}	4.03×10^{-6}

Fig. S1. DepMap portal (https://depmap.org/portal/interactive/) parameter inputs and data outputs used for comparison with the correlation analysis for *MGMT* gene express vs. KL-50 (**19**) AUC (Fig. 6a). [a]See https://depmap.org/portal/interactive/?filter=&regressionLine=true&associationTable=false&x=slice%2Fexpression%2F2158%2Fentity_id&y=slice%2FGDSC2_AUC%2F64354%2Fentity_id&color= for analysis. [b]See https://depmap.org/portal/interactive/?filter=&regressionLine=true&associationTable=false&x=slice%2Fexpression%2F7317%2Fentity_id&y=slice%2FGDSC2_AUC%2F50287%2Fentity_id&color= for analysis. [c]See https://depmap.org/portal/interactive/?filter=&regressionLine=true&associationTable=false&x=slice%2Fexpression%2F38016%2Fentity_id&y=slice%2FGDSC1_AUC%2F50549%2Fentity_id&color= for analysis. [d]This DepMap release contains data from CRISPR knockout screens from project Achilles, as well as genomic characterization data from the CCLE project. For more information, please see README.txt.. figshare. Dataset. https://doi.org/10.6084/m9.figshare.22765112.v2. [e]For citation, see Garnett and coworkers.[1] [f]For citation, see McDermott, Garnett, and coworkers.[2] [g]p-Value (linear regression).

a. Ranked by Absolute Correlation (22.2 μM)

Rank	Gene	Correlation	q value	$-\log_{10}(q)$
1	*MGMT*	0.567	3.71E-69	68.4
2	*CLECL1*	-0.279	1.06E-12	12.0
3	*CTDSPL*	0.269	8.50E-12	11.1
4	*SP140*	-0.268	8.50E-12	11.1
5	*CCR7*	-0.265	1.41E-11	10.9
6	*SHROOM3*	0.263	2.20E-11	10.7
7	*ARHGAP15*	-0.262	2.20E-11	10.7
8	*GTF2IRD1*	0.258	5.11E-11	10.3
9	*GPR65*	-0.256	6.80E-11	10.2
10	*EPB41L1*	0.256	6.80E-11	10.2

b. Ranked by q Value (22.2 μM)

Rank	Gene	Correlation	q value	$-\log_{10}(q)$
1	*MGMT*	0.567	3.71E-69	68.4
2	*CLECL1*	-0.279	1.06E-12	12.0
3	*CTDSPL*	0.269	8.50E-12	11.1
4	*SP140*	-0.268	8.50E-12	11.1
5	*CCR7*	-0.265	1.41E-11	10.9
6	*SHROOM3*	0.263	2.20E-11	10.7
7	*ARHGAP15*	-0.262	2.20E-11	10.7
8	*GTF2IRD1*	0.258	5.11E-11	10.3
9	*EPB41L1*	0.256	6.80E-11	10.2
10	*PLEKHA1*	0.256	6.80E-11	10.2

c. Ranked by Absolute Correlation (66.67 μM)

Rank	Gene	Correlation	q value	$-\log_{10}(q)$
1	*MGMT*	0.639	2.63E-94	93.6
2	*PRKCZ*	0.264	3.39E-11	10.5
3	*CTDSPL*	0.264	3.39E-11	10.5
4	*SP140*	-0.263	3.39E-11	10.5
5	*TMEM139*	0.255	1.60E-10	9.80
6	*GLB1L2*	0.252	3.00E-10	9.52
7	*SHROOM3*	0.249	4.53E-10	9.34
8	*PLEKHA1*	0.248	5.97E-10	9.22
9	*CLECL1*	-0.245	1.01E-09	9.00
10	*FAM241B*	0.242	1.83E-09	8.74

d. Ranked by q Value (66.67 μM)

Rank	Gene	Correlation	q value	$-\log_{10}(q)$
1	*MGMT*	0.639	2.63E-94	93.6
2	*SP140*	-0.263	3.39E-11	10.5
3	*PRKCZ*	0.264	3.39E-11	10.5
4	*CTDSPL*	0.264	3.39E-11	10.5
5	*TMEM139*	0.255	1.60E-10	9.80
6	*GLB1L2*	0.252	3.00E-10	9.52
7	*SHROOM3*	0.249	4.53E-10	9.34
8	*PLEKHA1*	0.248	5.97E-10	9.22
9	*CLECL1*	-0.245	1.01E-09	9.00
10	*FAM241B*	0.242	1.83E-09	8.74

e. Ranked by Absolute Correlation ($\log_2$(AUC))

Rank	Gene	Correlation	q value	$-\log_{10}(q)$
1	*MGMT*	0.637	5.89E-93	68.4
2	*WNK2*	0.233	6.24E-08	12.0
3	*HTR7*	-0.22	6.18E-07	11.1
4	*BICDL1*	0.217	8.82E-07	11.1
5	*OBSCN*	0.208	2.43E-06	10.9
6	*PDE9A*	0.208	2.43E-06	10.7
7	*CACFD1*	0.207	2.43E-06	10.7
8	*RNF103*	0.207	2.43E-06	10.3
9	*ARHGEF16*	0.206	2.43E-06	10.2
10	*RAB11FIP4*	0.206	2.43E-06	10.2

f. Ranked by q Value ($\log_2$(AUC))

Rank	Gene	Correlation	q value	$-\log_{10}(q)$
1	*MGMT*	0.637	5.89E-93	68.4
2	*WNK2*	0.233	6.24E-08	12.0
3	*HTR7*	-0.22	6.18E-07	11.1
4	*BICDL1*	0.217	8.82E-07	11.1
5	*OBSCN*	0.208	2.43E-06	10.9
6	*PDE9A*	0.208	2.43E-06	10.7
7	*CACFD1*	0.207	2.43E-06	10.7
8	*RNF103*	0.207	2.43E-06	10.3
9	*ARHGEF16*	0.206	2.43E-06	10.2
10	*RAB11FIP4*	0.206	2.43E-06	10.2

g. Ranked by Absolute Correlation ($\log_2(IC_{50})$)

Rank	Gene	Correlation	q value	$-\log_{10}(q)$
1	*MGMT*	0.643	1.26E-95	68.4
2	*TMEM139*	0.249	2.19E-09	12.0
3	*ARHGEF16*	0.236	2.03E-08	11.1
4	*CAMSAP3*	0.235	2.03E-08	11.1
5	*PRKCZ*	0.233	2.94E-08	10.9
6	*PARD6B*	0.231	3.52E-08	10.7
7	*HSBP1L1*	0.23	3.60E-08	10.7
8	*ZNF468*	0.227	6.40E-08	10.3
9	*FAM83H*	0.225	7.23E-08	10.2
10	*LLGL2*	0.223	1.10E-07	10.2

h. Ranked by q Value ($\log_2(IC_{50})$)

Rank	Gene	Correlation	q value	$-\log_{10}(q)$
1	*MGMT*	0.643	1.26E-95	68.4
2	*TMEM139*	0.249	2.19E-09	12.0
3	*ARHGEF16*	0.236	2.03E-08	11.1
4	*CAMSAP3*	0.235	2.03E-08	11.1
5	*PRKCZ*	0.233	2.94E-08	10.9
6	*PARD6B*	0.231	3.52E-08	10.7
7	*HSBP1L1*	0.23	3.60E-08	10.7
8	*ZNF468*	0.227	6.40E-08	10.3
9	*FAM83H*	0.225	7.23E-08	10.2
10	*LLGL2*	0.223	1.10E-07	10.2

Fig. S2. Top 10 genes ranked by increasing absolute correlation (**a**, **c**, **e**, and **g**) or decreasing q value (**b**, **d**, **f**, and **h**) as determined by correlation analysis for gene expression vs. viability for KL-50 (**19**) as measured at 22.2 µM (**a** and **b**), at 66.67 µM (**c** and **d**), by $\log_2$(AUC) (**e** and **f**), or by $\log_2(IC_{50})$ (**g** and **h**) in the PRISM multiplexed viability screen. AUC = area under the curve.

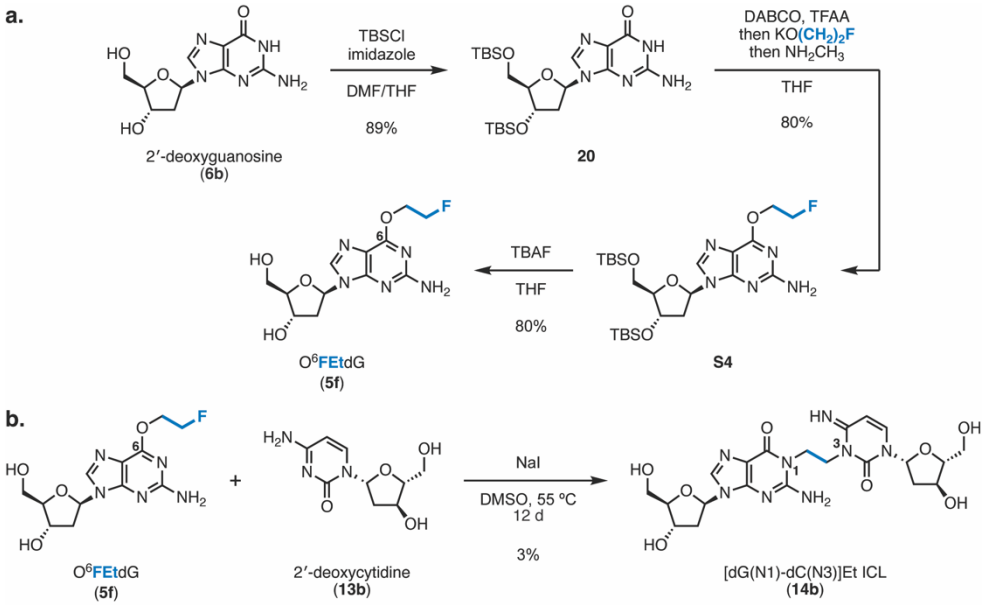
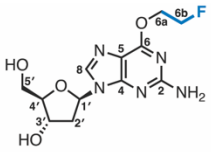


Fig. S3. Synthesis of synthetic standards of O^6-(2-fluoroethyl)-2′-deoxyguanosine (O^6FEtdG, **5f**) and the [dG(N1)-dC(N3)]Et ICL **14b**. **a.** Synthesis of O^6FEtdG (**5f**). **b.** Synthesis of the [dG(N1)-dC(N3)]Et ICL **14b**.



O^{6}FEtdG (**5f**)

Position	δ ^{1}H (Mult.)[a]	δ ^{13}C (Mult.)[b–c]	δ ^{19}F (Mult.)[d]	^{1}H-^{1}H gCOSY[e]	^{1}H-^{13}C gHMBC[f]
2	---	159.6 (s)	---	---	---
2-NH$_2$	6.47 (s)	---	---	---	2, 4
4	---	154.1	---	---	---
5	---	113.8	---	---	---
6	---	159.9	---	---	---
6a	4.65 (dt, $^3J_{HF}$ = 30.2, $^3J_{HH}$ = 3.9 Hz)	64.9 (d, $^2J_{CF}$ = 19.0 Hz)	---	6b	6
6b	4.79 (dt, $^2J_{HF}$ = 47.8, $^3J_{HH}$ = 4.0 Hz)	81.9 (d, $^1J_{CF}$ = 166.1 Hz)	–222.70 (tt, *J* = 48.0, 30.3 Hz)	6a	---
8	8.11 (s)	138.0	---	---	4, 5
1′	6.21 (*J* = dd, 7.8, 6.0 Hz)	82.8	---	2′x, 2′y	4, 8
2′ (^{1}H = x)[g]	2.21 (*J* = ddd, 13.2, 6.1, 3.1 Hz)	39.5[h]	---	1′ 2′y, 3′	3′, 4′
2′ (^{1}H = y)[g]	2.58 (ddd, *J* = 13.4, 7.9, 5.7 Hz)		---	1′ 2′x, 3′	1′, 3′, 4′
3′	4.35 (dq, *J* = 6.2, 3.1 Hz)	70.8	---	2′x, 2′y, 3′-OH, 4′	---
3′-OH	5.27 (d, *J* = 4.0 Hz)	---	---	3′	2′, 3′, 4′
4′	3.82 (td, *J* = 4.6, 2.7 Hz)	87.6	---	3′,5′x, 5′y	1′[i], 3′[i]
5′ (^{1}H = x)[f]	3.50 (dt, *J* = 11.7, 4.8 Hz)	61.7	---	4′, 5′y, 5′-OH	3′, 4′
5′ (^{1}H = y)[f]	3.57 (dt. *J* = 11.8, 5.1 Hz)		---	4′, 5′x, 5′-OH	3′, 4′
5′-OH	4.98 (t, *J* = 5.6 Hz)	---	---	5′x, 5′y	4′, 5′

Fig. S4. NMR assignment table for O^{6}FEtdG (**5f**). [a]Data acquired in DMSO-d_6 at 600 MHz; referenced to (CD$_3$)SO(CHD$_2$) (δ_H = 2.50). [b]Data acquired in DMSO-d_6 at 151 MHz with ^{1}H decoupling; referenced to DMSO-d_6 (δ_C = 39.52). [c]All ^{13}C signals were singlets unless otherwise indicated. [d]Data acquired in DMSO-d_6 at 376 MHz. [e]Data acquired in DMSO-d_6 at 600 MHz in both dimensions. [f]Data acquired in DMSO-d_6 at 600/151 MHz. [g]Arbitrary "x" and "y" designations used to distinguish diastereotopic protons attached to the same carbon. [h]Due to overlap with the DMSO-d_6 septet in the 1-D ^{13}C{^{1}H} spectrum, this ^{13}C NMR shift was estimated from an ^{1}H-^{13}C HSQC{^{1}H}{^{13}C} spectrum acquired in DMSO-d_6 at 600/151. [i]Relatively weak ^{1}H-^{13}C gHMBC correlation.

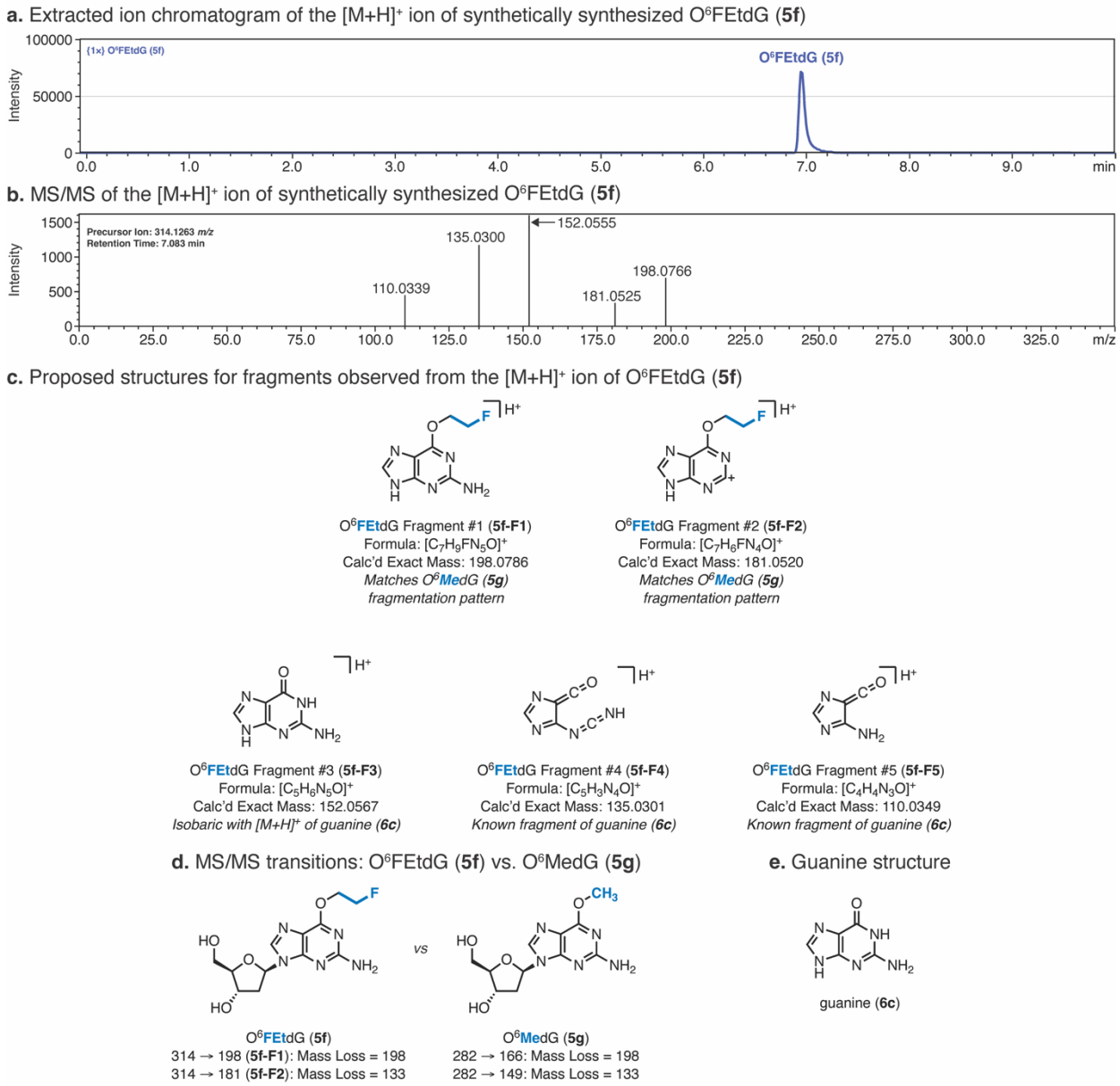
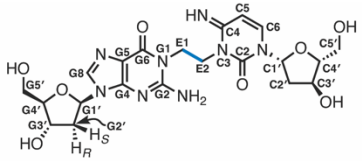


Fig. S5. LC/MS/MS characterization of synthetically prepared O^{6}FEtdG (**5f**). **a.** Positive mode extracted ion chromatogram (EIC) of synthetically prepared O^{6}FEtdG (**5f**). **b.** MS/MS fragmentation of the $[M+H]^+$ ion of synthetically prepared O^{6}FEtdG (**5f**). **c.** Proposed structures for masses observed from the fragmentation of the $[M+H]^+$ ion of O^{6}FEtdG (**5f**). Fragments #1 (**5f-F1**) and #2 (**5f-F2**) match the fragmentation pattern reported for O^{6}MedG (**5g**)[3]; see panel **d** for a side-by-side comparison. Fragments #4 (**5f-F4**) and #5 (**5f-F5**) match known fragments of guanine (**6c**, see panel **e** for structure).[4] For **a**, a positive mode EIC for the calculated *m/z* of the $[M+H]^+$ ion ± 20 ppm was plotted for O^{6}FEtdG (**5f**, $[M+H]^+$ calc'd for $[C_{12}H_{17}FN_5O_4]^+$ 314.1259); the intensity of the EIC is given in arbitrary units and was scaled by the factor given in curly brackets (i.e., {}). For **b**, data was acquired with a variable collision energy (CE) from 18.0 – 52.0 using the precursor ion 314.1263 *m/z* (retention time = 7.083 min). All data was acquired on LC/MS #1.



[dG(N1)–dC(N3)]Et ICL (**14b**)

Position	δ ^{1}H (Mult.)[a,b]	δ ^{13}C[a,c]	δ ^{15}N[a,d]	^{1}H-^{1}H gCOSY[a,e]	^{1}H-^{13}C gH2BC[15,14]	^{1}H-^{13}C gHMBC (25 °C)[a,f]	^{1}H-^{13}C gHMBC (65 °C)[f,g]	^{1}H-^{13}C LR-HSQMBC[a,f]	^{1}H-^{15}N gHMBC[a,h]	^{1}H-^{1}H ROESY[a,e]
E1	4.05 – 4.14 (br m, 2H)	36.8	---	---	E2	---	E2, G2, G6	---	---	---
E2	3.97 – 4.05 (br m, 2H)	38.5	---	---	E1	---	E1, C2, C4	---	---	---
G2	---	154.3[i]	---	---	---	---	---	---	---	---
G2-NH$_2$	7.83 (br s, 2H)	---	82.5	---	---	---	---	---	---	---
G4	---	149.3	---	---	---	---	---	---	---	---
G5	---	115.6	---	---	---	---	---	---	---	---
G6	---	156.4[i]	---	---	---	---	---	---	---	---
G7	---	---	172.4	---	---	---	---	---	---	---
G8	7.92 (s, 1H)	135.5	---	---	---	G4, G5	G2 or G6[j], G4, G5	G2, G4, G5, G6, G1′	G7, G9	G1′ G2′*S*, G3′
G9	---	---	249.5	---	---	---	---	---	---	---
G1′	6.12 (dd, *J* = 7.9, 6.0 Hz, 1H)	82.3	---	G2′*R*, G2′*S*	G2′	G4, G8	G4[k], G8	---	---	G8, G2′*R*, G2′*S*, G4′
G2′ (^{1}H = pro-*R*)	2.18 (ddd, *J* = 13.1, 6.0, 2.9 Hz, 1H)	39.5[m]	---	G1′, G2′S, G3′[n]	G1′	G3′, G4′	G3′	---	---	G1′, G2′*S*, G3′, G4′
G2′ (^{1}H = pro-*S*)	2.51 (1H)[l]		---	G1′ G2′*R*, G3′	G1′, G3′	G1′, G3′, G4′	G1′, G3′	---	G7	G8, G1′, G2′*R*, G3′
G3′	4.33 (dt, *J* = 5.6, 2.9 Hz, 1H)	70.8	---	G2′R[n], G2′S, G3′[n]	G2′, G4′	G1′, G5′	G1′[o]	---	---	G8, G2′R, G2′S, G4′
G4′	3.80 (td, *J* = 4.7, 2.6 Hz, 1H)	87.6	---	G3′[n], G5′[p]	G3′, G5′	G1′, G3′	---[q]	---	---	G1′, G2′*R*, G3′
G5′	3.47 – 3.58 (m, 2H)	61.7	---	G4′[p]	---[q]	---[q]	---[q]	---	---	---[r]
C1	---	---	138.7	---	---	---	---	---	---	---
C2	---	150.1	---	---	---	---	---	---	---	---
C4	---	157.6	---	---	---	---	---	---	---	---
C5	5.86 (d, *J* = 8.1 Hz, 1H)	101.0	---	C6	C6	C4, C6	C4, C6	C2	C1	C6
C6	7.34 (d, *J* = 8.1 Hz, 1H)	132.2	---	C5	C5	C2, C4, C6, C1′	C2, C4, C5, C1′	C2, C4, C5, C1′	C1	C5, C1′ C2′, C3′
C1′	6.16 (t, *J* = 6.9 Hz, 1H)	84.7	---	C2′	C2′	C2, C6	C2[l], C6	C2, C3′ C4′	---	C6, C2′, C4′
C2′	1.99 – 2.09 (m, 2H)	39.3[m]	---	C1′, C3′	C1′, C3′	C1′, C3′, C4′	C1′ C3′	---	C1	C6, C1′, C3′, C4′
C3′	4.21 (dt, *J* = 6.4, 3.4 Hz, 1H)	70.4	---	C2′, C4′	C2′, C4′	C1′, C5′	C1′[o]	---	---	C6, C2′, C4′
C4′	3.75 (q, *J* = 3.7 Hz, 1H)	87.2	---	C3′, C5′[p]	C3′, C5′	C3′	---[r]	C1′	---	C1′ C2′, C3′
C5′	3.47 – 3.58 (m, 2H)	61.4	---	C4′[p]	---[q]	---[q]	---[q]	---	---	---[r]

Fig. S6. NMR assignment table for [dG(N1)–dC(N3)]Et ICL (**14b**). [a]Data acquired at 25 ºC. [b]Data acquired in DMSO-d_6 at 600 MHz; referenced to $(CD_3)SO(CHD_2)$ (δ_H = 2.50). [c]Data acquired in DMSO-d_6 at 151 MHz with ^{1}H decoupling; referenced to DMSO-d_6 (δ_C = 39.52). [d]δ_N determined using a combination of ^{1}H-^{15}N HSQC[h] and ^{1}H-^{15}N gHMBC[h]. [e]Data acquired in DMSO-d_6 at 600 MHz in both dimensions. [f]Data acquired in DMSO-d_6 at 600/151 MHz. [g]Data acquired at 65 ºC. [h]Data acquired in DMSO-d_6 at 600/60 MHz. [j]Correlation from δ_H 7.92 (G8) to δ_C 156.4 (G2 or G6). [k]Assignment of these ^{1}H-^{13}C gHMBC correlations is tentative due to contour overlap. [l]Due to overlap with the $(CD_3)SO(CHD_2)$ pentet in the 1-D ^{1}H spectrum, no multiplicity is given for this peak and its integration is assumed to be 1H based on its assignment. [m]Due to overlap with the DMSO-d_6 septet in the 1-D ^{13}C{^{1}H} spectrum, this ^{13}C NMR shift was estimated from an ^{1}H-^{13}C HSQC{^{1}H}{^{13}C} spectrum[f]. [n]Cross-peak with G3′ in direct

dimension (f2) was weak. [o]Relatively weak ^{1}H-^{13}C gHMBC correlation. [p]Resolution insufficient to differentiate ^{1}H-^{1}H gCOSY cross-peaks to G5′/C5′; assignment of these ^{1}H-^{1}H gCOSY cross-peaks was aided by ^{1}H-^{13}C gH2BC. [q]Insufficient resolution to confidently assign ^{1}H-^{13}C gH2BC or ^{1}H-^{13}C gHMBC correlations from this proton. [r]Insufficient resolution to confidently assign ^{1}H-^{1}H ROESY cross-peaks involving G5′ and C5′.

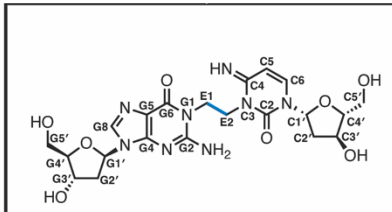
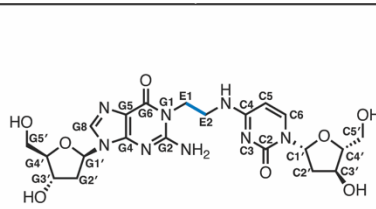
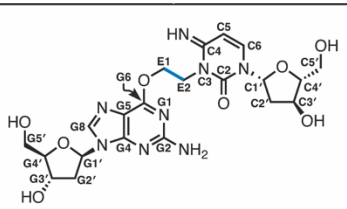
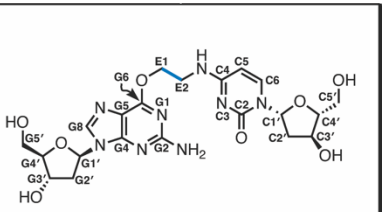
		 [dG(N1)–dC(N3)]Et ICL (**14b**)		 [dG(N1)–dC(N^4)]Et ICL (**S1**)		 [dG(O^6)–dC(N3)]Et ICL (**S2**)		 [dG(O^6)–dC(N^4)]Et ICL (**S3**)	
Position	**Experimental Data (δc)[a]**	**Calculated δc**	**Absolute Δδc[b,c]**	**Calculated δc**	**Absolute Δδc[b,c]**	**Calculated δc**	**Absolute Δδc[b,c]**	**Calculated δc**	**Absolute Δδc[b,c]**
E1	36.8	37.8	1.0	44.1	7.3	60.4	23.6	65.8	29.0
E2	38.5	36.8	1.7	38.4	0.1	38.7	0.2	43.3	4.8
G2	154.3[c]	156.9	2.6 (0.5)	157.2	2.9 (0.8)	162.4	8.1 (6.)	161.1	6.8 (4.7)
G4	149.3	146.9	2.4	147.9	1.4	154.5	5.2	153.3	4.0
G5	115.6	122.2	6.6	122.9	7.3	118.4	2.8	120.8	5.2
G6	156.4[c]	157.0	0.6 (2.7)	160.7	4.3 (7.3)	164.2	7.8 (9.9)	167.5	11.1 (13.2)
G8	135.5	134.6	0.9	135.8	0.3	140.4	4.9	138.1	2.6
G1′	82.3	86.3	4.0	87.8	5.5	88.9	6.6	90.0	7.7
G2′	39.5[d]	42.8	3.3	41.8	2.3	37.6	1.9	44.9	5.4
G3′	70.8	68.4	2.4	66.7	4.1	75.6	4.8	75.9	5.1
G4′	87.6	86.6	1.0	86.6	1.0	86.6	1.0	88.2	0.6
G5′	61.7	60.7	1.0	61.2	0.5	61.7	0.0	65.1	3.4
C2	150.1	153.6	3.5	157.7	7.6	153.3	3.2	156.3	6.2
C4	157.6	158.3	0.7	165.6	8.0	157.0	0.6	164.4	6.8
C5	101	99.4	1.6	90.1	10.9	101.2	0.2	88.9	12.1
C6	132.2	139.7	7.5	145.5	13.3	136.6	4.4	144.3	12.1
C1′	84.7	89.3	4.6	88.6	3.9	87.1	2.4	86.3	1.6
C2′	39.3[d]	40.5	1.2	39.7	0.4	39.1	0.2	39.8	0.5
C3′	70.4	65.2	5.2	65.6	4.8	65.3	5.1	65.1	5.3
C4′	87.2	86.9	0.3	84.0	3.2	87.3	0.1	84.8	2.4
C5′	61.4	61.6	0.2	61.7	0.3	59.2	2.2	60.2	1.2
Root Mean Square Deviation (RMSD):[c]		3.2 (3.2)		5.6 (5.6)		6.5 (6.5)		8.8 (8.9)	

Fig. S7. Comparison of experimental ^{13}C NMR data for [dG(N1)–dC(N3)]Et ICL **14b** to calculated ^{13}C NMR data for dG(N1)–dC(N3)]Et ICL **14b** as well as the isomeric deoxyribonucleoside ethyl crosslinks [dG(N1)–dC(N^4)]Et ICL (**S1**), [dG(O^6)–dC(N3)]Et ICL (**S2**), and [dG(O^6)–dC(N^4)]Et ICL (**S3**). Calculations performed using Spartan 20.[5] [a]Experimental data acquired at 151 MHz in DMSO-d_6 at 25 ºC from a 1–D ^{13}C{^{1}H} NMR experiment unless otherwise indicated. [b]Absolute $\Delta\delta_C$ = absolute difference between δ_C (experimental) and δ_C (calculated). [c]The δ_C assignments for G2 and G6 are interchangeable. Values in the table were calculated assuming δ_C (G2) = 154.3 and δ_C (G6) = 156.4; values in parenthesis were calculated using the inverted assignments (i.e., δ_C (G2) = 156.4 and δ_C (G6) = 156.4). [d]Due to overlap with the DMSO-d_6 septet in the 1-D ^{13}C{^{1}H} spectrum, this ^{13}C NMR shift was estimated from an ^{1}H-^{13}C HSQC{^{1}H}{^{13}C} spectrum acquired in DMSO-d_6 at 600/151 MHz at 25 ºC.

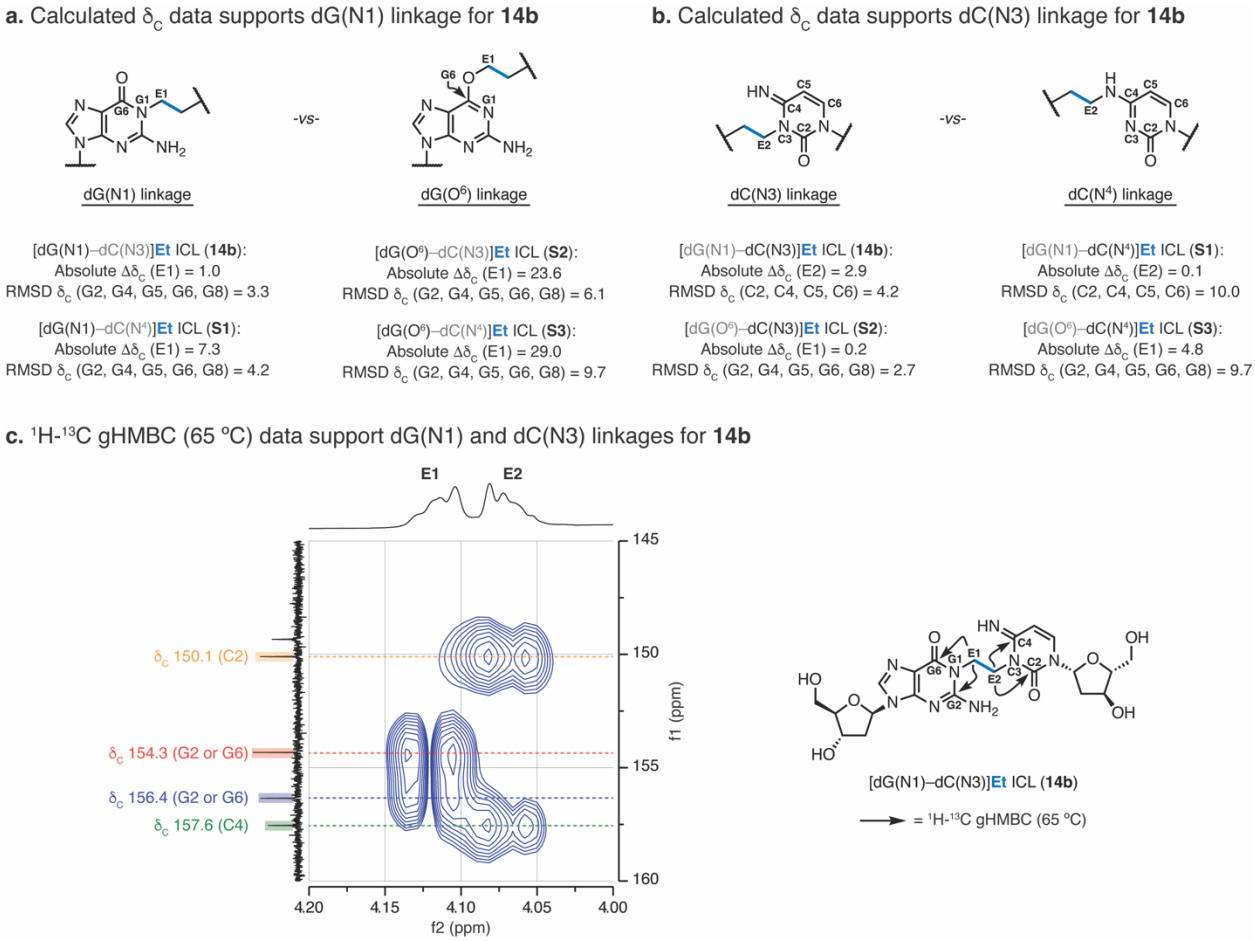


Fig. S8. Experimental and calculated NMR data support the assignment of the ethyl linkage of [dG(N1)–dC(N3)]Et ICL (**14b**) to the dG(N1) and dC(N3) positions. **a.** Comparison of experimental $^{13}C\{^{1}H\}$ NMR data (151 MHz, DMSO-d_6, 25 ºC) to calculated ^{13}C NMR shifts supports the assignment of the ethyl linker to the dG(N1) position. **b.** Comparison of experimental $^{13}C\{^{1}H\}$ (151 MHz, DMSO-d_6, 25 ºC) NMR data to calculated ^{13}C NMR shifts supports the assignment of the ethyl linker to the dC(N3) position. **c.** Experimental ^{1}H-^{13}C gHMBC (600/151 MHz, DMSO-d_6, 65 ºC) NMR data shows key gHMBC correlations from δ_H (E1) → δ_C (G2 & G6) and from δ_H (E2) → δ_C (C2 & C4).

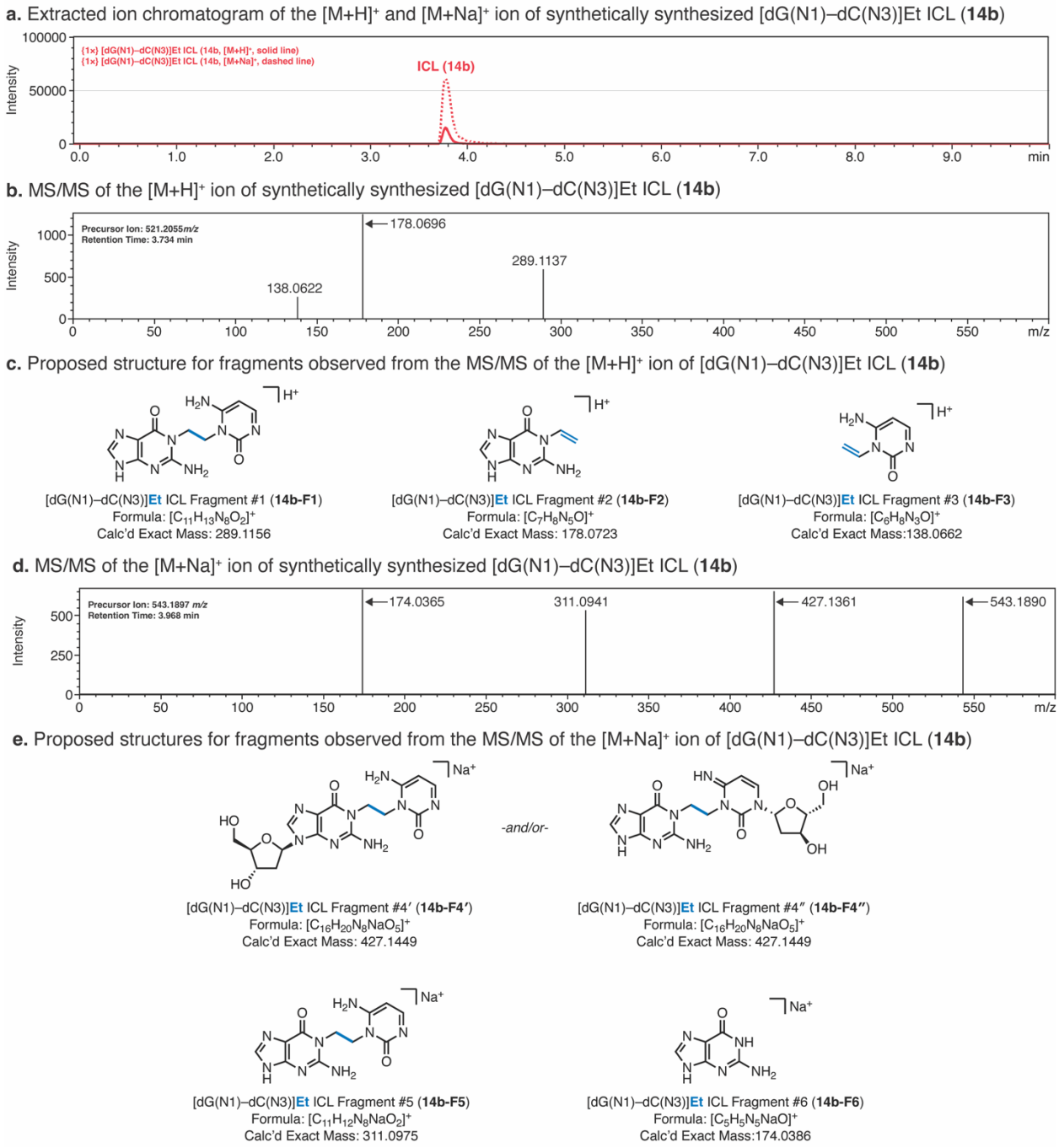


Fig. S9. LC/MS/MS characterization of synthetically prepared [dG(N1)–dC(N3)]Et ICL (**14b**). **a.** Positive mode extracted ion chromatograms (EICs) of the $[M+H]^+$ (solid line) and $[M+Na]^+$ (dashed line) ions of synthetically prepared [dG(N1)–dC(N3)]Et ICL (**14b**). **b.** MS/MS fragmentation of the $[M+H]^+$ ion of synthetically prepared [dG(N1)–dC(N3)]Et ICL (**14b**). **c.** Proposed structures for masses observed from the fragmentation of the $[M+H]^+$ ion of synthetically prepared [dG(N1)–dC(N3)]Et ICL (**14b**). **d.** MS/MS fragmentation of the $[M+Na]^+$ ion of synthetically prepared [dG(N1)–dC(N3)]Et ICL (**14b**). **e.** Proposed structures for masses observed from the fragmentation of the $[M+Na]^+$ ion of synthetically prepared [dG(N1)–dC(N3)]Et ICL (**14b**). For **a**, positive mode EICs for the calculated *m/z* of the $[M+H]^+$ and $[M+Na]^+$ ions ± 20

ppm were plotted for [dG(N1)–dC(N3)]Et ICL (**14b**, $[M+H]^+$ calc'd for $[C_{21}H_{29}N_8O_8]^+$ 521.2103; $[M+Na]^+$ calc'd for $[C_{21}H_{28}N_8NaO_8]^+$ 543.1922); the intensity of the EICs is given in arbitrary units and was scaled by the factor given in curly brackets (i.e., {}). For **b** and **d**, MS/MS data were acquired with a variable collision energy (CE) from 18.0 – 52.0 using the precursor ion 521.2055 *m/z* (retention time = 3.734 min) for **b** and the precursor ion 543.1897 *m/z* (retention time = 3.968 min) for **d**; intensity is given in arbitrary units. All data were acquired on LC/MS #2.

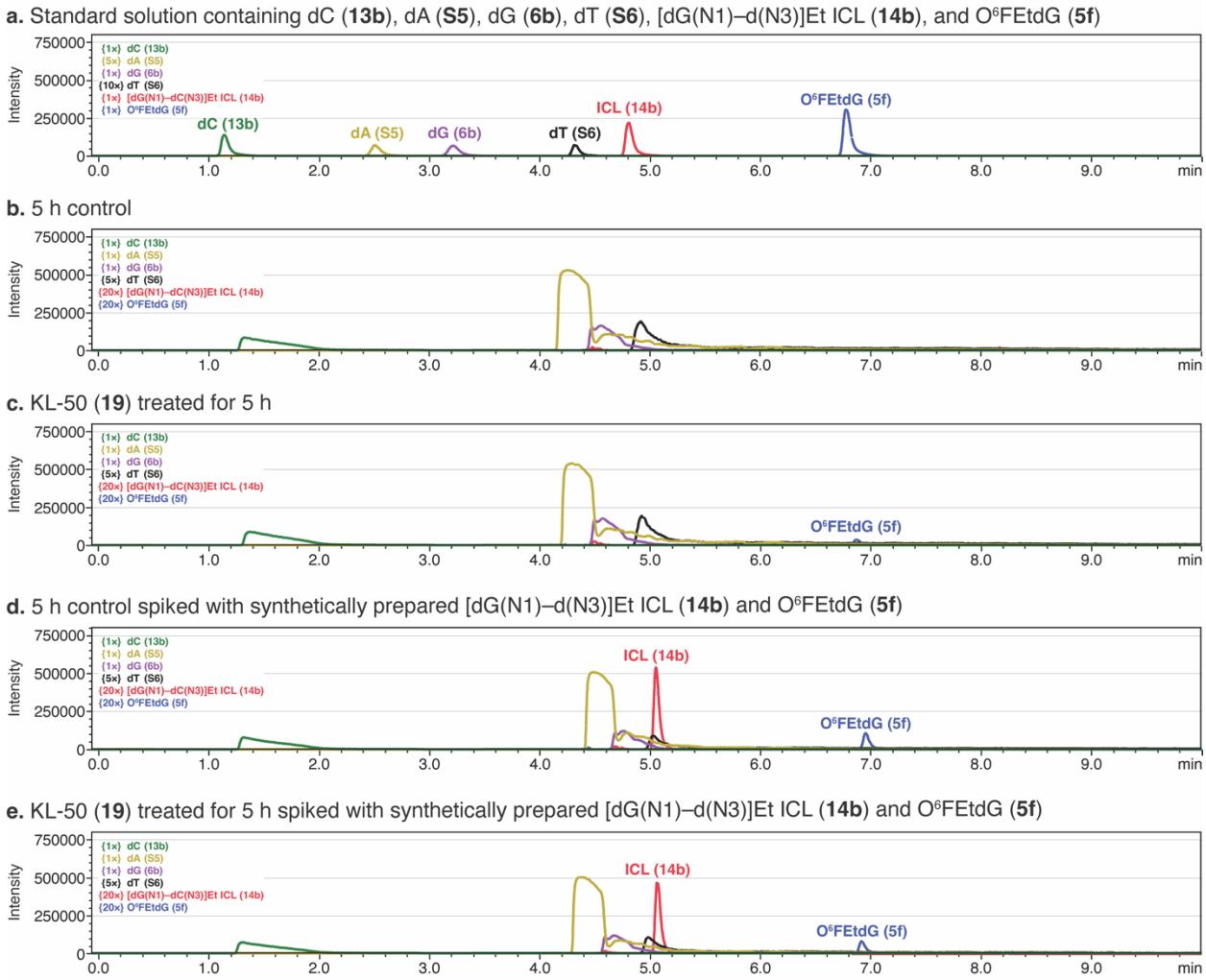


Fig. S10. Positive mode extracted ion chromatograms (EICs) for the four canonical DNA bases as well as [dG(N1)–dC(N3)]Et ICL (**14b**) and O^{6}FEtdG (**5f**) lesions from the nucleoside pools resulting from the enzymatic digestion of calf thymus DNA (ctDNA) incubated with vehicle or KL-50 (**19**) for 5 h. **a.** EICs for a standard solution of commercially available DNA bases (2′-deoxycytidine (dC, **13b**), 2′-deoxyadenosine (dA, **S5**), 2′-deoxyguanosine (dG, **6b**), and 2′-deoxythymidine (dT, **S6**)) and synthetically prepared [dG(N1)–dC(N3)]Et ICL (**14b**) and O^{6}FEtdG (**5f**). **b.** EICs of the nucleoside pool resulting from the enzymatic digestion of ctDNA incubated with vehicle for 5 h. **c.** EICs of the nucleoside pool resulting from the enzymatic digestion of ctDNA incubated with KL-50 (**19**, 1 mM) for 5 h. **d.** EICs of a mixture of the nucleoside pool resulting from the enzymatic digestion of ctDNA incubated with vehicle for 5 h and a standard solution of synthetically prepared [dG(N1)–dC(N3)]Et ICL (**14b**) and O^{6}FEtdG (**5f**). **e.** EICs of a mixture of the nucleoside pool resulting from the enzymatic digestion of ctDNA incubated with KL-50 (**19**, 1 mM) for 5 h and a standard solution of synthetically prepared [dG(N1)–dC(N3)]Et ICL (**14b**) and O^{6}FEtdG (**5f**). For **a**–**e**, a positive mode EIC for the calculated *m*/*z* of the $[M+H]^+$ ion ± 20 ppm was plotted for each species; the intensity of each EIC is given in arbitrary units and was scaled by the factor given in curly brackets (i.e., {}). The calculated *m*/*z*'s for the $[M+H]^+$ ion of each species are as follows: $[M+H]^+$ for dC (**13b**) calc'd for $[C_9H_{14}N_3O_4]^+$ 228.0979; $[M+H]^+$ for dA (**S5**) calc'd for $[C_{10}H_{14}N_5O_3]^+$ 252.1091; $[M+H]^+$ for dG

(**6b**) calc'd for $[C_{10}H_{14}N_5O_4]^+$ 268.1040; $M+H]^+$ for dT (**S6**) calc'd for $[C_{10}H_{15}N_2O_5]^+$ 243.0975; $[M+H]^+$ for [dG(N1)–dC(N3)]Et ICL (**14b**) calc'd for $[C_{21}H_{29}N_8O_8]^+$ 521.2103; $[M+H]^+$ for O^6FEtdG (**5f**) for $[C_{12}H_{17}FN_5O_4]^+$ 314.1259. All data were acquired on LC/MS #1.

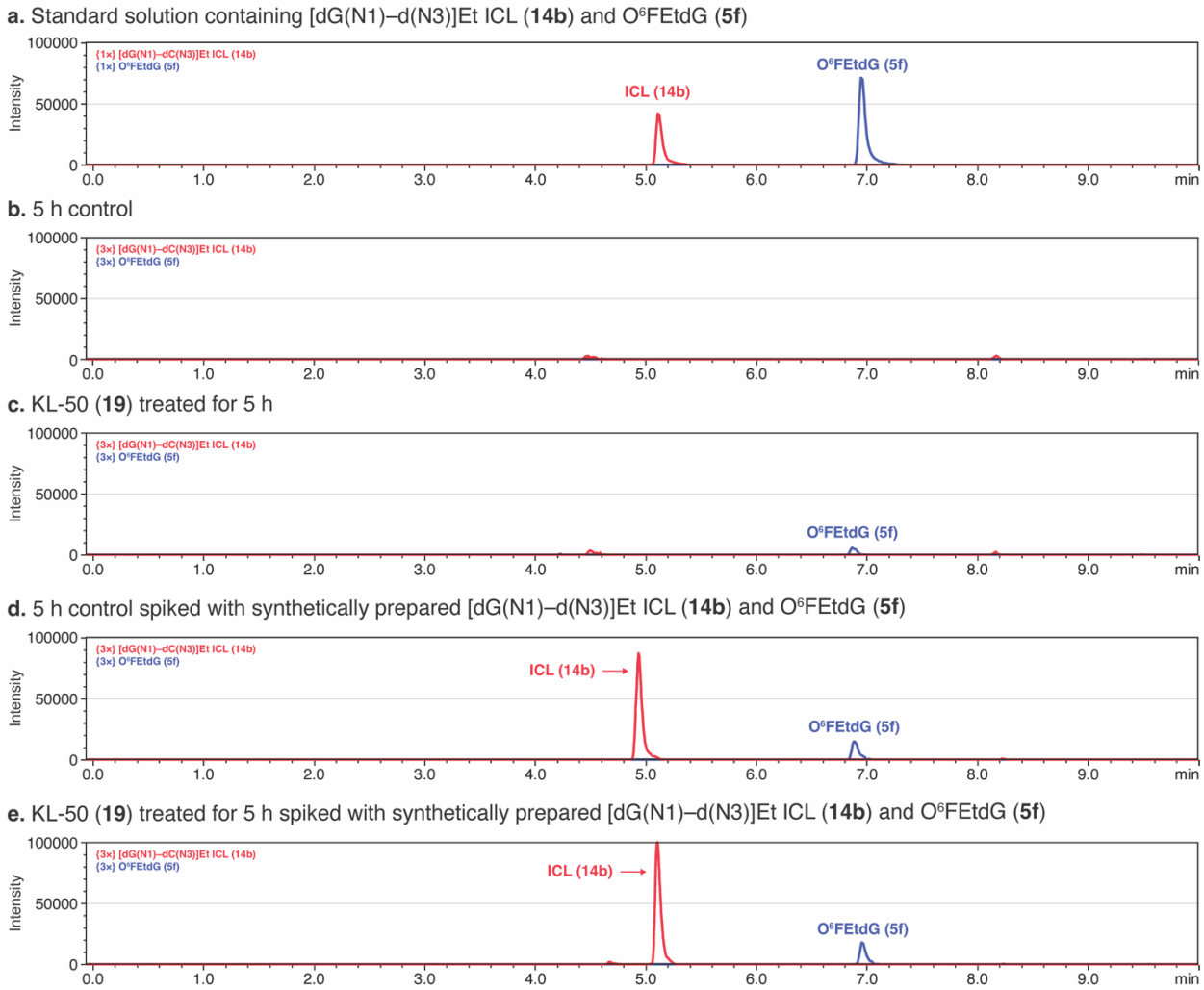


Fig. S11. Positive mode extracted ion chromatograms (EICs) for the [dG(N1)–dC(N3)]Et ICL (**14b**) and O^{6}FEtdG (**5f**) lesions from the nucleoside pools resulting from the enzymatic digestion of calf thymus DNA (ctDNA) incubated with vehicle or KL-50 (**19**) for 5 h. **a.** EICs for a standard solution of synthetically prepared [dG(N1)–dC(N3)]Et ICL (**14b**) and O^{6}FEtdG (**5f**). **b.** EICs of the nucleoside pool resulting from the enzymatic digestion of ctDNA incubated with vehicle for 5 h. **c.** EICs of the nucleoside pool resulting from the enzymatic digestion of ctDNA incubated with KL-50 (**19**, 1 mM) for 5 h. **d.** EICs of a mixture of the nucleoside pool resulting from the enzymatic digestion of ctDNA incubated with vehicle for 5 h and a standard solution of synthetically prepared [dG(N1)–dC(N3)]Et ICL (**14b**) and O^{6}FEtdG (**5f**). **e.** EICs of a mixture of the nucleoside pool resulting from the enzymatic digestion of ctDNA incubated with KL-50 (**19**, 1 mM) for 5 h and a standard solution of synthetically prepared [dG(N1)–dC(N3)]Et ICL (**14b**) and O^{6}FEtdG (**5f**). For **a**–**e**, a positive mode EIC for the calculated *m*/*z* of the $[M+H]^+$ ion ± 20 ppm was plotted for each species; the intensity of each EIC is given in arbitrary units and was scaled by the factor given in curly brackets (i.e., {}). The calculated *m*/*z*'s for the $[M+H]^+$ ion of each species are as follows: $[M+H]^+$ for [dG(N1)–dC(N3)]Et ICL (**14b**) calc'd for $[C_{21}H_{29}N_8O_8]^+$ 521.2103; $[M+H]^+$ for O^{6}FEtdG (**5f**) for $[C_{12}H_{17}FN_5O_4]^+$ 314.1259. All data were acquired on LC/MS #1.

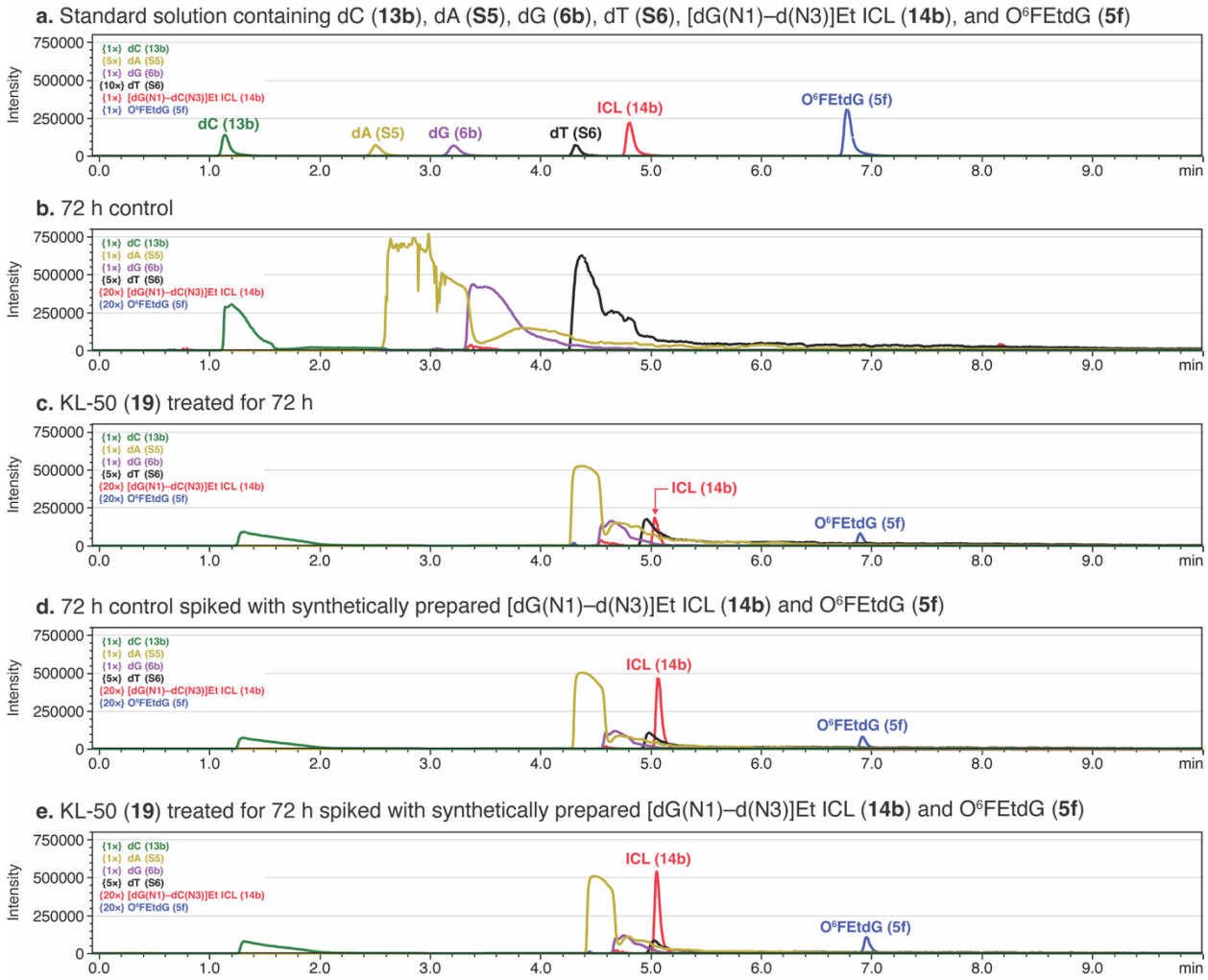


Fig. S12. Positive mode extracted ion chromatograms (EICs) for the four canonical DNA bases as well as [dG(N1)–dC(N3)]Et ICL (**14b**) and O^{6}FEtdG (**5f**) lesions from the nucleoside pools resulting from the enzymatic digestion of calf thymus DNA (ctDNA) incubated with vehicle or KL-50 (**19**) for 72 h. **a.** EICs for a standard solution of commercially available DNA bases (2′-deoxycytidine (dC, **13b**), 2′-deoxyadenosine (dA, **S5**), 2′-deoxyguanosine (dG, **6b**), and 2′-deoxythymidine (dT, **S6**)) and synthetically prepared [dG(N1)–dC(N3)]Et ICL (**14b**) and O^{6}FEtdG (**5f**). **b.** EICs of the nucleoside pool resulting from the enzymatic digestion of ctDNA incubated with vehicle for 72 h. **c.** EICs of the nucleoside pool resulting from the enzymatic digestion of ctDNA incubated with KL-50 (**19**, 1 mM) for 72 h. **d.** EICs of a mixture of the nucleoside pool resulting from the enzymatic digestion of ctDNA incubated with vehicle for 72 h and a standard solution of synthetically prepared [dG(N1)–dC(N3)]Et ICL (**14b**) and O^{6}FEtdG (**5f**). **e.** EICs of a mixture of the nucleoside pool resulting from the enzymatic digestion of ctDNA incubated with KL-50 (**19**, 1 mM) for 72 h and a standard solution of synthetically prepared [dG(N1)–dC(N3)]Et ICL (**14b**) and O^{6}FEtdG (**5f**). For **a**–**e**, a positive mode EIC for the calculated *m*/*z* of the $[M+H]^+$ ion ± 20 ppm was plotted for each species; the intensity of each EIC is given in arbitrary units and was scaled by the factor given in curly brackets (i.e., {}). The calculated *m*/*z*'s for the $[M+H]^+$ ion of each species are as follows: $[M+H]^+$ for dC (**13b**) calc'd for $[C_9H_{14}N_3O_4]^+$ 228.0979; $[M+H]^+$ for dA (**S5**) calc'd for $[C_{10}H_{14}N_5O_3]^+$ 252.1091; $[M+H]^+$ for dG

(**6b**) calc'd for $[C_{10}H_{14}N_5O_4]^+$ 268.1040; $M+H]^+$ for dT (**S6**) calc'd for $[C_{10}H_{15}N_2O_5]^+$ 243.0975; $[M+H]^+$ for [dG(N1)–dC(N3)]Et ICL (**14b**) calc'd for $[C_{21}H_{29}N_8O_8]^+$ 521.2103; $[M+H]^+$ for O^6FEtdG (**5f**) for $[C_{12}H_{17}FN_5O_4]^+$ 314.1259. All data were acquired on LC/MS #1.

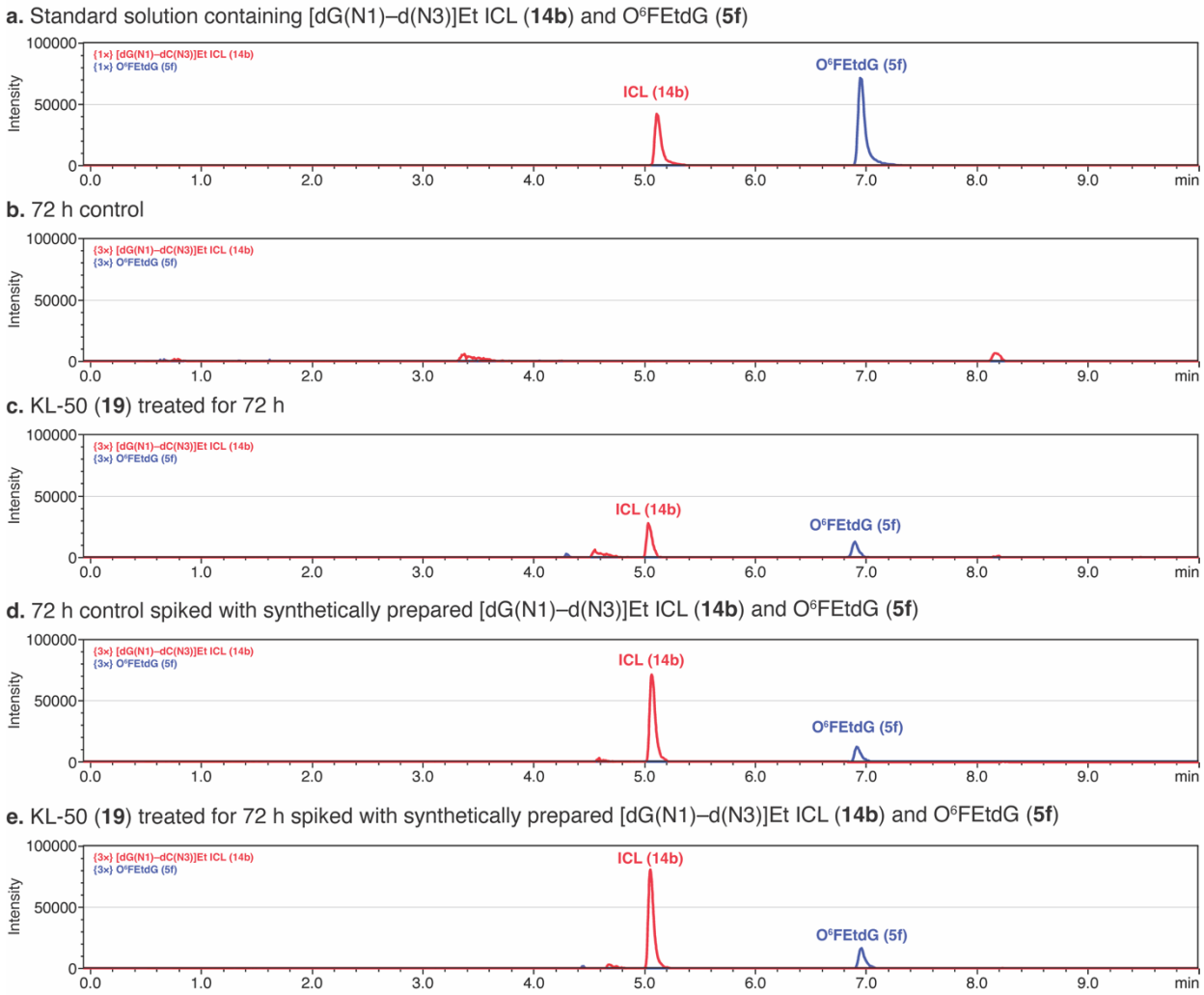


Fig. S13. Positive mode extracted ion chromatograms (EICs) for the [dG(N1)–dC(N3)]Et ICL (**14b**) and O^{6}FEtdG (**5f**) lesions from the nucleoside pools resulting from the enzymatic digestion of calf thymus DNA (ctDNA) incubated with vehicle or KL-50 (**19**) for 72 h. **a.** EICs for a standard solution of synthetically prepared [dG(N1)–dC(N3)]Et ICL (**14b**) and O^{6}FEtdG (**5f**). **b.** EICs of the nucleoside pool resulting from the enzymatic digestion of ctDNA incubated with vehicle for 72 h. **c.** EICs of the nucleoside pool resulting from the enzymatic digestion of ctDNA incubated with KL-50 (**19**, 1 mM) for 72 h. **d.** EICs of a mixture of the nucleoside pool resulting from the enzymatic digestion of ctDNA incubated with vehicle for 72 h and a standard solution of synthetically prepared [dG(N1)–dC(N3)]Et ICL (**14b**) and O^{6}FEtdG (**5f**). **e.** EICs of a mixture of the nucleoside pool resulting from the enzymatic digestion of ctDNA incubated with KL-50 (**19**, 1 mM) for 72 h and a standard solution of synthetically prepared [dG(N1)–dC(N3)]Et ICL (**14b**) and O^{6}FEtdG (**5f**). For **a**–**e**, a positive mode EIC for the calculated *m*/*z* of the $[M+H]^+$ ion ± 20 ppm was plotted for each species; the intensity of each EIC is given in arbitrary units and was scaled by the factor given in curly brackets (i.e., {}). The calculated *m*/*z*'s for the $[M+H]^+$ ion of each species are as follows: $[M+H]^+$ for [dG(N1)–dC(N3)]Et ICL (**14b**) calc'd for $[C_{21}H_{29}N_8O_8]^+$ 521.2103; $[M+H]^+$ for O^{6}FEtdG (**5f**) for $[C_{12}H_{17}FN_5O_4]^+$ 314.1259. All data were acquired on LC/MS #1.

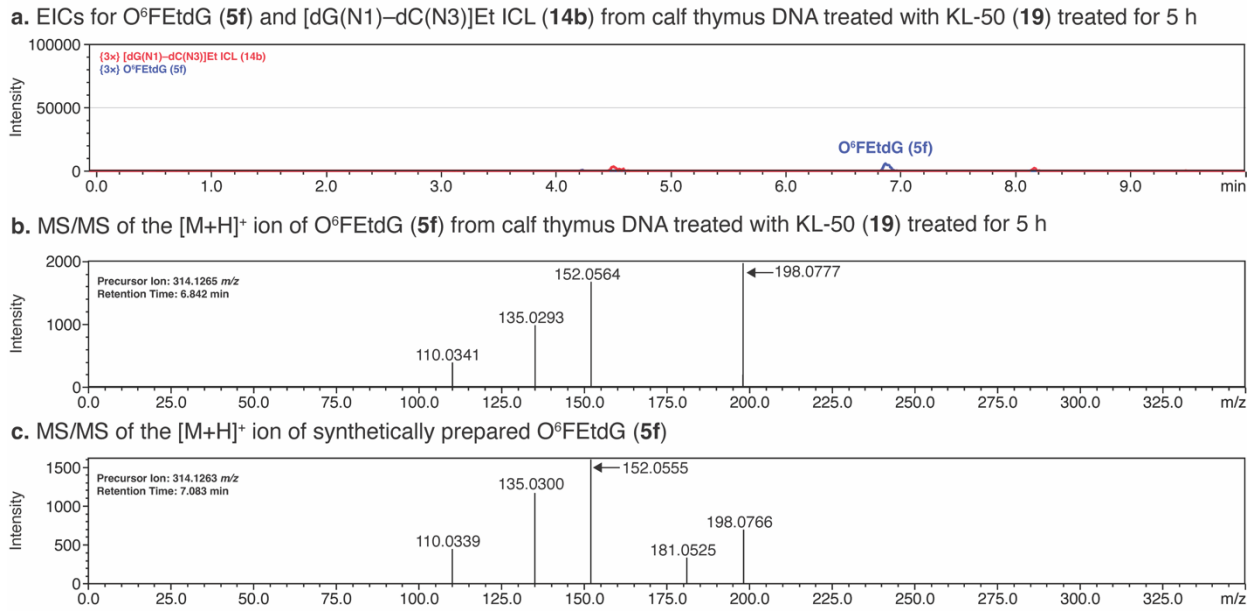


Fig. S14. LC/MS/MS characterization of the O^{6}FEtdG (**5f**) lesion resulting from the enzymatic digestion of calf thymus DNA (ctDNA) incubated with KL-50 (**19**) for 5 h. **a.** Positive mode extracted ion chromatogram (EIC) of the nucleoside pool resulting from the enzymatic digestion of ctDNA incubated with KL-50 (**19**, 1 mM) for 5 h. **b.** MS/MS fragmentation of the $[M+H]^+$ ion of O^{6}FEtdG (**5f**) observed in ctDNA incubated with KL-50 (**19**, 1 mM) for 5 h. **c.** MS/MS fragmentation of the $[M+H]^+$ ion of synthetically prepared O^{6}FEtdG (**5f**). For **a**, a positive mode EIC for the calculated *m/z* of the $[M+H]^+$ ion ± 20 ppm was plotted for O^{6}FEtdG (**5f**, $[M+H]^+$ calc'd for $[C_{12}H_{17}FN_5O_4]^+$ 314.1259); the intensity of the EIC is given in arbitrary units and was scaled by the factor given in curly brackets (i.e., {}). For **b**–**c**, MS/MS data were acquired with a variable collision energy (CE) from 18.0 – 52.0 using the precursor ion 314.1265 *m/z* (retention time = 6.842 min) for **b** and the precursor ion 314.1263 *m/z* (retention time = 7.083 min) for **c**. See Fig. S5c for proposed structures for masses observed from the fragmentation of the $[M+H]^+$ ion of O^{6}FEtdG (**5f**). All data were acquired on LC/MS #1.

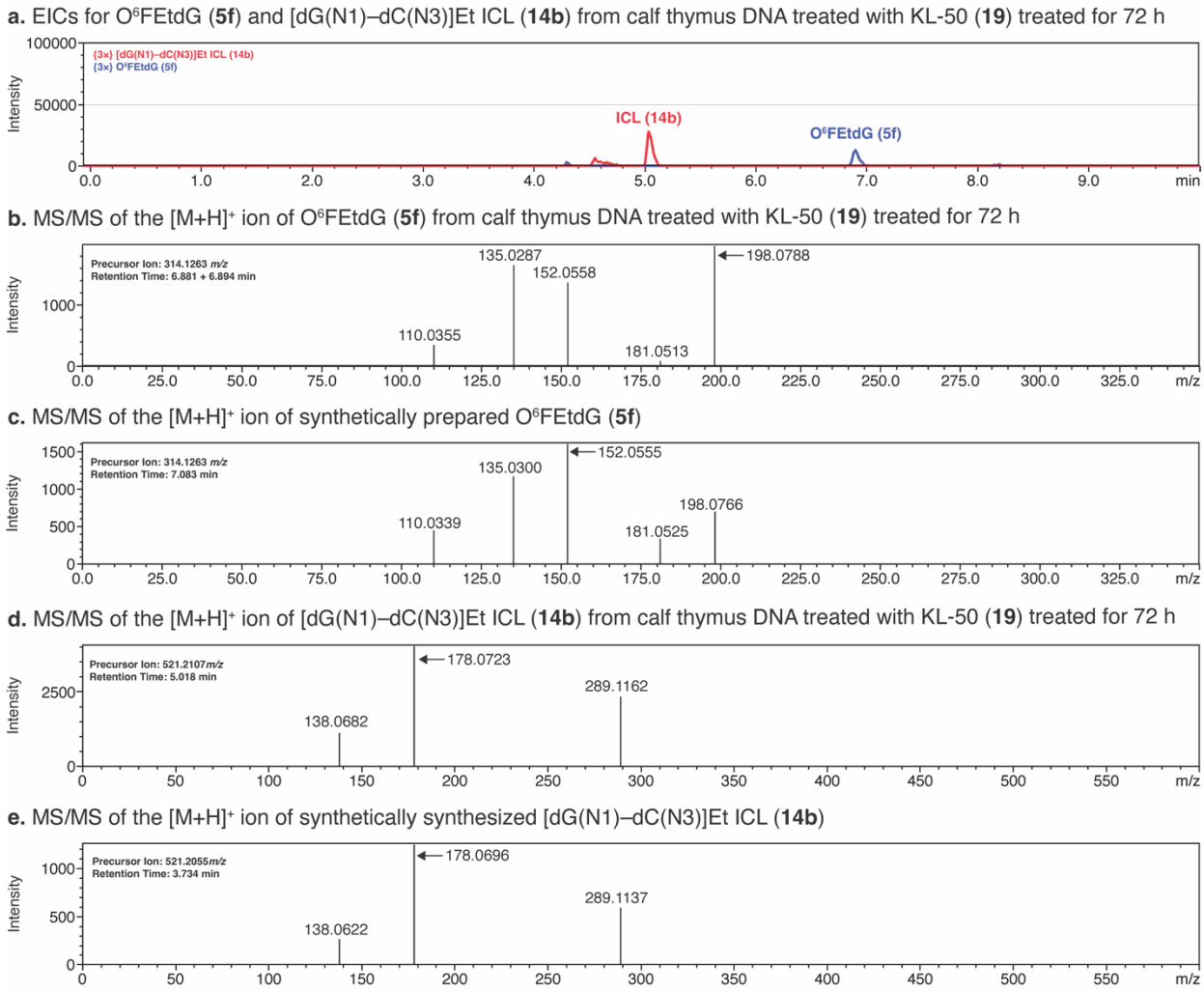


Fig. S15. LC/MS/MS characterization of the O^{6}FEtdG (**5f**) and [dG(N1)–dC(N3)]Et ICL (**14b**) lesions resulting from the enzymatic digestion of calf thymus DNA (ctDNA) incubated with KL-50 (**19**) for 72 h. **a.** Positive mode extracted ion chromatogram (EIC) of the nucleoside pool resulting from the enzymatic digestion of ctDNA incubated with KL-50 (**19**, 1 mM) for 72 h. **b.** MS/MS fragmentation of the $[M+H]^+$ ion of O^{6}FEtdG (**5f**) observed in ctDNA incubated with KL-50 (**19**, 1 mM) for 72 h. **c.** MS/MS fragmentation of the $[M+H]^+$ ion of synthetically prepared O^{6}FEtdG (**5f**). **d.** MS/MS fragmentation of the $[M+H]^+$ ion of [dG(N1)–dC(N3)]Et ICL (**14b**) observed in ctDNA incubated with KL-50 (**19**, 1 mM) for 72 h. **e.** MS/MS fragmentation of the $[M+H]^+$ ion of synthetically prepared [dG(N1)–dC(N3)]Et ICL (**14b**). For **a**, a positive mode EIC for the calculated *m/z* of the $[M+H]^+$ ion ± 20 ppm was plotted for each species; the intensity of each EIC is given in arbitrary units and was scaled by the factor given in curly brackets (i.e., {}). The calculated *m/z*'s for the $[M+H]^+$ ion of each species are as follows: $[M+H]^+$ for [dG(N1)–dC(N3)]Et ICL (**14b**) calc'd for $[C_{21}H_{29}N_8O_8]^+$ 521.2103; $[M+H]^+$ for O^{6}FEtdG (**5f**) calc'd for $[C_{12}H_{17}FN_5O_4]^+$ 314.1259. For **b**–**e**, MS/MS data were acquired with a variable collision energy (CE) from 18.0 – 52.0 using the precursor ion 314.1263 *m/z* (retention time = 6.881 + 6.894 min) for **b**, the precursor ion 314.1263 *m/z* (retention time = 7.083 min) for **c**, the precursor ion 521.2107 *m/z* (retention time = 5.018 min) for **d**, and the precursor ion 521.2055 *m/z* (retention time = 3.734 min) for **e**; intensity is given in arbitrary units. See Fig. S5c for proposed structures for masses

observed from the fragmentation of the $[M+H]^+$ ion of O^6FEtdG (**5f**); see Fig. 9c for proposed structures for masses observed from the fragmentation of the $[M+H]^+$ ion of [dG(N1)–dC(N3)]Et ICL (**14b**). The data in **a**–**d** were acquired on LC/MS #1 and the data in **e** was acquired on LC/MS #2; however, for both setups, the LC systems fed into the same HRMS instrument.

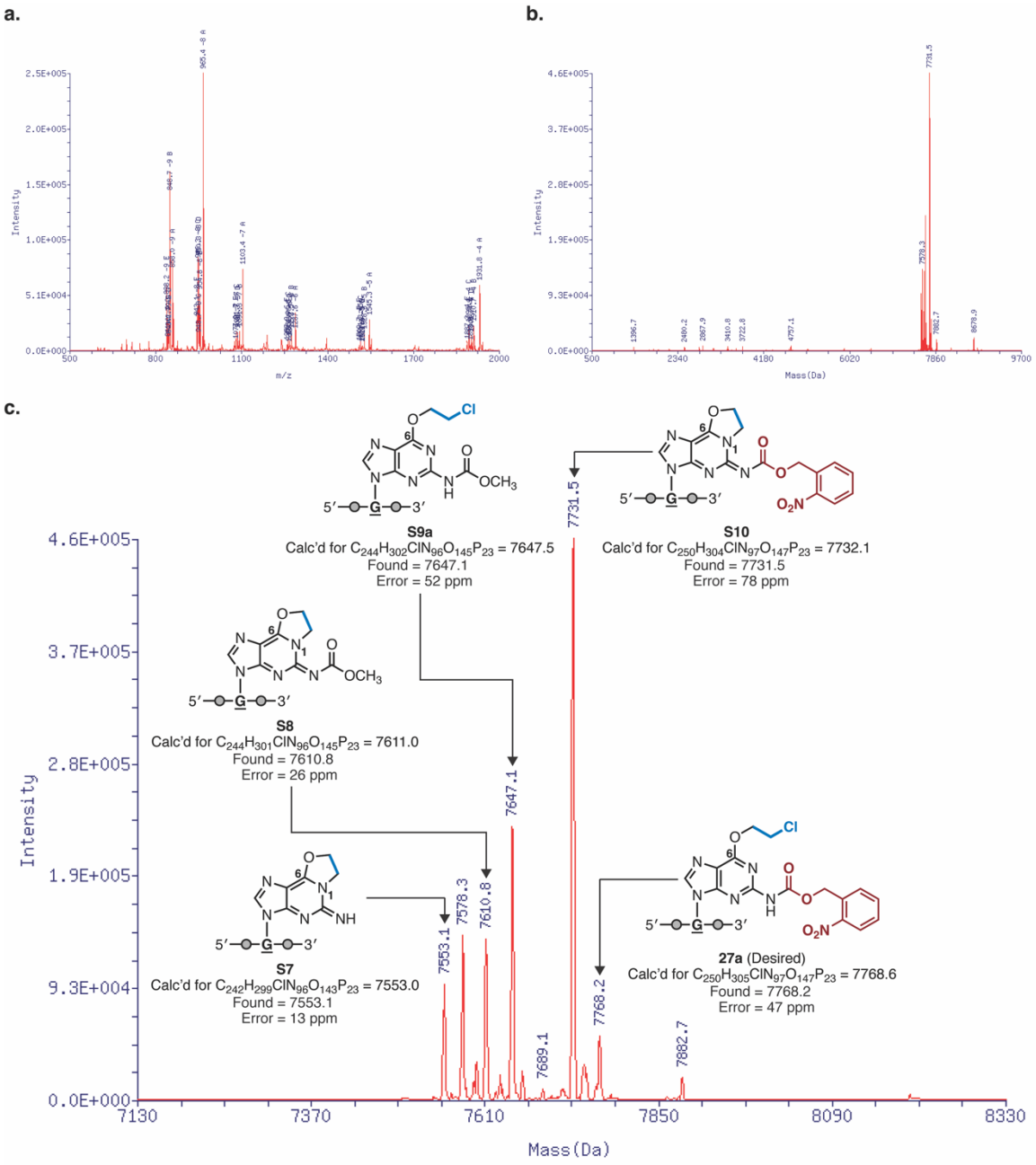


Fig. S16. Mass spectrometry analysis of the O^6ClEtG containing deoxyoligonucleotide **27a** following cleavage from the succinyl resin using potassium carbonate and methanol. **a.** Annotated mass spectrum (ESI, negative mode). **b.** Deconvoluted mass spectrum. **c.** Expansion of the deconvoluted mass spectrum with peak assignments. Sequence = 5′-d(GATGTGGTATA**G**ATAAGGAGTTGT)-3′ where **G** = modified guanine. For **c**, calculated masses are molecular weights and the structures of **S7–S10** are proposed only based on mass spectral data.

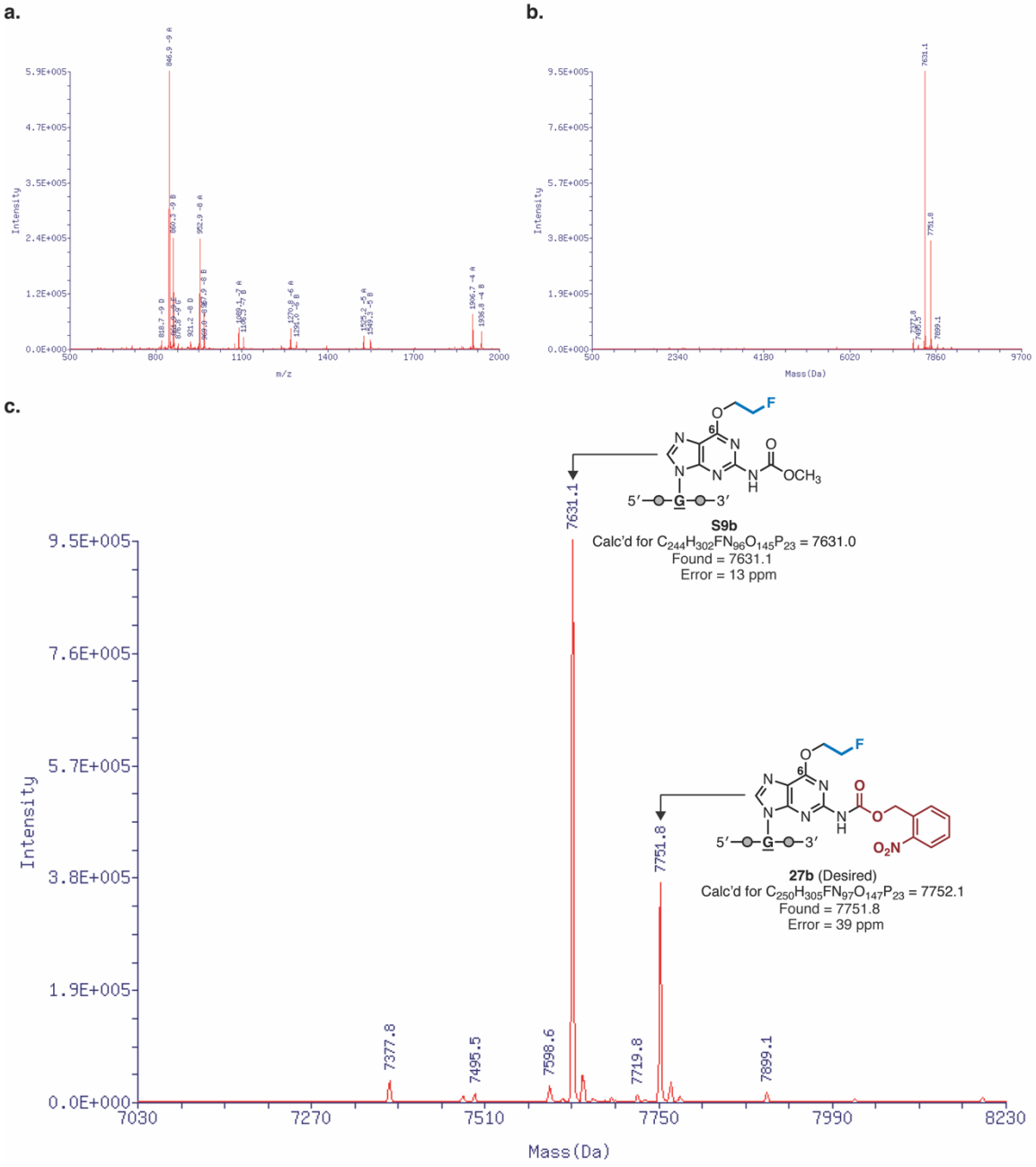


Fig. S17. Mass spectrometry analysis of the O^6FEtG containing deoxyoligonucleotide **27b** cleaved from a succinyl resin using potassium carbonate and methanol. **a.** Annotated mass spectrum (ESI, negative mode). **b.** Deconvoluted mass spectrum. **c.** Expansion of the deconvoluted mass spectrum with peak assignments. Sequence = 5′-d(GATGTGGTATA**G**ATAAGGAGTTGT)-3′ where **G** = modified guanine. For **c**, calculated masses are molecular weights and the structure of **S9b** is proposed only based on mass spectral data.

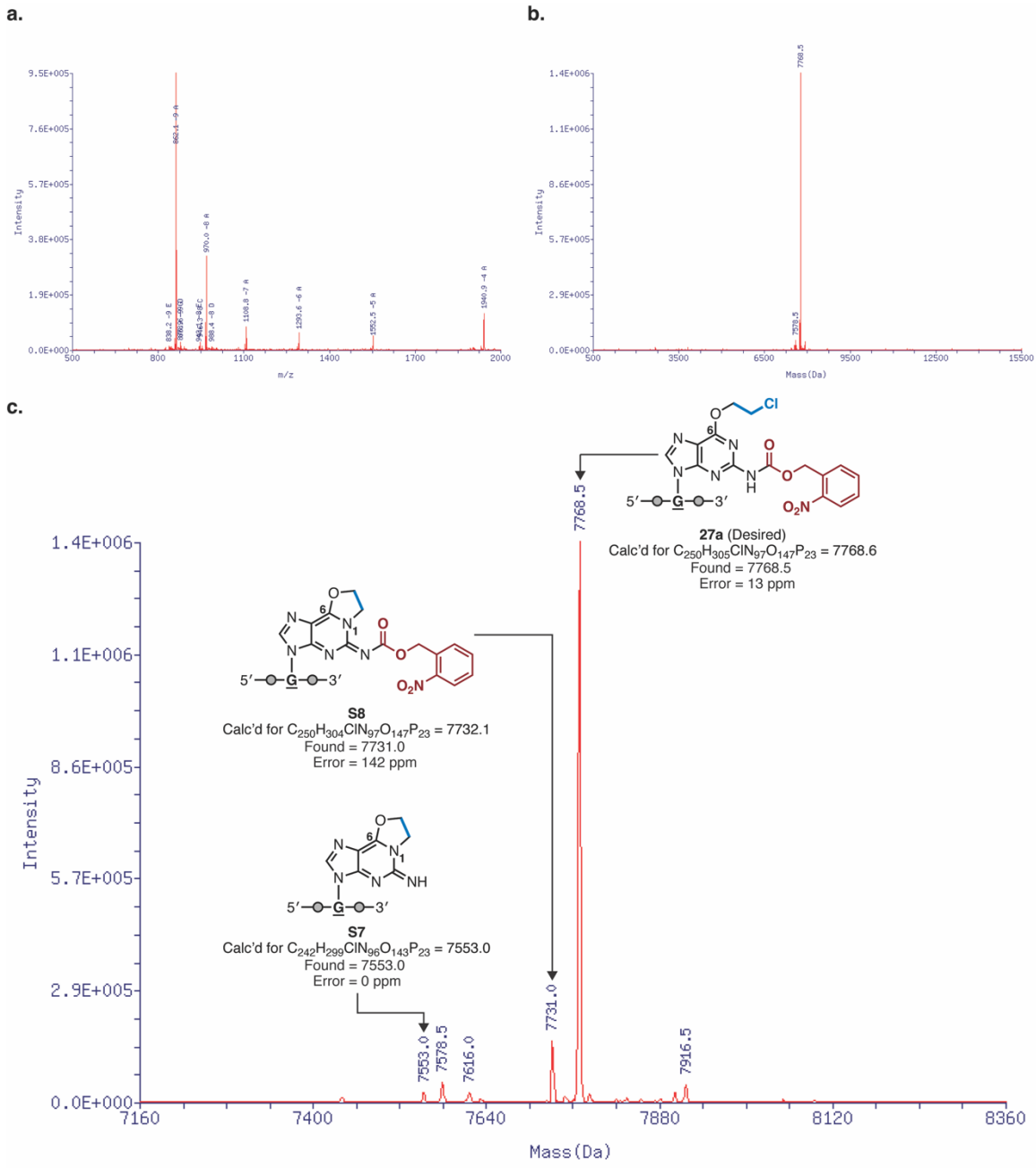


Fig. S18. Mass spectrometry analysis of O^{6}ClEtG containing deoxyoligonucleotide **27a** cleaved from a Q (hydroquinone) resin using 10:1 (v/v) methanol–diisopropylamine. **A.** Annotated mass spectrum (ESI, negative mode). **b.** Deconvoluted mass spectrum. **c.** Expansion of the deconvoluted mass spectrum and peak assignments. Sequence = 5′-d(GATGTGGTATA**G**ATAAGGAGTTGT)-3′ where **G** = modified guanine. For **c**, calculated masses are molecular weights and the structures of **S7–S8** are proposed only based on mass spectral data.

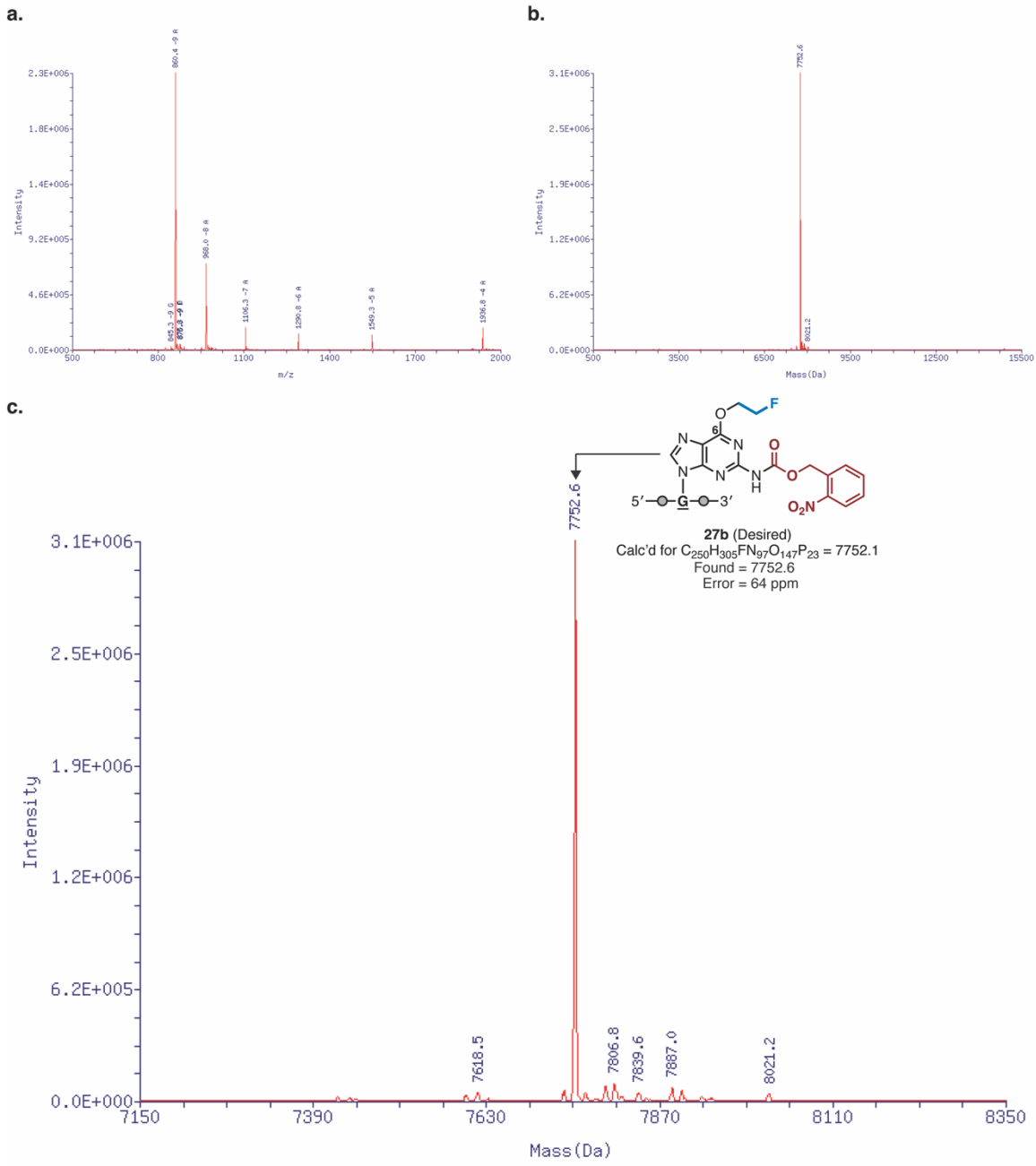


Fig. S19. Mass spectrometry analysis of the O^{6}FEtG containing deoxyoligonucleotide **27b** cleaved from a Q (hydroquinone) resin using 10:1 (v/v) methanol–diisopropylamine. **a.** Annotated mass spectrum (ESI, negative mode). **b.** Deconvoluted mass spectrum. **c.** Expansion of the deconvoluted mass spectrum with peak assignment. Sequence = 5′-d(GATGTGGTATA**G**ATAAGGAGTTGT)-3′ where **G** = modified guanine. For **c**, calculated masses are molecular weights.

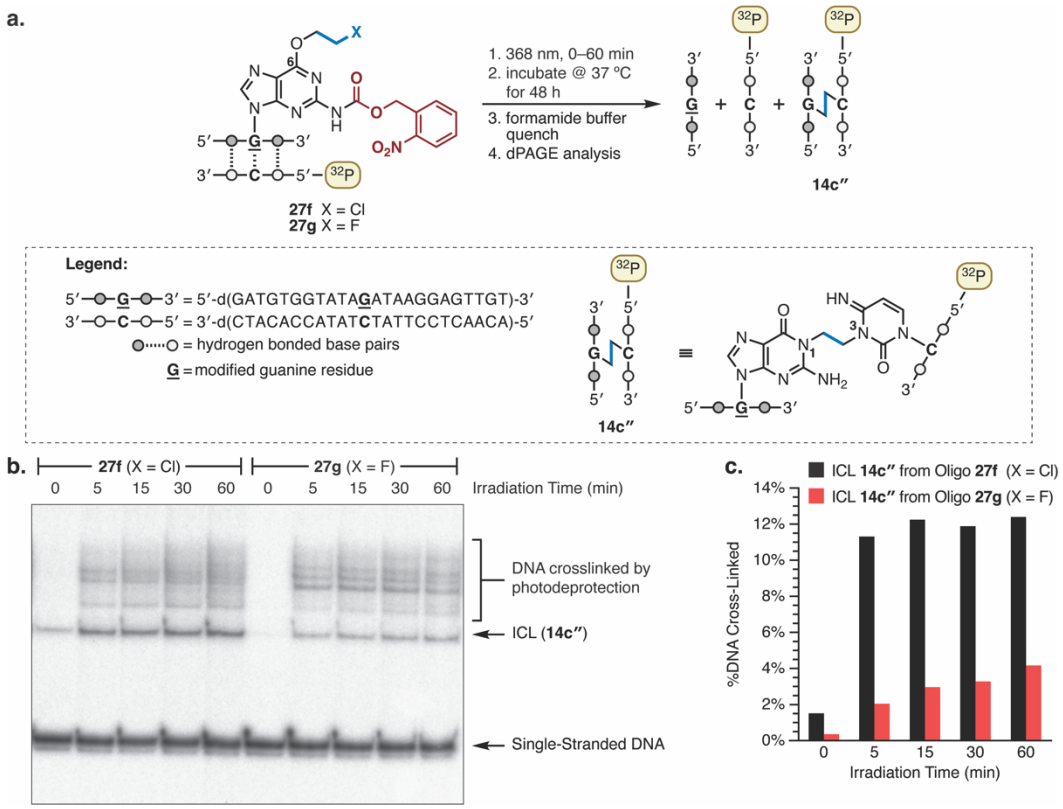


Fig. S20. Effect of irradiation time on the formation of the [G(N1)–C(N3)]Et ICL (**14c″**). **a.** Experimental design: the duplexes **27f** (X = Cl) or **27g** (X = F) were annealed, irradiated (368 nm light, 0, 5, 15, 30 or 60 min), incubated at 37 °C for 48 h, diluted with denaturing formamide buffer, and analyzed by dPAGE (18%). **b.** Representative denaturing gel visualized by autoradiography. **c.** Bands corresponding to single-stranded DNA and the [G(N1)–C(N3)]Et ICL (**14c″**) were quantified using ImageQuant software (Cytiva). For **a**–**c**, **14c″** is identical in structure to **14c** (Fig. 11) except that the opposite strand is 5′-^{32}P labeled. For **c**, bars indicate the amount of [G(N1)–C(N3)]Et ICL (**14c″**) divided by total amount of DNA expressed as a percentage.

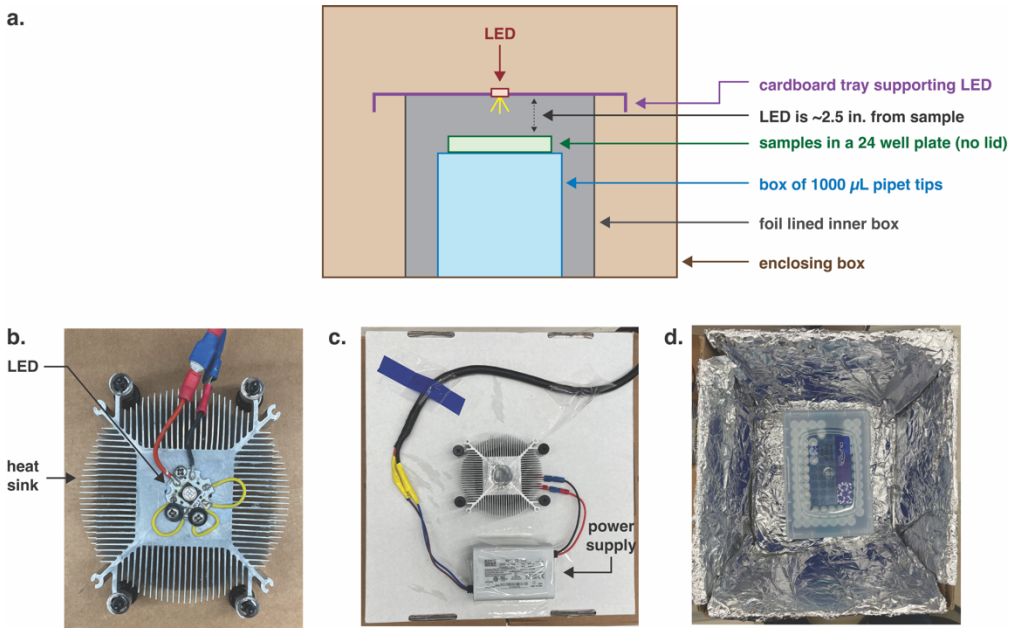


Fig. S21. Visual description of the photoirradiation setup. A 24 well plate containing samples to be irradiated was positioned on top of a box of 1000 µL pipet tips inside of a foil lined box. The LED light was affixed to a heat sink and suspended above the 24 well plate using a cardboard tray with cutout for the LED light. The entire apparatus was enclosed in a large cardboard box to contain stray light. The distance between the LED and the 24 well plate was ~2.5 in. **a.** Cartoon of the photoirradiation setup (not drawn to scale). **b.** Picture of the LED light affixed to the heat sink. **c.** Picture of the LED light connected to the power supply and suspended using the carboard tray. **d.** Picture of the foil lined box containing the box of 1000 µL pipet tips on which the 24 well plate was placed.

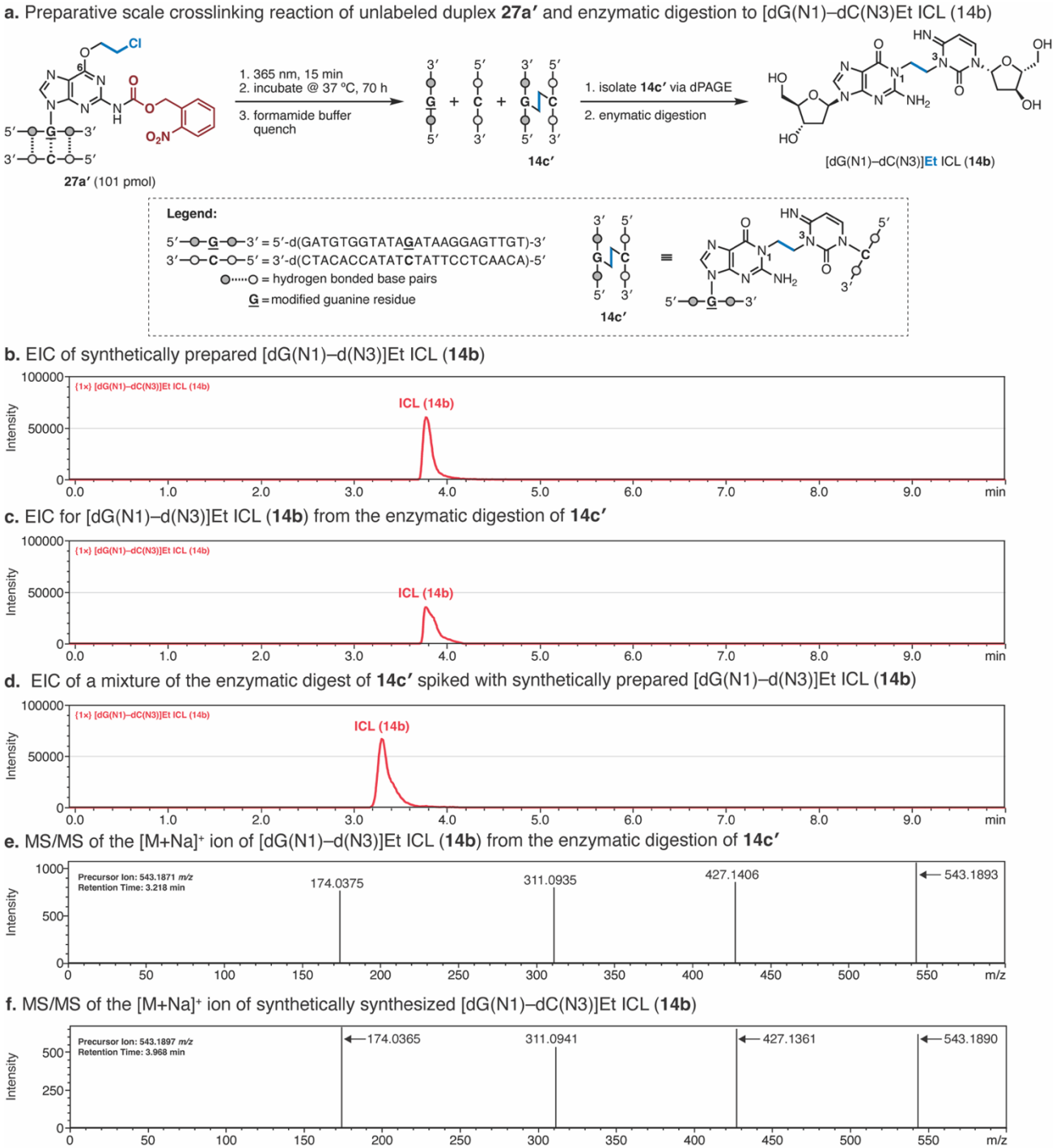
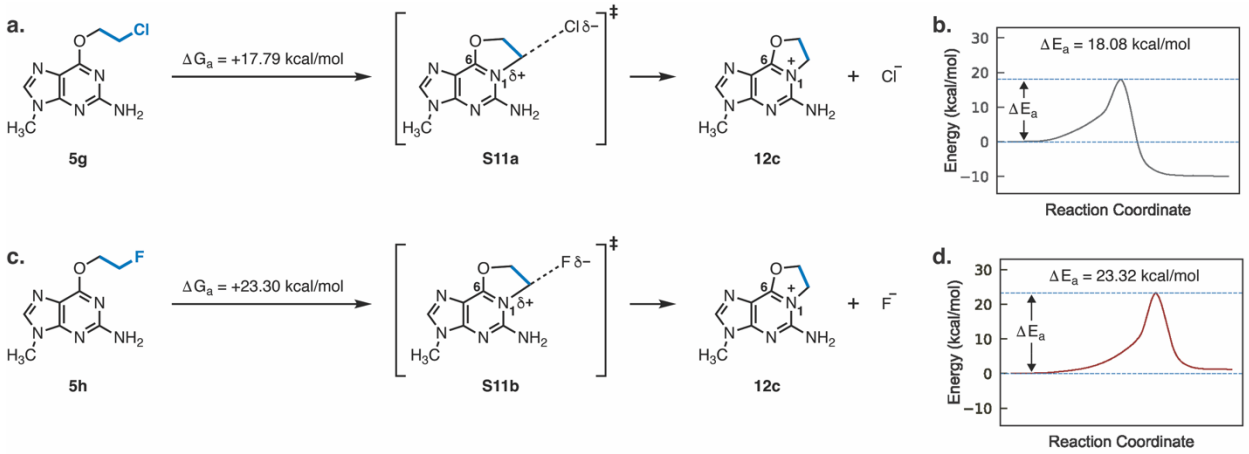


Fig. S22. Verifying the structure of the ICL **14c** via LC/MS/MS. **a.** Unlabeled N^2-NBOC-O^6-(2-chloroethyl)-2′-deoxyguanosine containing duplex **27a′** (101 pmol) was irradiated (368 nm, 15 min) then incubated (37 ºC, 70 h). The band corresponding to the unlabeled ICL **14c′** was isolated by dPAGE and enzymatically digested to yield a nucleoside pool containing [dG(N1)–dC(N3)]Et ICL (**14b**). **b.** Positive mode extracted ion chromatogram (EIC) for synthetically prepared [dG(N1)–dC(N3)]Et ICL (**14b**). **c.** Positive mode EIC of the nucleoside pool resulting from the enzymatic digestion of **14c′**. **d.** Positive mode EIC of a mixture of the nucleoside pool resulting

from the enzymatic digestion of **14c′** and a standard solution of synthetically prepared [dG(N1)–dC(N3)]Et ICL (**14b**). **e.** MS/MS fragmentation of the $[M+Na]^+$ ion of [dG(N1)–dC(N3)]Et ICL (**14b**) obtained from the enzymatic digestion of **14c′**. **f.** MS/MS fragmentation of the $[M+Na]^+$ ion of synthetically prepared [dG(N1)–dC(N3)]Et ICL (**14b**). For **a**, **c**, and **e**, **14c′** is identical in structure to **14c** (Fig. 11) except that **14c′** lacks the 5′-^{32}P label. For **b**–**d**, a positive mode EIC for the calculated *m/z* of the $[M+Na]^+$ ion ± 20 ppm was plotted for [dG(N1)–dC(N3)]Et ICL (**14b**; $[M+Na]^+$ calc'd for $[C_{21}H_{28}N_8NaO_8]^+$ 543.1922); the intensity of each EIC is given in arbitrary units and was scaled by the factor given in curly brackets (i.e., {}). For **e** and **f**, MS/MS data were acquired with a variable collision energy (CE) from 18.0 – 52.0 using the precursor ion 543.1871 *m/z* (retention time = 3.218 min) for **e** and the precursor ion 543.1897 *m/z* (retention time = 3.968 min) for **f**. See Fig. S9e for proposed structures for masses observed from the fragmentation of the $[M+Na]^+$ ion of [dG(N1)–dC(N3)]Et ICL (**14b**). For **b**–**f**, intensity is given in arbitrary units. All data were acquired on LC/MS #2.



e.

Functional/Basis Set	Optimized Energy of Reactant 5h (Hartree)	Optimized Approximate Transition State S11b (Hartree)	ΔE (Hartree)	ΔE (kcal/mol)
B3LYP/6-31G	-759.5392213	-759.4978171	-0.0414041	-25.9811
B3LYP/6-31G(d)	-759.7732389	-759.7181969	-0.0550421	-34.53891
B3LYP/6-31G(d,p)	-759.7911159	-759.7364253	-0.0546906	-34.318326
B3LYP/6-31+G(d)	-759.8154053	-759.7878515	-0.0275538	-17.289991
B3LYP/6-31+G(d,p)	-759.8327923	-759.8054558	-0.0273365	-17.153669
B3LYP/6-31++G(d)	-759.8156788	-759.7881915	-0.0274873	-17.248283
B3LYP/6-31++G(d,p)	-759.8330224	-759.8057357	-0.0272867	-17.122413
B3LYP/6-311+G(d)	-759.9844655	-759.9577311	-0.0267344	-16.775849
B3LYP/6-311++G(d,p)	-760.0005418	-759.9738693	-0.0266725	-16.736977
ωB97X-D/6-31G	-759.2568107	-759.2141016	-0.0427091	-26.799981
ωB97X-D/6-31G(d)	-759.4967304	-759.4397612	-0.0569692	-35.748154
ωB97X-D/6-31G(d,p)	-759.5135367	-759.4568722	-0.0566646	-35.557009
ωB97X-D/6-31+G(d)	-759.5308453	-759.5009733	-0.029872	-18.744709
ωB97X-D/6-31+G(d,p)	-759.5471656	-759.5174845	-0.029681	-18.624857
ωB97X-D/6-31++G(d)	-759.5310896	-759.5012834	-0.0298062	-18.703399
ωB97X-D/6-31++G(d,p)	-759.5473807	-759.5177479	-0.0296327	-18.594533
ωB97X-D/6-311+G(d)	-759.6915251	-759.6629754	-0.0285497	-17.914927
ωB97X-D/6-311++G(d,p)	-759.7070837	-759.6784838	-0.0285999	-17.94646

Fig. S23. Calculated energy barriers for the cyclization of O^6-(2-haloethyl)-N9-methylguanines **5g** and **5h** to N1,O^6-ethano-N9-methyleguanine (**12c**). **a.** Energy barrier from frequency calculations for the cyclization of O^6-(2-chloroethyl)-N9-methylguanine (**5g**) to N1,O^6-ethano-N9-methyleguanine (**12c**). **b.** Intrinsic reaction coordinate calculations for the cyclization of O^6-(2-chloroethyl)-N9-methylguanine (**5g**) to N1,O^6-ethano-N9-methyleguanine (**12c**). **c.** Energy barrier from frequency calculations for the cyclization of O^6-(2-fluoroethyl)-N9-methylguanine (**5h**) to N1,O^6-ethano-N9-methyleguanine (**12c**). **d.** Intrinsic reaction coordinate calculations for the cyclization of O^6-(2-fluoroethyl)-N9-methylguanine (**5h**) to N1,O^6-ethano-N9-methyleguanine (**12c**). **e.** Results of calculation benchmarking. For each benchmarking calculation, the reactant (**5h**) was first optimized with the respective basis set and function. Then an approximate transition state (**S11b**) was optimized with the C–F bond length fixed at 2.0Å. The energy difference between the optimized reactant and approximate transition state were compared. The polarization function and the diffusion function are shown to be important for both functionals. The energy differences converge as more basis functions are added. B3LYP and ωB97X-D are different by about 1 kcal/mol systematically. For **a**–**d**, density functional theory (DFT) calculations were performed with Gaussian 16[6] using the 6-311++G(d,p)[7] basis set and B3LYP functional[8] with GD3 dispersion[9]. For **a**–**e**, the SMD model for water[10] was used for all calculations with the temperature set to 298.15 K.

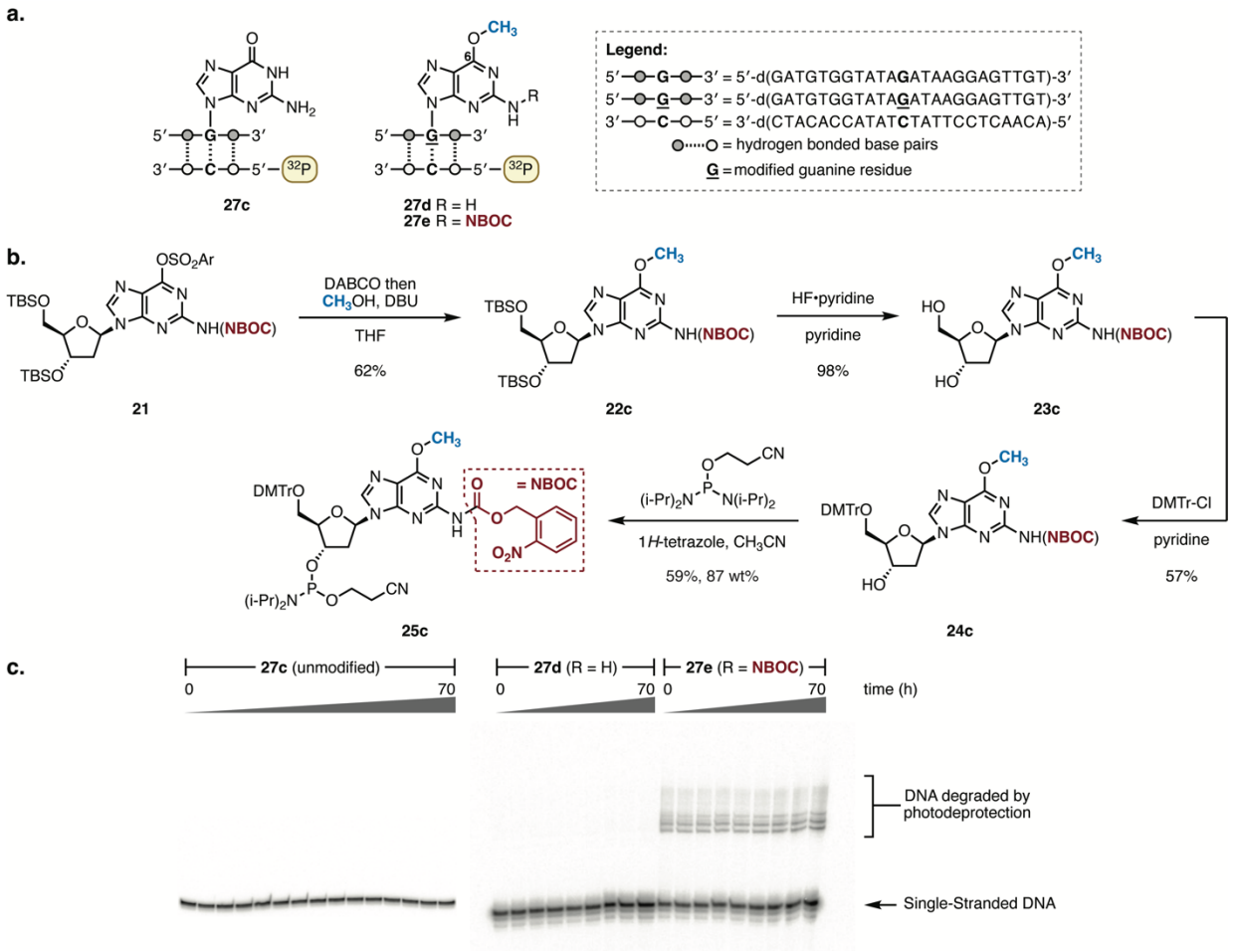


Fig. S24. **a.** Structures of the control duplexes **27c–27e**. **b.** Synthesis of the photocaged N^2-NBOC-O^6-methylguanine phosphoramidite **25c** required for the synthesis of the duplex **27e**. **c.** dPAGE analysis of the control duplexes **27c–27e** over the time 0–70 h. These results demonstrate that the time-independent bands result from DNA degraded by photodeprotection of the N^2-NBOC photolabile protecting group.

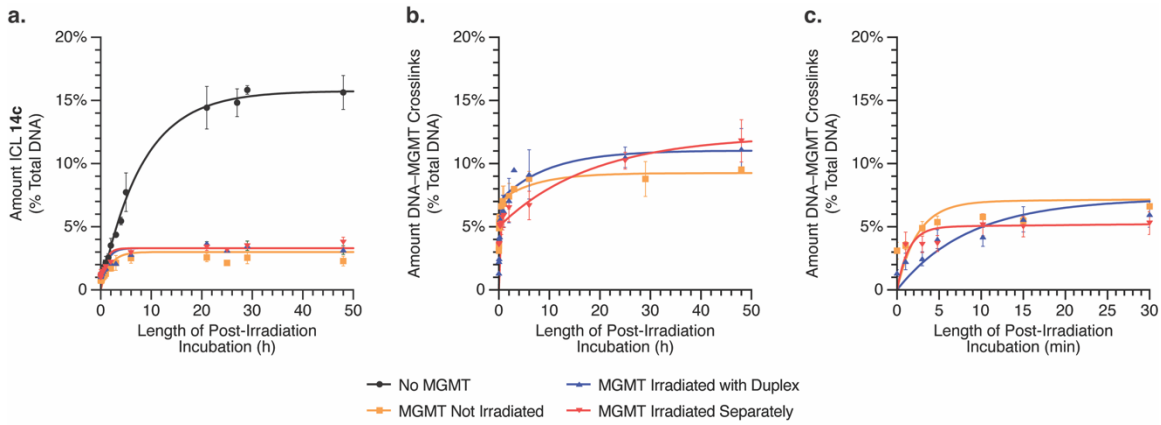


Fig. S25. Effect of timing of MGMT addition on the amount of DNA ICLs and DNA–MGMT DPCs. Prior to incubation at 37 ºC, O^6ClEtG containing duplex **27a** was irradiated (black circles), irradiated then treated with human MGMT (yellow squares), combined with human MGMT then irradiated (blue point-up triangles), or irradiated then combined with separately irradiated human MGMT (red point-down triangles). The samples were then incubated for 0–48 h. Aliquots were removed and immediately diluted with formamide denaturing buffer to halt crosslinking. At the conclusion of the time course, the aliquots were resolved by dPAGE (18%). **a.** Plot of amount ICL **14c** as a percentage of total DNA versus time. **b.** Plot of amount DNA–MGMT crosslinks as a percentage of total DNA versus time. **c.** Zoom in from 0–30 min of the plot in panel **b** to highlight the rapidity of the formation of MGMT–DNA crosslinks. For **a**–**c**, DNA was quantified by autoradiography; points indicate mean; error bars indicate standard deviation; n = 3 biological replicates.

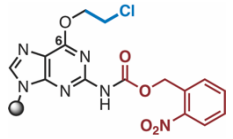
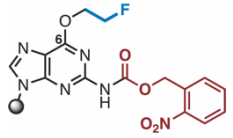
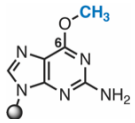
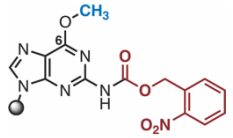
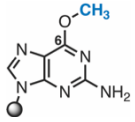
Supplementary Tables.

Table S1. Details of LC/MS experiments with KL-50 (**19**) treated calf thymus DNA (ctDNA) and characterization of ICL **14c**

Entry	Sample Description	LC/MS Machine	Injection Volume (μL)	DDA	Targeted MS/MS[a]	Appears in Figure(s)
1	O^6FEtdG (**5f**) and [dG(N1)–dC(N3)]Et (**14b**)	#1	25	Yes	Yes	Figs. 4c, S5a, S5b, S11a, S13a, S14c, S15c
2	dG (**6b**), dC (**13b**), dA (**S5**), dt (**S6**),O6FEtdG **5f**, and [dG(N1)–dC(N3)]Et (**14b**)	#1	25	No	No	Figs. S10a, S12a
3	[dG(N1)–dC(N3)]Et (**14b**)	#2	3	Yes	No	Figs. S9a, S9b, S9d, S15e, S22b, S22f
4	nucleoside pool (ctDNA + vehicle, 5h)	#1	25	Yes	Yes	Figs. S10b, S11b
5	nucleoside pool (ctDNA + KL-50 (**19**), 5h)	#1	25	Yes	Yes	Figs. 4d (solid series), S10c, S11c, S14a, S14b
6	nucleoside pool (ctDNA + vehicle, 5h) + O^6FEtdG (**5f**) and [dG(N1)–dC(N3)]Et (**14b**) spike	#1	25	No	No	Figs. S10d, S11d
7	nucleoside pool (ctDNA + KL-50 (**19**), 5h) + O^6FEtdG (**5f**) and [dG(N1)–dC(N3)]Et (**14b**) spike	#1	25	No	No	Figs. 4d (dashed series), S10e, S11e
8	nucleoside pool (ctDNA + vehicle, 72h)	#1	25	Yes	Yes	Figs. S12b, S13b
9	nucleoside pool (ctDNA + KL-50 (**19**), 72h)	#1	25	Yes	Yes	Figs. 4e (solid series), S12c, S13c, S15a, S15b, S15d
10	nucleoside pool (ctDNA + vehicle, 72h) + O^6FEtdG (**5f**) and [dG(N1)–dC(N3)]Et (**14b**) spike	#1	25	No	No	Figs. S12d, S13d
11	nucleoside pool (ctDNA + KL-50 (**19**), 72h) + O^6FEtdG (**5f**) and [dG(N1)–dC(N3)]Et (**14b**) spike	#1	25	No	No	Figs. 4e (dashed series), S12e, S13e
12	nucleoside pool (ICL **14c** from duplex **27a'**)	#2	3	Yes	No	Figs. S22c, S22e
13	nucleoside pool (ICL **14c** from duplex **27a'**) + [dG(N1)–dC(N3)]Et (**14b**) spike	#2	3	Yes	No	Fig. S22d

[a]The following *m/z* values (± 1 amu) were input for targeted MS/MS: 314.1200 ($[M+H]^+$ of O^6FEtdG (**5f**)), 521.2103 ($[M+H]^+$ of [dG(N1)–dC(N3)]Et (**14b**)), and 543.1920 ($[M+Na]^+$ of [dG(N1)–dC(N3)]Et (**14b**)).

Table S2. Deoxyoligonucleotides employed in this study.

Number	Sequence	Structure of Non-Natural Base (G)
26a	5′-d(GAT GTG GTA TAG ATA AGG AGT TGT)-3′	
26b	5′-d(GAT GTG GTA TAG ATA AGG AGT TGT)-3′	
26c	5′-d(GAT GTG GTA TAG ATA AGG AGT TGT)-3′	---
26d	5′-d(GAT GTG GTA TAG ATA AGG AGT TGT)-3′	
26e	5′-d(GAT GTG GTA TAG ATA AGG AGT TGT)-3′	
30′	5′-d(GAA CTG CAG CTC CGT GCT GGC CC)-3′	
S12	5′-d(ACA ACT CCT TAT CTA TAC CAC ATC)-3′	---
S13	5′-d(GGG CCA GCA CGG AGC TGC AGT TC)-3′	---

General chemical experimental procedures. All reactions were performed in single-neck, flame dried round-bottomed flasks fitted with rubber septa under a positive pressure of argon, unless otherwise specified. Air- and moisture-sensitive liquids were transferred via syringe or stainless-steel cannula. Organic solutions were concentrated by rotary evaporation at 30 °C, unless otherwise noted. Automated flash-column chromatography was performed using either a Biotage Isolera One system with ultraviolet (UV) monitoring or a Biotage Isolera One system with lambda-all monitoring, employing reusable cartridges hand-packed with silica gel (SiliaFlash® P60, 60 Å, 40-63 μm particle size) purchased from Silicycle (Québec, Canada). Analytical thin-layered chromatography (TLC) was performed using glass plates pre-coated with silica gel (250 μm, 60 Å pore size) embedded with a fluorescent indicator (254 nm). TLC plates were visualized by exposure to UV light or staining with cerium ammonium molybdate (CAM) stain, followed by brief heating with a heat gun. High pressure liquid chromatography (HPLC) was performed using an Agilent Prepstar HPLC system with a Phenomenex Luna $C_{18}(2)$ (100 Å) 10 μm (10 × 250 mm) column.

Chemical materials. Commercial solvents, chemicals, and reagents were used as received with the follow exceptions. Dichloromethane, tetrahydrofuran, and acetonitrile were purified according to the method of Pangborn et al.[11] Triethylamine and pyridine were distilled from calcium hydride under an atmosphere of dinitrogen prior to use. Methanol used in the synthesis of the N^2-NBOC protected O^6-methyl ether **22c** was distilled from magnesium under an atmosphere of dinitrogen prior to use. 2-Nitrobenzyl 1*H*-imidazole-1-carboxylate was prepared according to the procedure of Luedtke and co-workers.[12]

Chemistry instrumentation. Proton nuclear magnetic resonance (^{1}H NMR) were recorded at 400, 500, or 600 megahertz (MHz) at 25 °C, unless otherwise noted. Chemical shifts are expressed in parts per million (ppm, δ scale) downfield from tetramethylsilane and are referenced to residual proton in the NMR solvent ($CHCl_3$, δ 7.26; $(CD_3)SO(CHD_2)$, δ 2.50; CHD_2CN, δ 1.94). Data are represented as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet and/or multiple resonances, b = broad, app = apparent), coupling constant in Hertz (Hz), integration, and assignment. Proton-decoupled carbon nuclear magnetic resonance spectra ($^{13}C\{^1H\}$ NMR) were recorded at 126 or 151 MHz at 25 °C, unless otherwise noted. Chemical shifts are expressed in parts per million (ppm, δ scale) downfield from tetramethylsilane and are referenced to the carbon resonances of the solvent ($CHCl_3$, δ 77.16; DMSO-d_6, δ 39.52; CD_3CN, δ 1.3). ^{1}H–^{1}H gradient-selected correlation spectroscopy (^{1}H–^{1}H gCOSY) and ^{1}H–^{1}H rotating frame Overhauser enhancement spectroscopy (^{1}H–^{1}H ROESY) were recorded at 600 MHz, unless otherwise noted. ^{1}H–^{13}C heteronuclear single quantum coherence (^{1}H–^{13}C HSQC, $^1J_{CH}$ = 146 Hz or 210 Hz[a]), ^{1}H–^{13}C heteronuclear two-bond correlation (^{1}H–^{13}C H2BC), ^{1}H–^{13}C gradient-selected heteronuclear multiple bond correlation (^{1}H–^{13}C gHMBC, $^nJ_{CH}$ = 8 Hz), and ^{1}H–^{13}C long-range heteronuclear single quantum multiple bond correlation (^{1}H–^{13}C LR-HSQMBC, $^nJ_{CH}$ = 2 Hz)[13] were recorded at 500/126 or 600/151 MHz, unless otherwise noted. ^{1}H-^{15}N heteronuclear single quantum coherence (^{1}H–^{15}N HSQC, $^1J_{NH}$ = 100 Hz), and ^{1}H-^{15}N gradient heteronuclear single quantum coherence (^{1}H–^{15}N gHMBC, $^nJ_{NH}$ = 4 Hz) were recorded 600/61 MHz at 25 °C, unless

[a] For each compound, an ^{1}H–^{13}C HSQC spectrum was acquired with $^1J_{CH}$ = 146 Hz. In some cases, the HSQC cross peak from H8→C8 of the guanine nucleobase (traditional numbering scheme) was weak, likely because the typical $^1J_{CH}$ value for this bond is ~210 Hz (e.g., $^1J_{C8\text{-}H8}$ = 213 Hz for the 2-fluoroethyl ether **S16** and $^1J_{C8\text{-}H8}$ = 213 Hz for the O^6-(2-chloroethyl) diol **23a** as measured via ^{1}H-^{13}C gHBMC at 600/151 MHz in DMSO-d_6 and $CDCl_3$, respectively). In these cases, a second ^{1}H–^{13}C HSQC spectrum was acquired with $^1J_{CH}$ = 210 Hz.

otherwise noted; for $^{1}H–^{15}N$ HSQC and $^{1}H–^{15}N$ gHMBC experiments, δ_N are expressed in parts per million (ppm, δ scale) downfield from ammonia. Carbon-decoupled fluorine nuclear magnetic resonance spectra (^{19}F NMR) were recorded at 396 or 471 MHz at 25 °C, unless otherwise noted. Chemical shifts are expressed in parts per million (ppm, δ scale) downfield from tetramethylsilane. Proton-decoupled phosphorus nuclear magnetic resonance spectra ($^{31}P\{^{1}H\}$) were recorded at 202 MHz using triphenyl phosphate as an internal reference and, for quantitative experiments, as an internal standard, unless otherwise noted. Chemical shifts are parts per million (ppm, δ scale) downfield from tetramethylsilane and are referenced to the phosphorus resonance of triphenyl phosphate ($CDCl_3$, δ –17.7; CD_3CN, δ –17.0). For quantitative experiments, $^{31}P\{^{1}H\}$) were recorded with a relaxation delay of 25 s.

Attenuated total reflectance Fourier transform infrared (ATR-FTIR) spectra were obtained using a Shimadzu IRTracer-100 FTIR spectrometer referenced to a polystyrene standard. Data are represented as follows: frequency of absorption (cm^{-1}), intensity of absorption (s = strong, m = medium, w = weak, br = broad).

Optical rotations were measured on a Rudolph Research Analytical Autopol VI polarimeter equipped with a sodium (589 nm, D) lamp. Optical rotation data are represented as follows: specific rotation ($[a]_D^{20}$), concentration (g/mL), and solvent.

Liquid chromatography-mass spectroscopy (LC/MS) and liquid chromatography-tandem mass spectroscopy (LC/MS/MS) analyses were performed using four different instruments denoted LC/MS #1–#4 as described below. LC/MS #1 was used to analyze nucleoside pools obtained in the alkylation of calf thymus DNA with KL-50 (**19**) and to obtain high resolution mass spectrometry (HRMS) data for all chemically synthesized compounds.[b] LC/MS #2 was used to analyze the nucleoside pool obtained from the chemical characterization of the ICL **14c**.[b] LC/MS #3 was used to monitor the synthesis and purification of all chemically synthesized compounds with the exception of the [dG(N1)–dC(N3)]Et ICL (**14b**). LC/MS #4 was used to monitor the purification of [dG(N1)–dC(N3)]Et ICL (**14b**). All chromatographic separations were run using a binary gradient in which A = water containing 0.1% formic acid and B = acetonitrile containing 0.1% formic acid.

> LC/MS #1 was a Shimadzu Scientific Instruments QToF 9030 LC/MS system, equipped with a Nexera LC-40D xs UHPLC, consisting of a CBM-40 Lite system controller, a DGU-405 Degasser Unit, two LC-40D XS UHPLC pumps, a SIL-40C XS autosampler and a Column Oven CTO-40S. UV data was collected with a Shimadzu Nexera HPLC/UHPLC Photodiode Array Detector SPD M-40 in the range of 190 - 800nm. Mass spectra were subsequently recorded with the quadrupole time-of-flight (QToF) 9030 mass spectrometer. Samples were held at 4 ºC in the autosampler compartment. The ionization source was run in "ESI" mode, with the electrospray needle held at +4.5kV. Nebulizer gas was at 2 L/min, heating gas flow at 10 L/min and the interface at 300 ºC. Dry gas was at 10 L/min, the desolvation line at 250 ºC and the heating block at 400 ºC. Mass spectra were recorded in the range of 50 to 2000 m/z in positive or negative ion mode with an event time of 100 ms per scan. Chromatographic separation was achieved using a Shim-pack Scepter C18-120, 1.9 µm, 2.1 × 100 mm Column equilibrated at 40 ºC in a column oven. Samples were eluted at a constant flow rate of 400 µL/min using the following method: 5% B–A (0–2

[b] Please see Table S1 for a summary of details pertaining to the analysis of the nucleoside pools obtained in the alkylation of calf thymus DNA with KL-50 (**19**) and the chemical characterization of ICL **14c** using LC/MS#1–#2.

min), 5%→95% B–A (2–8 min), 95% B–A (8–10 min), 95%→5% B–A (10–10.01 min), and 5% B–A (10.01–12 min), in that order. Tandem mass spectrometry data (MS/MS) was acquired using Data Dependent Acquisition (DDA) performed with a variable collision energy of 35 ±17 V (internal units) with an event time of 100 ms per scan and a scan range of 50 to 1000 *m/z*. Targeted MS/MS combined with DDA performed with a variable collision energy of 35 ± 17 V (internal units) with an event time of 50 ms per event, a scan range of 50 to 600 *m/z* for all targeted ion fragmentations, and a scan range of 100 to 1000 *m/z* for DDA.

LC/MS #2 was a separate LC instrument with a configuration identical to that described for LC/MS #1 but connected by a longer transfer line into the same HRMS instrument as described for LC/MS #1. Samples were eluted at a constant flow rate of 400 µL/min using the following method: 2% B–A (0–2 min), 2%→50% B–A (2–15 min), 50%→98% B–A (15–16 min), 98% B–A (16–18 min), 98%→2% B–A (18–18.01 min), and 2%→(18.01–20 min).

LC/MS #3 was a Waters instrument equipped with a reverse-phase C_{18} column (1.7 µm particle size, 2.1 × 50 mm). Samples were eluted using the following method: 5%→95% B–A (0–4 min, 700 µL/min) then 95% B–A (4–5.5 min, 800 µL/min).

LC/MS #4 was an Agilent 1260 Infinity system using a Phenomenex Luna C_{18}(2) (100 Å) 5 µm (4.6 × 150 mm) (Phenomenex, CA, USA) column and a PDA detector coupled with a single quadrupole electrospray ionization mass spectrometry instrument (Agilent 6120). Samples were eluted at a constant flow rate of 700 µL/min using the following method: 10% B–A→100% B (30 min).

Chemicals for molecular biology experiments. Temozolomide (TMZ, **1**) and lomustine (**10**) were purchased from Selleck Chemicals. O^6-Benzylguanine (O^6BnG) was purchased from Selleck Chemicals and Combi-Blocks. Mitozolomide (MTZ, **11**) was purchased from Enamine. KL-50 (**19**) was synthesized as previously described[14] or obtained as a generous gift from Modifi Biosciences. Mitomycin C (MMC) was purchased from Sigma. For long-term clonogenic survival assays (Fig. 5), modified alkaline comet assays (Fig. 8a), and denaturing gel electrophoresis (Fig. 8b), TMZ (**1**, 100 mM stock stored at –80 ºC), lomustine (**10**, 100 mM stock stored at –20 ºC), MTZ (**11**, 100 mM stock stored at –80 ºC), KL-50 (**19**, 100 mM stock stored at –80 ºC), and MMC (10 mM stock stored at –20 ºC) were dissolved in dimethyl sulfoxide and stored at the indicated temperature. For the alkylation of calf thymus DNA (Fig. 9), KL-50 (**19**, 50 mM stock) was dissolved in dimethyl sulfoxide immediately before use. For the *Pst*I restriction assay (Fig. 12a) and the kinetic analysis of duplex **27a** in the presence of deactivated MGMT (Fig. 12b), O^6BnG (10 mM stock) was dissolved in dimethyl sulfoxide and stored at –20 ºC.

Biological reagents for molecular biology experiments. Calf thymus DNA was obtained from Invitrogen (Catalog No: 15633019). Recombinant human MGMT was obtained from Creative Biomart (Catalog No: MGMT-121H) and Abcam (Catalog No: ab79251). Upon receipt, the human MGMT was aliquoted, frozen in liquid dinitrogen, and stored at –80 ºC.

Light source for molecular biology experiments. Duplexes **27a–27g** and MGMT were irradiated using a UV LED emitter 368 nm (365 ~ 370 nm) (Digi-Key: catalog # 475-LZ4-V4UV0R-0000-ND) affixed to a heat sink and soldered to a 25 W single output switching power suppler (Digi-Key: catalog # 1866-1125-ND). The samples were positioned ~2.5 in. away from

the light source in a foil-lined box (9 in. length × 8 in. width × 7.75 in. height). See Fig. S21 for a visual depiction of the photoirradiation setup.

Cell culture. LN229 MGMT– and MGMT+ cell lines were provided by B. Kaina (Johannes Gutenberg University Mainz, Mainz, Germany). LN229 MGMT–/MMR– and LN229 MGMT+/MMR– cell lines generated through lentiviral-mediated shRNA knockdown of MSH2 have been previously described.[14] The cells were maintained in DMEM with 10% FBS (Gibco). Cell line identity was validated by short tandem repeat profiling, phenotypes were confirmed via Western blotting for MGMT and MSH2, and cells were confirmed negative for mycoplasma by quantitative RT-PCR.

Long-term clonogenic survival assays. Assays were performed as previously described.[14] Briefly, cells were trypsinized, washed, counted, diluted in a medium containing indicated concentrations of drug, and seeded in triplicate in six-well plates at three-fold dilutions, ranging from 9000 to 37 cells per well. Plates were incubated at 37 °C for 10-14 days for colony formation, then washed in PBS, stained with crystal violet solution, and manually counted. Surviving fractions were corrected for the plating efficiency under control condition. For conditions yielding no colony growth at 9000 cells/well, a single colony was counted to allow plotting and visualization of the data.

PRISM multiplexed viability screen. KL-50 (**19**) was screened in the PRISM assay[15] at 8-point dose (3-fold dilution (200 μM to 91.4 nM) with a 5 day treatment with 902 cancer cell lines passing quality control. Two PRISM cell line collections were used in the assay: PR500 (including only adherent cell lines), and PR300+ (including adherent and suspension cell lines). Three benchmark compounds, nutlin-3a, AZ-628, and imatinib, were also tested at dose to ensure high data quality. All compounds were run in triplicate, and each plate contained positive (bortezomib, 20 μM) and negative (DMSO) controls.

DepMap comparison of PRISM screens. To provide points of comparison for the Pearson's correlation coefficient of r = 0.642 observed between MGMT expression and KL-50 (**19**) area under the curve (AUC), this value was compared to Pearson correlations between the BCL2 inhibitor venetocalx vs. BCL2 expression, EGRF inhibitor erlotinib vs. EGFR expression, and the TOP2α inhibitor etoposide vs. TOP2α expression using the Dependency Map (DepMap) portal maintained by the Broad institute (https://depmap.org/portal/) via the "Data Explorer" function (https://depmap.org/portal/interactive/). See Fig. S1 for a list of X- and Y-axis inputs and linear regression outputs.

Modified alkaline comet assay. Assay was performed utilizing the CometAssay Kit (Trevigen), as previously described.[14] LN229 MGMT±/MMR± cells were treated in six-well plates with 0.1% DMSO control or KL-50 (**19**, 200 μM) for 24 h. Cells were then collected by trypsinization, washed with 1X PBS, suspended in melted Comet LMAgarose (Trevigen), and spread in a 50 μL volume containing 1000 cells on Trevigen CometSlides. Agarose was allowed to solidify at 4 °C, and slides were then incubated in Trevigen lysis solution with 10% DMSO overnight at 4 °C. Slides were removed from the lysis buffer and irradiated with 0 or 10 Gy with an XRAD 320 X-Ray System (Precision X-Ray) at 320 kV, 12.5 mM, and 50.0 cm SSD, with a 2 mm Al filter and 20 cm x 20 cm collimator. Slides were then incubated for 45 minutes in alkaline buffer (200 mM NaOH, 1 mM EDTA), subjected to alkaline electrophoresis at 21V for 45 minutes at 4 °C, washed with dH_2O twice followed by 70% EtOH, stained with 0.33X SYBR Gold (Invitrogen), and dried per Trevigen assay protocol. Slide images were acquired using a Cytation 3 imaging reader (BioTek), and the %DNA in comet tail was quantified using CometScore 2.0 software (TriTek).

Denaturing gel electrophoresis. Assay was adapted from Bossuet-Greif et al,[16] and performed as previously described.[14] LN229 MGMT±/MMR± cells were treated in 10-cm plates with 0.2%

DMSO control, KL-50 (**19**, 200 μM), MTZ (**11**, 200 μM), or TMZ (**1**, 200 μM) for 24 h or with mitomycin C (MMC, 50 μM) for 2 h. Cells were then collected by trypsinization, washed with 1X PBS, and stored at –80 °C prior to extraction of genomic DNA with the DNeasy Blood & Tissue Kit (Qiagen) and denaturing agarose gel electrophoresis. A 0.7% agarose gel was prepared in 100 mM NaCl-2mM EDTA (pH 8) and soaked in denaturing running buffer (40 mM NaOH–1 mM EDTA) for 2 h. Genomic DNA (400 ng/well) was loaded in 1X BlueJuice loading buffer (Invitrogen) and subjected to electrophoresis at 2 V/cm for 30 min, followed by 3 V/cm for 2 h. The gel was neutralized by immersion in 150 mM NaCl–100 mM Tris (pH 7.4) for 30 min, twice, and then stained with 1X SYBR Gold for 90 min. Stained gels were imaged on a ChemiDoc XRS+ Molecular Imager (Bio-Rad).

Animal studies. All animal studies were conducted in accordance with Mayo Clinic Institutional Animal Care and Use Committee policies (protocol #A00004593-19 for the study in Fig. 7a (TMZ (**1**) vs lomustine (**10**) vs vehicle); protocol #A00006634-22 for the study in Fig. 7b (TMZ (**1**) vs KL-50 (**19**) vs vehicle)). Female athymic nude mice were used for all studies. For the study in Fig. 7a, mice were purchased from Envigo (athymic nu/nu, aged 6–7 weeks); for the study in Fig. 7b, mice were purchased from Charles River (Wilmington, Massachusetts, United States: strain code 553, aged 6-7 weeks). Patient derived xenograft (PDX) maintenance and intracranial injections were performed as previously described.[17] Mice with established orthotopic tumors were dosed as indicated, observed daily, and euthanized at moribund.

Statistical Analysis. Comparison of cell viability in the PRISM multiplexed variability screen for "MGMT High" and "MGMT Low" groups (Fig. 6c) was performed using GraphPad Prism software by unpaired t-test. For in vivo intracranial studies (Fig. 7), survival was defined as the time from tumor implantation to reach a moribund state. Kaplan–Meier analysis was used to evaluate the probability of survival, and statistical comparisons were made using GraphPad Prism software by log-rank (Mantel-Cox) test with Bonferroni correction for multiple comparison.

Alkylation of calf thymus DNA with KL-50 (19). A mixture of calf thymus DNA (500 μg) and either DMSO (12.5 μL) or a solution of KL-50 (**19**) in DMSO (10 mM, 12.5 μL) was diluted to 125 μL final volume with aqueous Tris-HCl buffer (80 mM, pH 7.5); the final concentration of DMSO in the reaction mixture was 10% (v/v). The reaction mixture was incubated at 37 ºC for 5 or 72 h. The DNA was separated using a MilliporeUltra cartridge (3,000 Da MW cut-off). The DNA was digested with a mixture of Exonuclease III (NEB, 100 units/μL, 2 μL) and Exonuclease T5 (NEB, 10 units/μL, 3 μL) in CutSmart buffer (NEB, final volume = 100 μL) at 37 ˚C overnight. After digestion overnight, QuickCIP phosphatase (NEB, 5 units/μL, 6 μL) was added, and the reaction mixture was incubated for 2 h at 37 ºC. The resulting nucleoside pool was separated from the digestion enzymes using a MilliporeUltra cartridge (3,000 Da MW cut-off). The filtered nucleoside pool was used without dilution for LC/MS studies. Spiked samples were prepared by mixing the filtered nucleoside pool (27 μL) with solutions of synthetically prepared O^6FEtdG (**6f**) in HPLC grade water (5 μM, 1.8 μL) and [dG(N1)–dC(N3)]Et ICL (**14b**) in HPLC grade water (30 μM, 0.9 μL). Samples were analyzed on LC/MS #1 using an injection volume of 25 μL.[c]

Solid Phase Oligonucleotide Synthesis. All oligonucleotides were synthesized on a Mermade12 automated RNA-DNA synthesizer (BioAutomation). Commercial phosphoramidites were purchased from TxBio or Glen Research. Oligonucleotides containing an N^2-NBOC-O^6-alkyl-2′-

[c] Please see Table S1 for a summary of details pertaining to the analysis of nucleoside pools obtained in the alkylation of calf thymus DNA with KL-50 (**19**) and the chemical characterization of ICL **14c** using LC/MS#1–#2.

deoxyguanosine residue (**26a**, **26b,** and **26e**) were synthesized using the phosphoramidites **25a**, **25b**, and **25c**, respectively, using a Q-dT polymer support (Glen Research). Phosphoramidites and reagents compatible with ultra-mild oligonucleotide deprotection (Pac-dA, iPr-Pac-dG and Pac_2O Cap A, Glen Research) were used for the synthesis of oligonucleotides containing an O^6-methyl-2′-deoxyguanosine residue (**26d** and **30′**) or an N^2-NBOC-O^6-alkyl-2′-deoxyguanosine residue (**26a**, **26,** and **26e**). Following the completion of solid phase oligonucleotide synthesis, the oligonucleotides were deprotected as described below. Molecular weight and purity of the oligonucleotides was determined by mass-spectrometry (Novatia LLC). Table S2 presents the oligonucleotides synthesized in this study.

Deprotection of unmodified deoxyoligonucleotides 26c, S12, and S13. Following the completion of solid phase oligonucleotide synthesis, the polymer support was transferred to a 4 mL glass vial. Then 4 mL of a 1:1 mixture containing 28–30% aqueous ammonium hydroxide (JT Baker) and 40% aqueous methylamine (Millipore Sigma) were added and the mixture was incubated at 65 °C for 15 min. The supernatant was transferred to a clean vial, cooled at –80 °C for 1 h, and dried using a SpeedVac. The oligonucleotide was redissolved in 750 µL of double-distilled water, desalted using GlenPak 1.0 columns according to the manufacturer's protocol (Glen Research) with an elution volume of 1.75 mL, and precipitated by addition of 100 µL of 5 M NaCl and 10 mL of ethanol. After precipitation, the pellet was dissolved in the storage buffer (20 mM MOPS, pH 6.0, 1 mM EDTA), and the resulting solution was stored at –80 ºC.

Deprotection of deoxyoligonucleotides 26d and 30′ containing an O^6-methyl-2′-deoxyguanosine residue. Following the completion of solid phase oligonucleotide synthesis, the polymer support was transferred to a 4-mL glass vial. Then 1 mL of a solution of 1,8-diazabicyclo[5.4.0]undec-7-ene in anhydrous methanol (10% v/v) was added, and the mixture was incubated at 25 °C in the dark for 7 d. The supernatant was transferred to a clean vial, and 1.5 mL of 2 M triethylammonium acetate buffer (Applied Biosystems) was added. The oligonucleotides were desalted using GlenPak 2.5 columns according to the manufacturer's protocol (Glen Research) with an elution volume of 2.75 mL, and precipitated by addition of 150 µL of 5 M NaCl and 10 mL of ethanol. The pellet was dissolved in the storage buffer (20 mM MOPS, pH 6.0, 1 mM EDTA) and stored at –80 ºC.

Deprotection of deoxyoligonucleotides 26a, 26b, and 26e containing an N^2-NBOC-O^6-alkyl-2′-deoxyguanosine residue. Following the completion of solid phase oligonucleotide synthesis, the polymer support was transferred to a 4-mL glass vial. Then 1 mL of anhydrous methanol–diisopropyl amine (10:1, v/v) was added, and the mixture was incubated at 25 °C in the dark for 24 h. The supernatant was transferred to a clean vial, then 1.5 mL of 2 M triethylammonium acetate buffer (Applied Biosystems) was added, and the oligos were desalted, ethanol-precipitated and stored as above.[d]

5′-End labeling of deoxyoligonucleotides. 100 pmol of DNA oligonucleotide was mixed with 1 µL of the 10x buffer for T4 polynucleotide kinase (NEB), 1 µL of γ-^{32}P-ATP (6000 Ci/mmol, Perkin Elmer) and 1 µL of T4 Polynucleotide kinase (NEB) in a final volume of 10 µL. The reaction mixture was incubated at 37 ºC for 1 h. The 5′-end labeled deoxyoligonucleotides were

[d] Attempts to deprotected deoxyoligonucleotides **26a** and **26b** using the Glen Research protocol for ultra-mild deprotection with potassium carbonate in methanol caused extensive degradation (see Figs. S16–S17).

purified on an 18% denaturing polyacrylamide gel, eluted overnight in 10 mM MOPS, pH 6.0, 1 mM EDTA, 300 mM NaCl at 4 °C, and precipitated by adding three volumes of ethanol.

Kinetic analysis of the formation of the ICL and DNA–MGMT crosslinks. The formation of the ICL **14c** in the absence of the MGMT protein was monitored using the DNA duplex **27a–g** in which either the modified strand (**26a–e**) or the complementary stand (**S12**) was 5′-end labeled with ^{32}P. The formation of the ICL **14c** and DNA–MGMT crosslinks in the presence of the MGMT protein was monitored using the DNA duplex (**27a–b**) in which the modified strand (**26a–b**) was 5′-end labeled with ^{32}P.

The formation of the ICL **14c** and DNA–MGMT crosslinks over time was measured essentially as described[12] in 20 mM sodium phosphate buffer, pH 7.2, 250 mM NaCl in a final volume of 100 µL at 37 ºC and in the presence or absence of 5 µM MGMT. The ^{32}P-labeled DNA duplex with the first strand containing the photocaged nucleotide was at 1 µM. First, the duplex strands were mixed with buffer, sodium chloride, and sterile water, heated at 80 °C for 1 min, and slowly cooled to 37 °C. The mixture was transferred to a well of a 24-well plate and irradiated at 368 nM for 5–15 min (UV LED emitter 368 nm, Digi-Key; see Fig. S22 for visual description of photoirradiation setup); our data suggest that varying the irradiation time from 5–60 min does not affect cross-linking (Fig. S21). After irradiation, the reaction mixture was immediately transferred to an Eppendorf tube, the MGMT protein was added as necessary, and the mixture was incubated at 37 °C. Aliquots (3 µL) at specific time points were withdrawn, mixed with 5 µL of the formamide quench buffer (82% (v/v) deionized formamide, 0.16% (w/v) xylene cyanol (XC), 0.16 % (w/v) bromophenol blue (BB), 10 mM EDTA pH 8.0), chilled on ice, and analyzed on an 18% denaturing polyacrylamide gel. The gel was dried and exposed to a PhosphoImager screen (Cytiva). Bands corresponding to the ICL **14c**, DNA–MGMT crosslinks, and free oligo were quantified using ImageQuant software (Cytiva). Fraction of the ICL **14c** or DNA–MGMT crosslink was plotted vs. time and fit to a single or double exponential equation to determine the rate constants for the formation of ICL **14c** and DNA–MGMT crosslinks, respectively, using KaleidaGraph software.

To determine whether inactivation of MGMT via pre-treatment with O^{6}BnG would increase the formation of ICL **14c** and decrease the formation of DNA–MGMT crosslinks relative to the addition of active MGMT, the experiment was carried out as above with duplex **27a**, except that MGMT was incubated with O^6-BnG (100 µM) for 1 h at 37 ºC prior to its addition to irradiated duplex **27a**.

To determine whether addition of the MGMT prior to the removal of the NBOC photolabile protecting group by irradiation affects the cross-link formation, the experiment was carried out as above with duplex **27a**, except that MGMT was added to the reaction mixture prior to irradiation. To test for potential irradiation damage to the protein in this experiment, a control experiment was carried out in which the MGMT was irradiated prior to addition to the reaction mixture and then the experiment was carried out as above.

Chemical characterization of the ICL 14c. 72 µL of the 1.4 mM N^2-NBOC-O^6-(2-chloroethyl)-2′-deoxyguanosine containing deoxyoligonucleotide **26a** was mixed with 42 µL of the 2.5 mM complementary strand **S12** in 300 µL final volume of 20 mM PBS, pH 7.2, 250 mM NaCl. The reaction mixture was heated at 80 ºC for 1 min and slowly cooled to 37 ºC. The mixture was transferred to a well of a 24-well plate, irradiated at 368 nm for 15 min (UV LED emitter 368 nm, Digi-Key; see Fig. S22 for visual description of photoirradiation setup), transferred to an

Eppendorf tube, and incubated at 37 ºC for 70 h. The samples were ethanol-precipitated and the unlabeled ICL **14c′**[e] was isolated on an 18% denaturing polyacrylamide gel, eluted overnight in 10 mM MOPS, pH 6.0, 1 mM EDTA, 300 mM NaCl at 4 °C, and precipitated by adding three volumes of ethanol. The isolated ICL **14c′**[e] (1 μM final concentration) was digested in the final volume of 100 μL with the mixture of Exonuclease III (NEB, 100 units/μL, 2 μL) and Exonuclease T5 (NEB, 10 units/∝L, 3 μL). Reaction was carried out in the rCutSmart buffer (NEB), at 37 °C overnight. QuickCIP phosphatase (NEB, 5 units/μL, 6 μL) was added, the reaction mixture was incubated at 37 °C for 2 h. The resulting nucleoside pool was separated from the digestion enzymes using a MilliporeUltra cartridge (3,000 Da MWCO). The filtered nucleoside pool was used without dilution for LC/MS studies. A spiked sample was prepared by mixing the filtered nucleoside pool (5 μL) with a solution of synthetically prepared [dG(N1)–dC(N3)]Et ICL (**14b**) in HPLC grade H_2O (1 μM, 5 μL). Samples were analyzed on LC/MS #2 using an injection volume of 3 μL.[f]

Analysis of the alkylation of MGMT by duplexes 27a′–27b′. Step 1 (MGMT alkylation). N^2-NBOC-O^6-(2-haloethyl)-2′-deoxyguanosine containing deoxyoligonucleotide **26a** (X = Cl) or **26b** (X = F) was mixed with the complementary deoxyoligonucleotide **S12** in 40 μL of a buffer containing 50 mM HEPES, pH 7.5, 50 mM KCl, 0.025% Triton X-100 and 1 μg/mL BSA. The final concentration of the resulting duplex **27a′** or **27b′** was 4 μM.[g] The solution was heated at 80 °C for 1 min and slowly cooled to 37 °C. Half of the cooled duplex solution was transferred to a well of a 24-well plate and irradiated at 368 nM for 5–15 min (UV LED emitter 368 nm, Digi-Key; see Fig. S22 for visual description of photoirradiation setup); the other half was not irradiated. Then 4 μM MGMT solution in the same buffer (2 μL) was mixed with 4 μL of the duplex solution (with or without prior irradiation) and incubated at 37 °C for 3 h. The residual MGMT activity of the resulting reaction mixture was assessed by the *Pst*I restriction assay as described below in Step 2.

Step 2 (*Pst*I restriction assay). The deoxyoligonucleotide **30′** containing an O^6-methyl-2′-deoxyguanosine residue in the *Pst*I restriction site was 5′-end labeled as described above and mixed with the complementary oligonucleotide **S13** in 50 mM HEPES, pH 7.5, 50 mM KCl, 0.025% Triton X-100 and 1 μg/mL BSA. The final concentration of the resulting duplex **28** was 66 nM. 6 μL of the resulting solution was mixed with 4 μL of the reaction mixture prepared in Step 1 and incubated at 37 ºC for 2.5 h. Then 1 μL of the 10x rCutSmart buffer and 1 μL of the *Pst*I restriction enzyme (20 U/μL, NEB) were added, and the mixture was incubated at 37 ºC for 1 h. Samples were analyzed on an 18% denaturing polyacrylamide (dPAGE) gel.

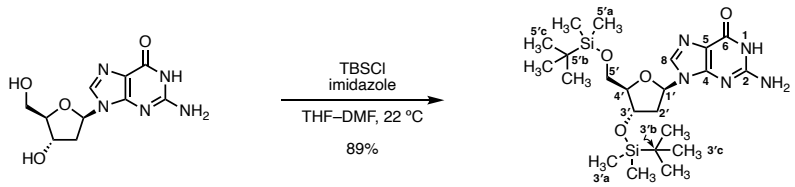
In one control experiment for the *Pst*I restriction assay, duplex **28** was treated directly with the *Pst*I restriction enzyme; no formation of the *Pst*I cleavage product **31** was observed upon dPAGE analysis. In a second control experiment, duplex **28** was treated directly with MGMT followed by the *Pst*I restriction enzyme; formation of the *Pst*I cleavage product **31** was observed upon dPAGE analysis. In a third control experiment, duplex **28** was treated with deactivated MGMT (pre-treated

[e] **14c′** is identical in structure to **14c** (Fig. 11) except that **14c′** lacks the 5′-^{32}P label.
[f] Please see Table S1 for a summary of details pertaining to the analysis of nucleoside pools obtained in the alkylation of calf thymus DNA with KL-50 (**19**) and the chemical characterization of ICL **14c** using LC/MS#1–#2.
[g] Duplexes **27a′** and **27b′** are identical in structure to duplexes **27a** or **27b**, respectively (Fig. 11a), except that **27a′** and **27b′** lack the 5′-^{32}P label.

with O^6BnG (100 μM) for 1 h at 37 ºC) followed by the *Pst*I restriction enzyme; no formation of the *Pst*I cleavage product **31** was observed upon dPAGE analysis.

Synthetic Procedures.

Synthesis of the bis(silyl ether) ***20****.*



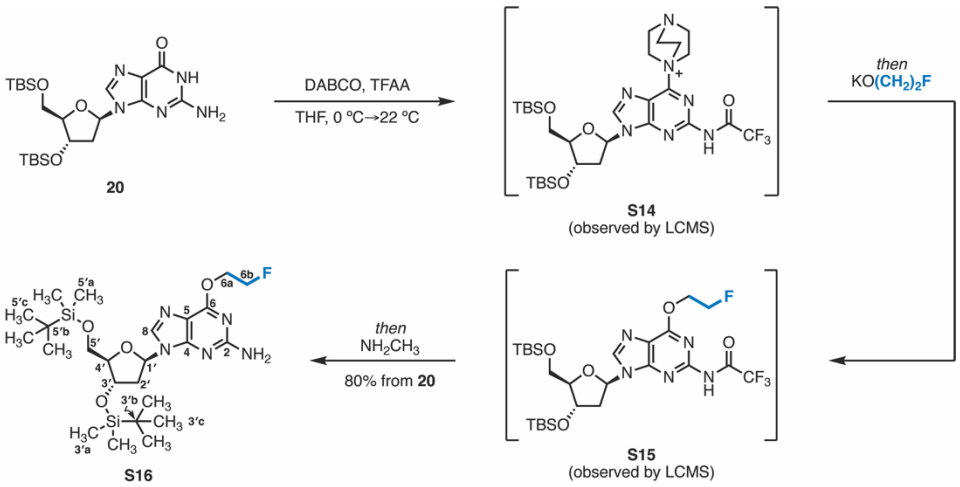
2′-deoxyguanosine (**6b**) **20**

In an adaption of a published procedure[18], *tert*-butyldimethylsilyl chloride (22.5 g, 149 mmol, 3.95 equiv) was added to a mixture of 2′-deoxyguanosine (**6b**, 10.1 g, 37.9 mmol, 1 equiv; dried by azeotropic distillation with toluene (3 × 100 mL, 50 ºC) and then overnight under vacuum) and imidazole (18.2 g, 262 mmol, 7.07 equiv) in tetrahydrofuran (150 mL) and dimethylformamide (28 mL) at 22 ºC. The inside walls of the flask were rinsed with additional dimethylformamide (18 mL). The reaction mixture was stirred for 22 h at 22 ºC. The product mixture was diluted with absolute ethanol (60 mL) and the diluted product mixture was stirred for 1 h at 22 ºC. The mixture was concentrated (50 ºC) to a constant volume. The partially concentrated solution was co-evaporated with toluene (4 × 150 mL, 50 ºC). The concentrated product mixture was suspended in water (850 mL) and stirred vigorously at 22 ºC for 1 h. The heterogenous mixture was filtered through a fritted funnel and the filter cake was washed with water (3.0 L). The washed solids were suction-dried for 30 min then rinsed sequentially with hexanes (2.0 L) and pentanes (1.0 L). The solids were collected and dried under vacuum over phosphorous pentoxide at 22 ºC to provide the bis(silyl ether) **20** as a white solid (16.7 g, 89%).

R_f = 0.61 (15% methanol–dichloromethane; UV). ^{1}H NMR (500 MHz, DMSO-d_6): δ 10.61 (s, 1H, NH), 7.88 (s, 1H, H_8), 6.47 (br s, 2H, NH_2), 6.10 (dd, J = 7.7, 6.0 Hz, 1H, $H_{1'}$), 4.48 (dt, J = 5.7, 3.0 Hz, 1H, $H_{3'}$), 3.81 (ddd, J = 5.9, 4.4, 2.6 Hz, 1H, $H_{4'}$), 3.69 (ABXdd, ^2J = 11.0 Hz, ^3J = 5.9 Hz, Δν = 25.4 Hz, 1H, $H_{5'}$), 3.64 (ABXdd, ^2J = 11.0 Hz, ^3J = 4.5 Hz, Δν = 25.4 Hz, 1H, $H_{5'}$), 2.64 (ddd, J = 13.2, 7.8, 5.6 Hz, 1H, $H_{2'}$), 2.23 (ddd, J = 13.2, 6.0, 3.2 Hz, 1H, $H_{2'}$), 0.88 (s, 9H, $H_{3'c}$ or $H_{5'c}$), 0.87 (s, 9H, $H_{3'c}$ or $H_{5'c}$), 0.10 (s, 6H, $H_{3'a}$ and/or $H_{5'a}$), 0.041 (s, 3H, $H_{3'a}$ or $H_{5'a}$), 0.035 (s, 3H, $H_{3'a}$ or $H_{5'a}$).[a] ^{13}C{^{1}H} NMR (126 MHz, DMSO-d_6): δ 156.7 (C), 153.7 (C), 151.0 (C), 134.9 (C_8), 116.6 (C), 87.0 ($C_{4'}$), 82.1 ($C_{1'}$), 72.1 ($C_{3'}$), 62.7 ($C_{5'}$), 38.9 ($C_{2'}$), 25.8 ($C_{3'c}$ or $C_{5'c}$), 25.7 ($C_{3'c}$ or $C_{5'c}$), 17.9 (C), 17.7 (C), –4.8 ($C_{3'a}$ or $C_{5'a}$), –5.0 ($C_{3'a}$ or $C_{5'a}$), –5.49 ($C_{3'a}$ or $C_{5'a}$), –5.54 ($C_{3'a}$ or $C_{5'a}$).[h] HRMS-ESI (m/z): $[M+H]^+$ calc'd for $[C_{22}H_{42}N_5O_4Si_2]^+$ 496.2770; found 496.2769 (error = 0.2 ppm).

[h] ^{1}H and ^{13}C NMR spectroscopic data for the bis(silyl) ether **20** obtained in this work were in good agreement with those previously reported.[18]

*Synthesis of the 2-fluoroethyl ether **S16**.*



Trifluoroacetic anhydride (1.40 mL, 10.1 mmol, 2.49 equiv) was added dropwise via syringe pump over 15 min to a suspension of the bis(silyl ether) **20** (2.00 g, 4.03 mmol, 1 equiv) and 1,4-diazabicyclo[2.2.2]octane (DABCO, 1.62 g, 14.4 mmol, 3.57 equiv) in tetrahydrofuran (36 mL) at 0 °C. The reaction mixture was stirred for 1.3 h at 0 °C. Additional DABCO (474 mg, 4.23 mmol, 1.05 equiv) and trifluoroacetic anhydride (360 µL, 2.55 mmol, 0.630 equiv) were added[i], and the reaction mixture was stirred for an additional 1.2 h at 0 °C, by which point LC/MS indicated conversion to the N^2-trifluoroacetyl DABCO adduct **S14**. A freshly prepared solution of potassium 2-fluoroethanolate in tetrahydrofuran (1.61 M, 16.0 mL, 25.7 mmol, 6.36 equiv)[j] was added dropwise via syringe pump over 25 min at 0 °C; the reaction thickened as the solution of potassium 2-fluoroethanolate was added. The reaction mixture was stirred for 2.4 h at 0 °C, by which point LC/MS indicated consumption of the N^2-trifluoroacetyl DABCO adduct **S14** and formation of a mixture of the N^2-trifluoroacetamide **S15** and the 2-fluoroethyl ether **S16**. Acetic acid (1.40 mL, 24.4 mmol, 6.03 equiv) was then added dropwise via syringe at 0 °C. A solution of methylamine in water (11.6 M, 17.0 mL, 197 mmol, 48.7 equiv) was then added dropwise over 20 min via syringe pump at 0 °C. The reaction mixture was stirred for 5 min at 0 °C then warmed to 22 °C. The warmed solution was stirred for 50 min at 22 °C, by which point LC/MS indicated complete conversion of the N^2-trifluoroacetamide **S15** to the 2-fluoroethyl ether **S16**. The product mixture was transferred to a separatory funnel and partitioned between dichloromethane (180 mL), water (120 mL), and saturated aqueous sodium chloride solution (200 mL). The layers that formed were separated, and the aqueous layer was extracted with dichloromethane (5 × 60 mL). The organic layers were combined, and the combined organic layers were dried over sodium sulfate. The dried solution was filtered. The filtrate was concentrated and the residue obtained was purified

[i] Additional portions of DABCO and trifluoroacetic anhydride were added because LC/MS analysis indicated the presence of a small amount of the bis(silyl ether) **20**, the N^2-trifluoroacetyl DABCO adduct **S14**, and a mass corresponding to trifluoroacetylated bis(silyl ether) **20** (not pictured).

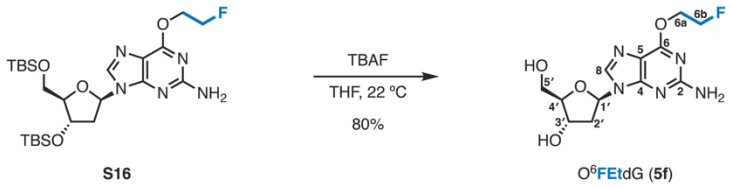
[j] The solution of potassium 2-fluoroethanolate in tetrahydrofuran was prepared immediately before use in the following manner. In a dinitrogen filled glove box, a flask was charged with potassium *tert*-butoxide (2.89 g, 25.7 mmol, 6.36 equiv). The charged flask was stoppered with a septum, brought out of the glovebox, placed under argon, and charged with tetrahydrofuran (14 mL). The resulting mixture was chilled in an ice-water bath (0 °C). Neat 2-fluoroethanol (2.00 mL, 34.1 mmol, 8.43 equiv) was added dropwise via syringe pump over 17 min to the chilled solution at 0 °C. The resulting mixture was aged for 1 h at 0 °C. The aged mixture was drawn into a syringe using a needle and added dropwise to the reaction, as described above.

by automated flash-column chromatography (eluting with 10% ethyl acetate hexanes→100% ethyl acetate[k]) to provide the 2-fluoroethyl ether **S16** as a white solid (1.75 g, 80%).

R_f = 0.37 (50% ethyl acetate–hexanes; UV). ^{1}H NMR (600 MHz, DMSO-d_6): δ 8.08 (s, 1H, H_8), 6.49 (s, 2H, NH_2), 6.21 (t, J = 6.8 Hz, 1H, $H_{1'}$), 4.79 (dt, $^2J_{HF}$ = 47.7 Hz, $^3J_{HH}$ = 4.0, 2H, H_{6b}), 4.64 (dt, $^3J_{HF}$ = 30.2 Hz, $^3J_{HH}$ = 3.9, 2H, H_{6a}), 4.52 (dt, J = 5.7, 3.1 Hz, 1H, $H_{3'}$), 3.82 (ddd, J = 6.8, 5.0, 3.1 Hz, 1H, $H_{4'}$), 3.72 (ABXdd, ^2J = 11.1 Hz, ^3J = 6.0 Hz, Δν = 44.9 Hz, 1H, $H_{5'}$), 3.65 (ABXdd, ^2J = 11.1 Hz, ^3J = 4.3 Hz, Δν = 44.9 Hz, 1H, $H_{5'}$), 2.73 (ddd, J = 13.2, 7.5, 5.7 Hz, 1H, $H_{2'}$), 2.26 (ddd, J = 13.3, 6.1, 3.3 Hz, 1H, $H_{2'}$), 0.89 (s, 9H, $H_{3'c}$ or $H_{5'c}$), 0.86 (s, 9H, $H_{3'c}$ or $H_{5'c}$), 0.11 (s, 6H, $H_{3'a}$ and/or $H_{5'a}$), 0.05 – 0.01 (m, 6H, $H_{3'a}$ and/or $H_{5'a}$). ^{13}C{^{1}H} NMR (151 MHz, DMSO-d_6): δ 159.9 (C_6), 159.7 (C_2), 154.2 (C_4), 137.6 (C_8), 113.7 (C_5), 87.0 ($C_{4'}$), 82.3 ($C_{1'}$), 81.9 (d, $^1J_{CF}$ = 166.2 Hz, C_{6b}), 72.1 ($C_{3'}$), 64.9 (d, $^2J_{CF}$ = 18.9 Hz, C_{6a}), 62.7 ($C_{5'}$), 38.7 ($C_{2'}$), 25.8 ($C_{3'c}$ or $C_{5'c}$), 25.7 ($C_{3'c}$ or $C_{5'c}$), 18.0 ($C_{3'b}$ or $C_{5'b}$), 17.7 ($C_{3'b}$ or $C_{5'b}$), –4.8 ($C_{3'a}$ or $C_{5'a}$), –5.0 ($C_{3'a}$ or $C_{5'a}$), –5.48 ($C_{3'a}$ or $C_{5'a}$), –5.53 ($C_{3'a}$ or $C_{5'a}$). ^{19}F NMR (376 MHz, DMSO-d_6): δ –222.73 (tt, J = 47.9, 30.2 Hz). IR (ATR-FTIR, film), cm^{-1}: 3495 (br w), 3332 (br w), 3203 (br w), 2954 (m), 2929 (m), 2858 (m), 1613 (s), 1583 (s), 1252 (s), 1109 (s), 1077 (s), 1027 (s), 843 (s), 777 (s). HRMS-ESI (m/z): [M+H]$^+$ calc'd for $[C_{24}H_{45}FN_5O_4Si_2]^+$ 542.2989; found 542.2978 (error = 2.0 ppm). $[a]_D^{20}$ = +1.48 (c = 0.321, $CHCl_3$).

[k] Detailed gradient: 10% B–A (528 mL), 10%→30% B–A (13 mL), 30%→50% B–A (712 mL), 50%→60% B–A (13 mL), 60% B–A →100% B (396 mL), and 100% B (9 mL), in that order; A = hexanes and B = ethyl acetate.

*Synthesis of O⁶FEtdG (**5f**).*

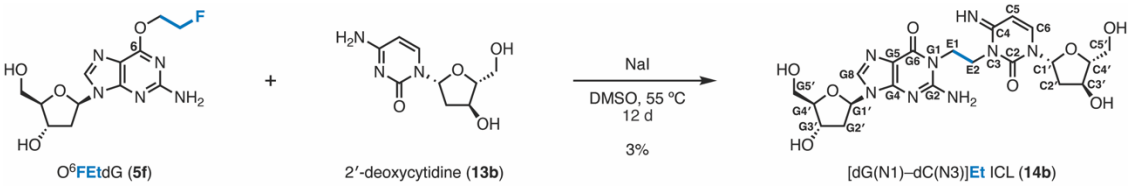


A solution of tetra-*n*-butylammonium fluoride in tetrahydrofuran (1.0 M, 920 µL, 920 µmol, 4.99 equiv) was added dropwise via syringe to a suspension of the 2-fluoroethyl ether **S16** (100 mg, 185 µmol, 1 equiv) in tetrahydrofuran (1.0 mL) at 22 ºC. The reaction mixture was stirred for 30 min at 22 ºC. Solid calcium carbonate (22.8 mg, 228 µmol, 1.23 equiv), DOWEX resin (65 mg), and methanol (2.0 mL) were then added in sequence.[l] The resulting mixture was stirred for 1 h at 22 ºC. The stirred mixture was filtered through a pad of Celite. The filtrate was concentrated and the residue obtained was purified by automated flash-column chromatography (eluting with 100% dichloromethane→50% methanol–dichloromethane)[m] to provide O⁶FEtdG (**5f**) as a white solid. The white solid was triturated with cold (–20 °C) dichloromethane (3 × 800 µL) to remove residual tetra-*n*-butylammonium salts, to provide O⁶FEtdG (**5f**) as a white solid (46.0 mg, 80%).

R_f = 0.43 (5% methanol–dichloromethane; UV, stains blue with CAM). ^{1}H NMR (600 MHz, DMSO-d_6): δ 8.11 (s, 1H, H_8), 6.47 (s, 2H, NH_2), 6.21 (dd, J = 7.8, 6.0 Hz, 1H, $H_{1'}$), 5.27 (d, J = 4.0 Hz, 1H, $C_{3'}OH$), 4.98 (t, J = 5.6 Hz, 1H, $C_{5'}OH$), 4.79 (dt, $^2J_{HF}$ = 47.8 Hz, $^3J_{HH}$ = 4.0, 2H, H_{6b}), 4.65 (dt, $^3J_{HF}$ = 30.2, $^3J_{HH}$ = 3.9, 2H, H_{6a}), 4.35 (dq, J = 6.2, 3.1 Hz, 1H, $H_{3'}$), 3.82 (td, J = 4.6, 2.7 Hz, 1H, $H_{4'}$), 3.57 (dt, J = 11.8, 5.1 Hz, 1H, $H_{5'}$), 3.50 (dt, J = 11.7, 4.8 Hz, 1H, $H_{5'}$), 2.58 (ddd, J = 13.4, 7.9, 5.7 Hz, 1H, $H_{2'}$), 2.21 (ddd, J = 13.2, 6.1, 3.1 Hz, 1H, $H_{2'}$). ^{13}C{^{1}H} NMR (151 MHz, DMSO-d_6): δ 159.9 (C_6), 159.6 (C_2), 154.1 (C_4), 138.0 (C_8), 113.8 (C_5), 87.6 ($C_{4'}$), 82.8 ($C_{1'}$), 81.9 (d, $^1J_{CF}$ = 166.1 Hz, C_{6b}), 70.8 ($C_{3'}$), 64.9 (d, $^2J_{CF}$ = 19.0 Hz, C_{6a}), 61.7 ($C_{5'}$), 39.5 ($C_{2'}$)[n]. ^{19}F NMR (376 MHz, DMSO-d_6): δ –222.70 (tt, J = 48.0, 30.3 Hz).[o] HRMS-ESI (m/z): $[M+H]^+$ calc'd for $[C_{12}H_{17}FN_5O_4]^+$ 314.1259; found 314.1252 (error = 2.2 ppm). IR (ATR-FTIR), cm^{-1}: 3344 (br m), 3217 (br m), 1616 (s), 1586 (s), 1254 (s). $[a]_D^{20}$ = –3.45 (c = 0.110, CH_3OH).

[l] This workup procedure for removing tetra-*n*-butylammonium salts was adapted from Kishi and coworkers.[19]
[m] Detailed gradient: 100% A (132 mL), 100% A→5% B–A (33 mL), 5% B–A (66 mL), 5%→10% B–A (33 mL), 10% B–A (66 mL), 10%→20% B–A (33 mL), 20% B–A (66 mL), 20%→50% B–A (33 mL), and 50% B–A (99 mL), in that order; A = dichloromethane and B = methanol.
[n] Due to overlap with the DMSO-d_6 septet in the 1-D ^{13}C{^{1}H} NMR, this ^{13}C NMR shift was estimated from an ^{1}H-^{13}C HSQC{^{1}H}{^{13}C} spectrum (DMSO-d_6, 600/151 MHz).
[o] For 2-D NMR data, please see NMR assignment table (Fig. S4).

*Synthesis of [dG(N1)–dC(N3)]Et ICL (**14b**).*

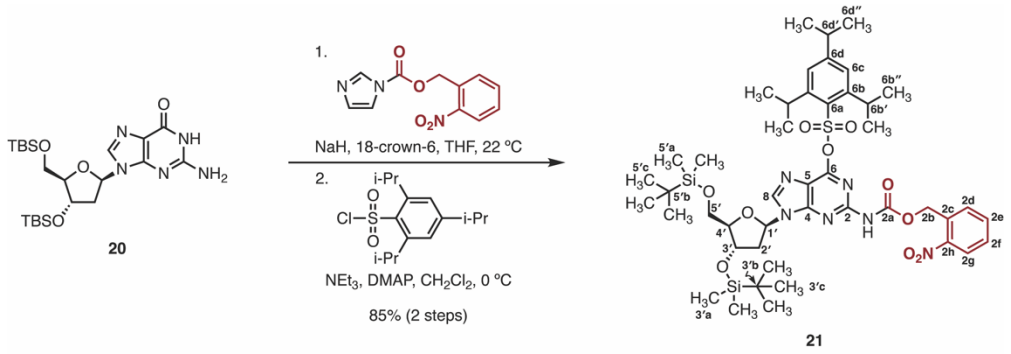


In an adaption of published procedures[12, 20], a 25-mL round-bottomed flask fused to a Teflon-coated valve[p] was charged with dried O^{6}FEtdG[q] (**5f**, 303 mg, 967 µmol, 1 equiv), dried 2′-deoxycytidine[q] (**13b**, 577 mg, 2.54 mmol, 2.63 equiv), sodium iodide (291 mg, 1.94 mmol, 2.01 equiv), and dimethyl sulfoxide (6.0 mL). The Teflon-coated valve was closed, and the reaction mixture was stirred for 12 d at 55 ºC, with protection from light. The product mixture was purified by iterative reverse phase high pressure liquid chromatography (C_{18}, 200 µL injections, 5%→20% B–A; A = 10 mM aqueous ammonium acetate and B = 100% acetonitrile).[r] Fractions containing [dG(N1)–dC(N3)]Et ICL (**14b**), as determined by LC/MS analysis, were combined and concentrated. The residue obtained was further purified by a second round of iterative reverse phase high pressure liquid chromatography (C_{18}, 50 µL injections, 8%→10% B–A; A = 10 mM aqueous ammonium acetate and B = 100% acetonitrile).[s] Fractions containing [dG(N1)–dC(N3)]Et ICL (**14b**) were combined, concentrated, and lyophilized to provide [dG(N1)–dC(N3)]Et ICL **14b** as a fluffy white powder (13.5 mg, 3%).

^{1}H NMR (600 MHz, DMSO-d_6): δ 7.92 (s, 1H, H_{G8}), 7.83 (br s, 2H, NH_2), 7.34 (d, J = 8.1 Hz, 1H, H_{C6}), 6.16 (t, J = 6.9 Hz, 1H, $H_{C1'}$), 6.12 (dd, J = 7.9, 6.0 Hz, 1H, $H_{G1'}$), 5.86 (d, J = 8.1 Hz, 1H, H_{C5}), 4.33 (dt, J = 5.6, 2.9 Hz, 1H, $H_{G3'}$), 4.21 (dt, J = 6.4, 3.4 Hz, 1H, $H_{C3'}$), 4.14 – 4.04 (br m, 2H H_{E1}), 4.05 – 3.97 (br m, 2H, H_{E2}), 3.80 (td, J = 4.7, 2.6 Hz, 1H, $H_{G4'}$), 3.75 (q, J = 3.7 Hz, 1H, $H_{C4'}$), 3.58 – 3.47 (m, 4H, $H_{G5'}$ and $H_{C5'}$), 2.51 ($H_{G2'S}$)[t], 2.18 (ddd, J = 13.1, 6.0, 2.9 Hz, 1H, $H_{G2'R}$), 2.09 – 1.99 (m, 2H, $H_{C2'}$).[u] ^{13}C{^{1}H} NMR (151 MHz, DMSO-d_6): δ 157.6 (C_{C4}), 156.4 (C_{G2} or C_{G6}), 154.3 (C_{G2} or C_{G6}), 150.1 (C_{C2}), 149.3 (C_{G4}), 135.5 (C_{G8}), 132.2 (C_{C6}), 115.6 (C_{G5}), 101.0 (C_{C5}), 87.6 ($C_{G4'}$), 87.2 ($C_{C4'}$), 84.7 ($C_{C1'}$), 82.3 ($C_{G1'}$), 70.8 ($C_{G3'}$), 70.4 ($C_{C3'}$), 61.7 ($C_{G5'}$), 61.4 ($C_{C5'}$), 39.5 ($C_{G2'}$)[v], 39.3 ($C_{C2'}$)[v], 38.5 (C_{E2}), 36.8 (C_{E1}).[u] HRMS-ESI (*m*/*z*): $[M+H]^+$ calc'd for $[C_{21}H_{29}N_8O_8]^+$ 521.2103; found 521.2111 (error = 1.5 ppm).

[p] A picture of this style of flask was previously published by Herzon and coworkers.[21]
[q] O^{6}FEtdG (**5f**) and 2′-deoxycytidine (**13b**) were dried separately by azeotropic distillation with benzene (3 ×10 mL) and then under vacuum overnight over phosphorous pentoxide at 22 ºC prior to their use in this reaction.
[r] Detailed gradient: 5% B–A (5 min to remove residual dimethyl sulfoxide) then 5%→20% B–A (30 min); A = 10 mM aqueous ammonium acetate and B = 100% acetonitrile; flow rate = 4 mL/min; 1 min automated fraction collection; fractions 8–10 contained semi-pure [dG(N1)–dC(N3)]Et ICL (**14b**).
[s] Detailed gradient: 8%→10% B–A; A = 10 mM aqueous ammonium acetate and B = 100% acetonitrile; flow rate = 4 mL/min; 1 min automated fraction collection; sub-fractions 8–10:11–15 contained pure [dG(N1)–dC(N3)]Et ICL (**14b**).
[t] Due to overlap with the $(CD_3)SO(CHD_2)$ pentet in the 1-D ^{1}H spectrum, neither multiplicity nor integration is given for this peak.
[u] For 2-D NMR data, please see NMR assignment table (Fig. S6).
[v] Due to overlap with the DMSO-d_6 septet in the 1-D ^{13}C{^{1}H} spectrum, these ^{13}C NMR shifts were estimated from an ^{1}H-^{13}C HSQC{^{1}H}{^{13}C} spectrum (DMSO-d_6, 600/151 MHz).

*Synthesis of the sulfonate **21**.*



Step 1. In an adaption of a published procedure on a similar substrate[12], a solution of 18-crown-6 (6.49 g, 24.6 mmol, 1.22 equiv) in tetrahydrofuran (24 mL)[w] was added to a suspension of sodium hydride (dry, 95%, 665 mg, 26.3 mmol, 1.30 equiv)[x] in tetrahydrofuran (120 mL) at 22 ºC. The transfer was quantitated using additional tetrahydrofuran (4 × 6.0 mL). The bis(silyl ether) **20** (10.0 g, 20.2 mmol, 1 equiv) was added to the resulting mixture at 22 ºC. The inside walls of the flask were rinsed with additional tetrahydrofuran (15 mL). The bubbling reaction mixture was stirred vigorously for 1.1 h at 22 ºC, at which time the bubbling had largely subsided. 2-Nitrobenzyl 1*H*-imidazole-1-carboxylate (7.61 g, 30.8 mmol, 1.53 equiv) was added to the reaction mixture at 22 ºC. The inside walls of the flask were rinsed with additional tetrahydrofuran (15 mL). The resulting black reaction mixture was stirred for 24 h at 22 ºC. The product mixture was diluted with glacial acetic acid (2.40 mL, 41.0 mmol, 2.03 equiv), resulting in a color change from dark brown to orange. The diluted reaction mixture was concentrated (35 ºC) and co-evaporated with dichloromethane (3 × 75 mL) to provide an orange oil. Purification by automated flash-column chromatography (eluting with 0.5%→5% methanol-dichloromethane[y]) provided the partially purified carbamate (not pictured) as a yellow foam (13.2 g) contaminated with 18-crown-6 (^{1}H NMR analysis). This material was used in the next step without further purification.

Step 2. In an adaption of a published procedure on a similar substrate[12], 2,4,6-triisopropylbenzenesulfonyl chloride (15.6 g, 51.4 mmol, 2.64 equiv) was added to a bright orange solution of the carbamate obtained in the previous step (13.2 g, nominally 19.5 mmol, 1 equiv), 4-(dimethylamino)pyridine (696 mg, 5.70 mmol, 0.292 equiv), and triethylamine (8.50 mL, 61.1 mmol, 3.13 equiv) in dichloromethane (74 mL) at 0 ºC. The inside walls of the flask were rinsed with additional dichloromethane (20 mL). The resulting yellow reaction mixture was stirred for 4.8 h at 0 ºC. The product mixture was removed from the cooling bath, Celite was added, and the resulting mixture was concentrated. The residue obtained was purified by automated flash-column

[w] The solution of 18-crown-6 in tetrahydrofuran was prepared immediately before use in the following manner. In a dinitrogen filled glove box, a flask was charged with 18-crown-6 (6.49 g). The charged flask was stoppered with a septum, brought out of the glovebox, placed under argon, and charged with tetrahydrofuran (24 mL). The resulting mixture was sonicated to give a colorless solution that was then added to the reaction flask, as described above.

[x] Sodium hydride (dry, 95%, 665 mg) was charged to the reaction flask in a dinitrogen filled glovebox. The charged flask was stopped with a septum, brought out of the glovebox, placed under argon, and charged with tetrahydrofuran (122 mL) prior to the addition of the solution of 18-crown-6 in tetrahydrofuran.

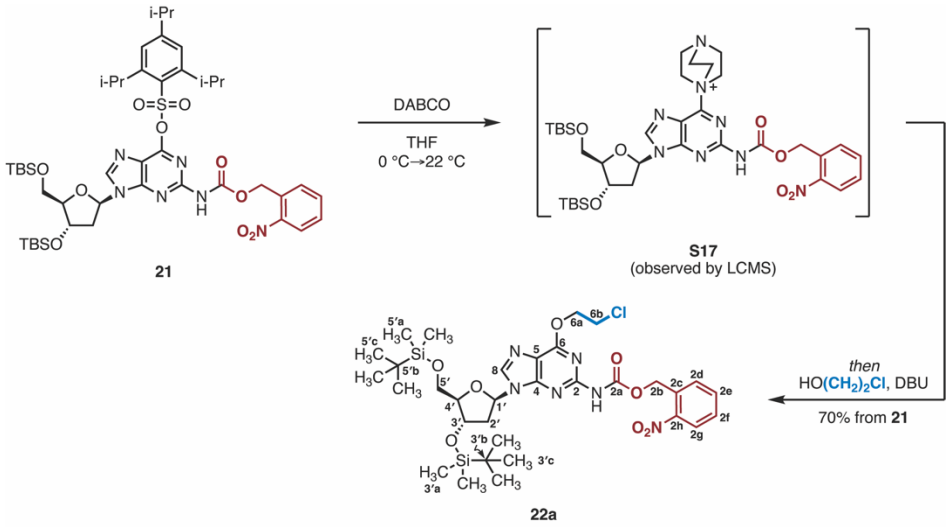
[y] Detailed gradient: 5% B–A (4230 mL) then 5%→50% B–A (5730 mL); A = dichloromethane and B = 10% methanol–dichloromethane.

chromatography (solid loading on Celite, eluting with 5%→15% ethyl acetate–hexanes[z]) to provide the sulfonate **21** as a pale yellow foam (15.5 g, 16.5 mmol, 85% over two steps).

R_f = 0.83 (50% ethyl acetate–hexanes; UV). ^{1}H NMR (600 MHz, $CDCl_3$): δ 8.19 (s, 1H, H_8), 8.15 (dd, *J* = 8.2, 1.2 Hz, 1H, H_{2g}), 7.72 – 7.65 (m, 2H, H_{2d} and H_{2e}), 7.50 (ddd, *J* = 8.5, 6.9, 1.9 Hz, 1H, H_{2f}), 7.48 (s, 1H, NH), 7.22 (s, 2H, H_{6c}), 6.39 (t, *J* = 6.4 Hz, 1H, $H_{1'}$), 5.67 (ABq, *J* = 15.2 Hz, Δν = 7.0 Hz, 2H, H_{2b}), 4.64 (dt, *J* = 6.1, 3.8 Hz, 1H, $H_{3'}$), 4.28 (hept, *J* = 6.7 Hz, 2H, $H_{6b'}$), 3.97 (q, *J* = 3.7 Hz, 1H, $H_{4'}$), 3.87 (dd, *J* = 11.2, 4.4 Hz, 1H, $H_{5'}$), 3.74 (dd, *J* = 11.2, 3.6 Hz, 1H, $H_{5'}$), 2.93 (hept, *J* = 6.9 Hz, 1H, $H_{6d'}$), 2.74 (dt, *J* = 12.9, 6.3 Hz, 1H, $H_{2'}$), 2.40 (ddd, *J* = 13.3, 6.4, 4.0 Hz, 1H, $H_{2'}$), 1.284 (d, *J* = 6.7 Hz, 6H, $H_{6b''}$), 1.279 (d, *J* = 6.8 Hz, 6H, $H_{6b''}$), 1.26 (d, *J* = 6.9 Hz, 6H, $H_{6d''}$), 0.885 (s, 9H, $H_{3'c}$ or $H_{5'c}$), 0.878 (s, 9H, $H_{3'c}$ or $H_{5'c}$), 0.08 (s, 3H, $H_{3'a}$ or $H_{5'a}$), 0.07 (s, 3H, $H_{3'a}$ or $H_{5'a}$), 0.05 (s, 3H, $H_{3'a}$ or $H_{5'a}$), 0.04 (s, 3H, $H_{3'a}$ or $H_{5'a}$). ^{13}C NMR (151 MHz, $CDCl_3$): δ 154.9 (C_2, C_4, C_6, or C_{6d}), 154.84 (C_2, C_4, C_6, or C_{6d}), 154.81 (C_2, C_4, C_6, or C_{6d}), 151.2 (C_{6b}), 150.8 (C_2 or C_6), 150.5 (C_{2a}), 147.4 (C_{2h}), 143.1 (C_8), 134.1 (C_{2e}), 132.5 (C_{2c}), 131.1 (C_{6a}), 128.9 (C_{2f}), 128.7 (C_{2d}), 125.3 (C_{2g}), 124.1 (C_{6c}), 119.7 (C_5), 88.4 ($C_{4'}$), 84.9 ($C_{1'}$), 72.0 ($C_{3'}$), 64.0 (C_{2b}), 63.0 ($C_{5'}$), 40.9 ($C_{2'}$), 34.5 ($C_{6d'}$), 30.0 ($C_{6b'}$), 26.1 ($C_{3'c}$ or $C_{5'c}$), 25.9 ($C_{3'c}$ or $C_{5'c}$), 24.78 ($C_{6b''}$), 24.75 ($C_{6b''}$), 23.7 ($C_{6d''}$), 18.6 ($C_{3'b}$ or $C_{5'b}$), 18.1 ($C_{3'b}$ or $C_{5'b}$), –4.6 ($C_{3'a}$ or $C_{5'a}$), –4.7 ($C_{3'a}$ or $C_{5'a}$), –5.3 ($C_{3'a}$ or $C_{5'a}$), –5.4 ($C_{3'a}$ or $C_{5'a}$). HRMS-ESI (*m*/*z*): $[M + Na]^+$ calc'd for $[C_{45}H_{68}N_6NaO_{10}SSi_2]^+$ 963.4148; found 963.4150 (error = 0.2 ppm). IR (ATR-FTIR), cm^{-1}: 3441 (br w), 2957 (m), 2930 (m), 2858 (m), 1767 (m), 1621 (m), 1580 (s), 1527 (s), 1463 (m), 1386 (s), 1361 (m), 1343 (m), 1259 (m), 1198 (s), 1182 (s), 1098 (s), 1075 (s), 837 (s), 778 (s). $[\alpha]_D^{20}$ = +17.63 (*c* = 0.416, $CHCl_3$).

[z] Detailed gradient: 5% B–A (2820 mL), 5%→10% B–A (940 mL), 10% B–A (1175 mL), 10%→15% B–A (470 mL), and 15% B–A (2309 mL), in that order; A = hexanes and B = ethyl acetate.

Synthesis of the N^2-NBOC protected O^6-(2-chloroethyl) ether ***22a****.*



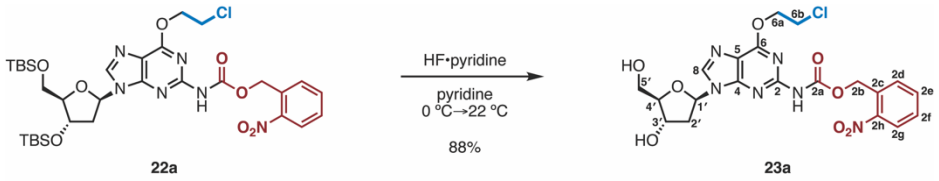
A solution of 1,4-diazabicyclo[2.2.2]octane (DABCO, 851 mg, 7.59 mmol, 1.51 equiv) in tetrahydrofuran (10 mL) was added dropwise via syringe pump over 10 min to a solution of the sulfonate **21** (4.73 g, 5.03 mmol, 1 equiv) in tetrahydrofuran (34 mL) at 0 ºC. The transfer was quantitated using additional tetrahydrofuran (3 × 2.0 mL). The reaction mixture was stirred for 15 min at 0 ºC. The cooling bath was then removed, and the reaction mixture was allowed to warm to 22 °C. The warmed mixture was stirred for 1.8 h at 22 ºC, by which point LC/MS analysis indicated consumption of the sulfonate **21** and formation of the N^2-NBOC DABCO adduct **S17**. 2-Chloroethanol (3.40 mL, 50.3 mmol, 10.0 equiv) was then added, and the resulting solution was stirred for 10 min at 22 ºC. 1,8-Diazabicyclo(5.4.0)undec-7-ene (DBU, 2.20 mL, 14.7 mmol, 2.93 equiv) was then added dropwise via syringe pump over 12 min at 22 ºC. The reaction mixture was stirred for 3.3 h at 22 ºC. The product mixture was transferred to a separatory funnel and diluted with ethyl acetate (300 mL). The diluted product mixture was washed sequentially with saturated aqueous ammonium chloride solution–water (1:1 v/v, 5 × 100 mL) and saturated aqueous sodium chloride solution (100 mL). The washed organic layer was dried over sodium sulfate. The dried solution was filtered through a pad of Celite containing a layer of sodium sulfate on top. The filtrate was collected and concentrated onto Celite. The residue obtained was purified by automated flash-column chromatography (solid loading on Celite, eluting with 10% ethyl acetate–hexanes→100% ethyl acetate[aa]) to provide the N^2-NBOC protected O^6-(2-chloroethyl) ether **22a** as light yellow foam (2.60 g, 70%).

R_f = 0.47 (50% ethyl acetate–hexanes; UV). ^{1}H NMR (600 MHz, $CDCl_3$): δ 8.145 (s, 1H, H_8), 8.142 (dd, J = 8.2, 1.3 Hz, 1H, H_{2g}), 7.72 (dd, J = 7.8, 1.4 Hz, 1H, H_{2d}), 7.68 (td, J = 7.6, 1.3 Hz, 1H, H_{2e}), 7.55 (s, 1H, NH), 7.53 – 7.49 (m, 1H, H_{2f}), 6.39 (t, J = 6.5 Hz, 1H, $H_{1'}$), 5.68 (s, 2H, H_{2b}), 4.80 (td, J = 6.2, 2.4 Hz, 2H, H_{6a}), 4.62 (dt, J = 5.9, 3.6 Hz, 1H, $H_{3'}$), 3.98 (q, J = 3.5 Hz, 1H, $H_{4'}$), 3.91 (t, J = 6.2 Hz, 2H, H_{6b}), 3.87 (ABX dd, ^2J = 11.2 Hz, ^3J = 4.2, Δν = 58.8 Hz, 1H, $H_{5'}$), 3.77 (ABX dd, ^2J = 11.2 Hz, ^3J = 3.3, Δν = 58.8 Hz, 1H, $H_{5'}$), 2.65 (ddd, J = 12.9, 6.8, 5.9 Hz, 1H, $H_{2'}$),

[aa] Detailed gradient: 10% B–A (792 mL), 10%→25% B–A (264 mL), 25% B–A (528 mL), 25%→60% B–A (396 mL), 60% B–A (765 mL), 60% B–A →100% B (264 mL), and 100% B (425 mL), in that order; A = hexanes and B = ethyl acetate.

2.40 (ddd, J = 13.1, 6.2, 3.9 Hz, 1H, $H_{2'}$), 0.902 (s, 9H, $H_{3'c}$ or $H_{5'c}$), 0.899 (s, 9H, $H_{3'c}$ or $H_{5'c}$), 0.09 (s, 3H, $H_{3'a}$ or $H_{5'}$), 0.08 (s, 3H, $H_{3'a}$ or $H_{5'}$), 0.07 (s, 6H, $H_{3'a}$ or $H_{5'}$). ^{13}C NMR (151 MHz, $CDCl_3$): δ 160.3 (C_6), 152.9 (C_4), 151.4 (C_2), 150.9 (C_{2a}), 147.5 (C_{2h}), 140.7 (C_8), 134.0 (C_{2e}), 132.5 (C_{2c}), 129.1 (C_{2d}), 129.0 (C_{2f}), 125.3 (C_{2g}), 118.5 (C_5), 88.2 ($C_{4'}$), 84.5 ($C_{1'}$), 72.0 ($C_{3'}$), 66.7 (C_{6a}), 64.0 (C_{2b}), 63.0 ($C_{5'}$), 41.3 ($C_{2'}$ or C_{6b}), 41.2 ($C_{2'}$ or C_{6b}), 26.1 ($C_{3'c}$ or $C_{5'c}$), 25.9 ($C_{3'c}$ or $C_{5'c}$), 18.6 ($C_{3'b}$ or $C_{5'b}$), 18.1 ($C_{3'b}$ or $C_{5'b}$), –4.5 ($C_{3'a}$ or $C_{5'a}$), –4.7 ($C_{3'a}$ or $C_{5'a}$), –5.2 ($C_{3'a}$ or $C_{5'a}$), –5.3 ($C_{3'a}$ or $C_{5'a}$). IR (ATR-FTIR), cm^{-1}: 2954 (m), 2929 (m), 2857 (m), 1761 (m), 1611 (m), 1587 (m), 1527 (s), 1462 (m), 1447 (m), 1418 (m), 1390 (m), 1361 (m), 1342 (m), 1305 (m), 1252 (m), 1238 (m), 1190 (s), 1103 (s), 1073 (s), 1029 (m), 1007 (m), 836 (s), 813 (m), 778 (s), 729 (m), 668 (m). HRMS-ESI (m/z): $[M+H]^+$ calc'd for $[C_{32}H_{50}ClN_6O_8Si_2]^+$ 737.2912; found 737.2882 (error = 4.1 ppm). $[a]_D^{20}$ = +14.88 (c = 0.402, $CHCl_3$).

*Synthesis of the O^6-(2-chloroethyl) diol **23a**.*



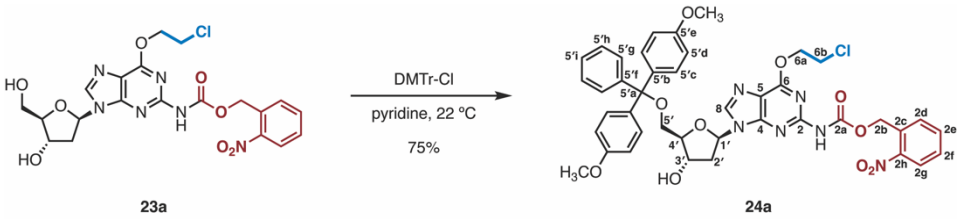
22a 23a

Hydrogen fluoride–pyridine complex (~70% HF, 3.20 mL, 24.9 mmol, 20.5 equiv) was added portion-wise over 8 min to a 50-mL Falcon tube[bb] containing a solution of the N^2-NBOC protected O^6-(2-chloroethyl) ether **22a** (893 mg, 1.21 mmol, 1 equiv) in pyridine (21 mL) at 0 ºC. The reaction mixture was stirred for 30 min at 0 ºC. The cooling bath was then removed, and the reaction mixture was allowed to warm to 22 °C. The reaction mixture was stirred for 14.9 h at 22 ºC. The product mixture was slowly added over 2 min to a vigorously stirring mixture of sodium bicarbonate (5.18 g, 61.7 mmol, 50.9 equiv), methanol (40 mL), and water (20 mL) at 0 ºC. The transfer was quantitated using methanol (40 mL). The resulting mixture was stirred for 5 min at 0 ºC. The cooling bath was then removed, and the mixture was allowed to warm to 22 ºC. The warmed mixture was stirred for 30 min at 22 °C. The product mixture was concentrated to remove most of the solvent. The partially concentrated reaction mixture was azeotroped with methanol (3 × 25 mL) to remove water. The azeotroped product mixture was purified by automated flash-column chromatography (solid loading on Celite, eluting with 1%→10% methanol–dichloromethane)[cc] to provide the O^6-(2-chloroethyl) diol **23a** as an off-white foam (544 mg, 88%).

R_f = 0.65 (15% methanol–dichloromethane; UV, stains blue with CAM). ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.14 – 8.10 (m, 1H, H_{2g}), 7.94 (s, 1H, H_8), 7.75 (s, 1H, NH), 7.71 – 7.64 (m, 2H, H_{2d} and H_{2e}), 7.51 (ddd, J = 8.6, 6.0, 2.9 Hz, 1H, H_{2f}), 6.31 (dd, J = 8.0, 6.1 Hz, 1H, $H_{1'}$), 5.66 (s, 2H, H_{2b}), 4.90 (dt, J = 5.6, 2.7, 1H, $H_{3'}$), 4.80 (t, J = 6.1 Hz, 2H, H_{6a}), 4.14 (q, J = 2.7 Hz, 1H, $H_{4'}$), 3.93 (dd, J = 12.5, 2.6, 1H, $H_{5'}$), 3.90 (t, J = 6.1 Hz, 2H, H_{6b}), 3.81 (dd, J = 12.5, 2.7 Hz, 1H, $H_{5'}$), 2.99 (ddd, J = 13.7, 8.0, 5.9 Hz, 1H, $H_{2'}$), 2.38 (ddd, J = 13.5, 6.2, 2.8 Hz, 1H, $H_{2'}$).[dd] $^{13}C\{^1H\}$ NMR (151 MHz, $CDCl_3$): δ 160.7 (C_6), 152.3 (C_4), 151.3 (C_2 or C_{2a}), 151.1 (C_2 or C_{2a}), 147.6 (C_{2h}), 141.7 (C_8), 134.0 (C_{2e}), 132.1 (C_{2c}), 129.3 (C_{2d}), 129.1 (C_{2f}), 125.3 (C_{2g}), 119.5 (C_5), 88.5 ($C_{4'}$), 86.6 ($C_{1'}$), 72.5 ($C_{3'}$), 67.0 (C_{6a}), 64.2 (C_{2b}), 63.0 ($C_{5'}$), 41.3 (C_{6b}), 40.5 ($C_{2'}$).[ee] **HRMS-ESI** (*m*/*z*): $[M+Na]^+$ calc'd for $[C_{20}H_{21}ClN_6NaO_8]^+$ 531.1002; found 531.1006 (error = 0.8 ppm).

[bb] The Falcon tube was sealed using an inverted 24/40 septum and secured with electrical tape.
[cc] Detailed gradient: 5% B–A (264 mL), 5%→15% B–A (261 mL), 15% B–A (605 mL), 15%→50% B–A (924 mL), and 50% B–A (303 mL), in that order; A = dichloromethane and B = 20% methanol–dichloromethane.
[dd] ^{1}H NMR spectroscopic data for **23a** obtained in this work were in agreement with those previously reported.[12]
[ee] ^{13}C NMR spectroscopic data for **23a** obtained in this work were in agreement with those previously reported[12] with the exception of one peak in the ^{13}C NMR. The previous report listed peaks at δ 120.1 and δ 119.2; in this region we observed only one peak at δ 119.5. Additionally, we observed a peak at δ 147.6 not reported in previous report.

*Synthesis of the O^6-(2-chloroethyl) trityl ether **24a**.*



The O^6-(2-chloroethyl) diol **23a** (649 mg, 1.28 mmol, 1 equiv) was dried by sequential azeotropic distillation with pyridine (4 × 6.0 mL, 32 ºC) and benzene (3 × 6.0 mL, 32 ºC). Pyridine (3.6 mL) was then added, to provide a pale-yellow solution. 4,4′-Dimethoxytrityl chloride (652 mg, 1.93 mmol, 1.51 equiv) was then added in one portion at 22 ºC. The inside walls of the flask were rinsed with pyridine (1.0 mL). Within 2 min, the reaction mixture became dark orange. The reaction mixture was stirred with protection from light for 4 h at 22 ºC. Absolute ethanol (1.9 mL) was then added in one portion. The diluted product mixture was stirred for 1 h at 22 ºC. The product mixture was concentrated. The concentrated product mixture was dissolved in dichloromethane (30 mL) and Celite was added. The resulting suspension was concentrated and the residue obtained was purified by automated flash-column chromatography (solid loading on Celite, eluting with 10% B–A→100% B; A = 100:1 v/v hexanes–triethylamine and B = 100:1 v/v ethyl acetate–triethylamine[ff]). Fractions containing the O^6-(2-chloroethyl) trityl ether **24a** were combined and concentrated (31 ºC). Residual triethylamine[gg] was removed by sequential co-evaporation with hexanes–ethyl acetate (1:1 v/v, 5 × 100 mL, 30 ºC) and dichloromethane (4 × 50 mL) to provide the O^6-(2-chloroethyl) trityl ether **24a** as a yellow foam (780 mg, 75%).

R_f = 0.53 (100% ethyl acetate; UV, stains orange with CAM). ^{1}H NMR (600 MHz, DMSO-d_6): δ 10.59 (s, 1H, NH), 8.35 (s, 1H, H_8), 8.19 (dd, J = 8.2, 1.2 Hz, 1H, H_{2g}), 7.88 (dd, J = 7.9, 1.5 Hz, 1H, H_{2d}), 7.84 (td, J = 7.5, 1.3 Hz, 1H, H_{2e}), 7.64 (ddd, J = 8.6, 7.3, 1.6 Hz, 1H, H_{2f}), 7.28 (dd, J 7.9, 1.9 Hz, 2H, $H_{5'g}$), 7.18 – 7.14 (m, 7H, $H_{5'c}$, $H_{5'h}$, and $H_{5'i}$), 6.77 – 6.73 (m, 2H, $H_{5'd}$), 6.71 – 6.68 (m, 2H, $H_{5'd}$), 6.37 (dd, J = 7.0, 5.7 Hz, 1H, $H_{1'}$), 5.53 (ABq, J = 15.1 Hz, Δν = 12.3 Hz, 2H, H_{2b}), 5.29 (d, J = 4.6 Hz, 1H, OH), 4.78 – 4.74 (m, 2H, H_{6a}), 4.54 (dq, J = 9.1, 4.7 Hz, 1H, $H_{3'}$), 4.08 (t, J = 5.5 Hz, 2H, H_{6b}), 3.97 (ddd, J =7.3, 4.3, 3.2 Hz, 1H, $H_{4'}$), 3.694 (s, 3H, $C_{5'e}OCH_3$), 3.690 (s, 3H, $C_{5'e}OCH_3$), 3.33 ($H_{5'}$)[hh], 3.09 (dd, J = 10.3, 3.1 Hz, 1H, $H_{5'}$), 2.89 (dt, 12.8, 6.2 Hz, 1H, $H_{2'}$), 2.33 (ddd, J = 13.5, 7.0, 4.8 Hz, 1H, $H_{2'}$).[ii] ^{13}C NMR (151 MHz, DMSO-d_6): δ 159.5 (C_6), 157.93 ($C_{5'e}$), 157.87 ($C_{5'e}$), 152.7 (C_4), 151.6 (C_2), 151.2 (C_{2a}), 146.8 (C_{2h}), 145.0 ($C_{5'f}$), 141.7 (C_8), 135.6 ($C_{5'b}$), 135.5 ($C_{5'b}$), 134.4 (C_{2e}), 132.5 (C_{2c}), 129.7 ($C_{5'c}$, $C_{5'h}$, or $C_{5'i}$), 129.6 ($C_{5'c}$, $C_{5'h}$, or $C_{5'i}$), 129.0 (C_{2f}), 128.7 (C_{2d}), 127.62 ($C_{5'c}$, $C_{5'g}$, $C_{5'h}$, or $C_{5'i}$), 127.55 ($C_{5'c}$, $C_{5'g}$, $C_{5'h}$, or $C_{5'i}$), 126.5 ($C_{5'c}$, $C_{5'h}$, or $C_{5'i}$), 124.9 (C_{2g}), 117.5 (C_5), 112.9 ($C_{5'd}$), 112.8 ($C_{5'd}$), 86.5 ($C_{4'}$), 85.3 ($C_{5'a}$), 83.7 ($C_{1'}$), 70.6 ($C_{3'}$), 66.5 (C_{6a}), 64.5 ($C_{5'}$), 62.6 (C_{2b}), 54.93 ($C_{5'e}$OCH3), 54.91 ($C_{5'e}$OCH3), 42.4

[ff] Detailed gradient: 10% B–A (396 mL), 10%→20% B–A (462 mL), 20% B–A (462 mL), 20%→50% B–A (396 mL), 50%→80% B–A (528 mL), 80% B–A (528 mL), 80% B–A→100% B (396 mL), and 100% B (2640 mL), in that order; A = 100:1 v/v hexanes–triethylamine and B = 100:1 v/v ethyl acetate–triethylamine.

[gg] Removing triethylamine is critical as residual triethylamine will neutralize the 1*H*-tetrazole used in the conversion of the O^6-(2-chloroethyl) trityl ether **24a** to the O^6-(2-chloroethyl) phosphoramidite **25a**.

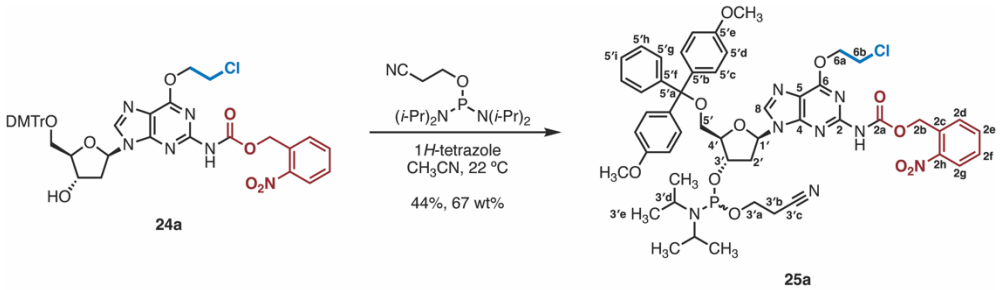
[hh] Due to overlap with the water peak in the 1-D ^{1}H spectrum, this ^{1}H shift was estimated from an ^{1}H-^{13}C HSQC{^{1}H}{^{13}C} spectrum (DMSO-d_6, 600/151 MHz) and neither multiplicity nor integration are given.

[ii] ^{1}H NMR spectroscopic data for **24a** obtained in this work were in agreement with those previously reported.[12]

(C_{6b}), 38.7 ($C_{2'}$).[jj] HRMS-ESI (*m*/*z*): $[M+H]^+$ calc'd for $[C_{41}H_{40}ClN_6O_{10}]^+$ 811.2489; found 811.2459 (error = 3.7 ppm).

[jj] ^{13}C NMR spectroscopic data for **24a** obtained in this work were in agreement with those previously reported[12] except that we observed peaks at δ 152.7 and δ 83.7 not reported in the previous report as well as the resolution of diastereotopic pairs of peaks at δ 157.93/157.87 and δ 54.93/54.91 that were not resolved in the previous report.

*Synthesis of O^6-(2-chloroethyl) phosphoramidite **25a**.*



The O^6-(2-chloroethyl) trityl ether **24a** (664 mg, 818 μmol, 1 equiv) was dried by azeotropic distillation with dry acetonitrile (6 × 10 mL, 35 ºC) and then under vacuum for 19.5 h. To the dried starting material was added acetonitrile (5.1 mL) followed by 2-cyanoethyl *N*,*N*,*N*′,*N*′-tetraisopropylphosphorodiamidite (41 wt% by qNMR, 760 μL, 981 μmol, 1.20 equiv). The reaction mixture was stirred for 10 min at 22 ºC. 1*H*-Tetrazole (3 wt% in acetonitrile, 2.90 mL, 981 μmol, 1.20 equiv) was added in one portion to the reaction mixture at 22 ºC. Within 1 min, precipitates were observed. The reaction mixture was stirred with protection from light for 5.8 h at 22 ºC. Methanol (2.2 mL) was then added. The diluted product mixture was concentrated (25 ºC) and the residue obtained was purified by automated flash-column chromatography[kk] (eluting with 20% B–A→100% B; A = 100:1 v/v hexanes–triethylamine and B = 100:1 v/v ethyl acetate–triethylamine [ll]). Fractions containing both diastereomers of the O^6-(2-chloroethyl) phosphoramidite **25a** were combined and concentrated. Residual triethylamine[mm] was removed by sequential co-evaporation with hexanes–ethyl acetate (1:1 v/v, 4 × 130 mL), hexanes (4 × 130 mL), and dichloromethane (4 × 50 mL). The co-evaporated material was dried by azeotropic distillation with dry acetonitrile (4 × 5.0 mL). The azeotroped product was co-evaporated with dry dichloromethane (4 × 5.0 mL) to provide the O^6-(2-chloroethyl) phosphoramidite **25a** as a white foam (545 mg @ 67 wt% by qNMR, 44%).

R_f = 0.80 & 0.89 (90% ethyl acetate–hexanes; UV, stains orange with CAM[nn]). ^{1}H NMR (500 MHz, CD_3CN): δ 8.50 – 8.43 (m, 1H, NH), 8.15 (dt, J = 8.2, 1.5 Hz, 1H, H_{2g}), 8.01 (s, 1H, H_8), 7.82 – 7.72 (m, 2H, H_{2d} and H_{2e}), 7.60 – 7.54 (m, 1H, H_{2f}), 7.38 – 7.27 (m, 2H, $H_{5'g}$), 7.24 – 7.10 (m, 7H, $H_{5'c}$, $H_{5'h}$ and $H_{5'i}$), 6.75 – 6.69 (m, 2H, $H_{5'd}$), 6.69 – 6.64 (m, 2H, $H_{5'd}$), 6.36 – 6.31 (m, 1H, $H_{1'}$), 5.55 (ABq, J = 15.1 Hz, Δν = 9.8 Hz, 2H, H_{2b}), 4.95 – 4.87 (m, 1H, $H_{3'}$), 4.80 – 4.75 (m, 2H, H_{6a}), 4.24 – 4.15 (m, 1H, $H_{4'}$), 4.00 (t, J = 5.6 Hz, 2H, H_{6b}), 3.81 – 3.63 (m, 8H, $H_{3'a}$ and $C_{5'e}OCH_3$), 3.61 – 3.44 (m, $H_{3'd}$ and $H_{5'}$)[oo], 3.27 – 3.19 (m, 1H, $H_{5'}$), 3.11 – 3.01 (m, 1H, $H_{2'}$), 2.62

[kk] The silica gel used to run the column was dried in an oven for >48 h then cooled in a desiccator immediately prior to use. The solvent mixtures used to run the column were prepared immediately prior to use in the following manner: for 100:1 v/v hexanes–triethylamine, a new bottle (4 L) of hexanes was sparged with dinitrogen for ≥1 h. Then an appropriate amount was measured out and combined with triethylamine (distilled from CaH_2 under dinitrogen immediately before use) to give 100:1 v/v hexanes–triethylamine. The 100:1 v/v ethyl acetate–triethylamine mixture was prepared in parallel in the same manner.

[ll] Detailed gradient: 20% B–A (330 mL), 20%→60% B–A (330 mL), 60% B–A (528 mL), 60%→75% B–A (198 mL), 75%→85% B–A (132 mL), 85% B–A (594 mL), 85% B–A→100% B (198 mL), and 100% B (273 mL), in that order; A = 100:1 hexanes–triethylamine and B = 100:1 ethyl acetate–triethylamine.

[mm] Removing triethylamine is critical as residual triethylamine will neutralize the tetrazole-based activator used in the solid phase oligonucleotide synthesis of deoxyoligonucleotide **26a**.

[nn] Orange color likely results from deprotection of the dimethoxytrityl ether to give the dimethoxytrityl cation.

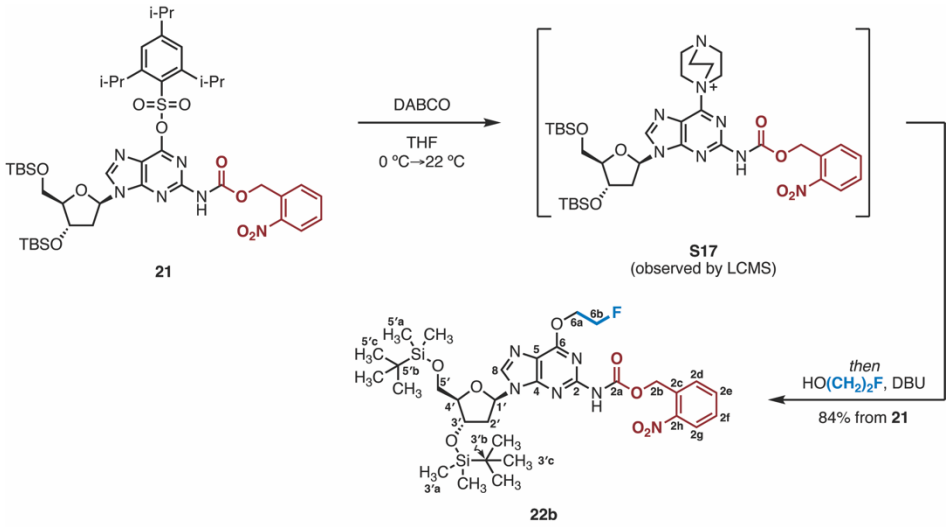
[oo] Due to overlap with impurities in the 1-D ^{1}H spectrum, it was not possible to obtain an accurate integration for this multiplet.

– 2.51 (m, 2H, $H_{2'}$ and $H_{3'b}$), 2.49 (t, J = 6.0 Hz, 1H, $H_{3'b}$), 1.15 (d, J = 6.8 Hz, 3H, $H_{3'e}$), 1.13 (d, J = 6.8 Hz, 3H, $H_{3'e}$), 1.12 (d, J = 6.8 Hz, 3H, $H_{3'e}$), 1.04 (d, J = 6.8 Hz, 3H, $H_{3'e}$). ^{13}C NMR (151 MHz, CD_3CN): δ 161.0 (C_6), 159.54 ($C_{5'e}$), 159.52 ($C_{5'e}$), 159.47 ($C_{5'e}$), 159.46 ($C_{5'e}$), 153.8 (C_4), 152.5 (C_2), 152.4 (C_2), 151.8 (C_{2a}), 151.7 (C_{2a}), 148.33 (C_{2h}), 148.29 (C_{2h}), 146.24 ($C_{5'f}$), 146.21 ($C_{5'f}$), 142.6 (C_8), 136.9 ($C_{5'b}$), 136.82 ($C_{5'b}$), 136.81 ($C_{5'b}$), 135.20 (C_{2e}), 135.19 (C_{2e}), 133.6 (C_{2c}), 133.5 (C_{2c}), 131.04 ($C_{5'c}$), 131.00 ($C_{5'c}$), 130.9 ($C_{5'c}$), 129.93 (C_{2f}), 129.90 (C_{2f}), 129.7 (C_{2d}), 129.6 (C_{2d}), 128.94 ($C_{5'g}$), 128.87 ($C_{5'g}$), 128.60 ($C_{5'h}$ or $C_{5'i}$), 128.57 ($C_{5'h}$ or $C_{5'i}$), 127.7 ($C_{5'h}$ or $C_{5'i}$), 127.6 ($C_{5'h}$ or $C_{5'i}$), 126.0 (C_{2g}), 119.6 (C5 or C3′c), 119.45 (C5 or C3′c), 119.42 (C_5 or $C_{3'c}$), 119.40 (C_5 or $C_{3'c}$), 113.801 ($C_{5'd}$), 113.797 ($C_{5'd}$), 113.75 ($C_{5'd}$), 113.73 ($C_{5'd}$), 87.24 ($C_{4'}$), 87.21 ($C_{4'}$), 87.0 ($C_{4'}$), 86.94 ($C_{4'}$), 86.86 ($C_{5'a}$), 86.8 ($C_{5'a}$), 85.9 ($C_{1'}$), 85.8 ($C_{1'}$), 74.8 ($C_{3'}$), 74.7 ($C_{3'}$), 74.6 ($C_{3'}$), 74.5 ($C_{3'}$), 67.7 (C_{6a}), 65.3 ($C_{5'}$), 65.2 ($C_{5'}$), 64.23 (C_{2b}), 64.18 (C_{2b}), 59.5 ($C_{3'a}$), 59.43 ($C_{3'a}$), 59.35 ($C_{3'a}$), 59.3 ($C_{3'a}$), 55.8 ($C_{5'e}OCH_3$), 44.1 ($C_{3'd}$), 44.02 ($C_{3'd}$), 43.97 ($C_{3'd}$), 43.9 ($C_{3'd}$), 43.1 (C_{6b}), 39.28 ($C_{2'}$), 39.26 ($C_{2'}$), 39.04 ($C_{2'}$), 39.01 ($C_{2'}$), 24.90 ($C_{3'e}$), 24.88 ($C_{3'e}$), 24.86 ($C_{3'e}$), 24.84 ($C_{3'e}$), 24.81 ($C_{3'e}$), 24.79 ($C_{3'e}$), 21.03 ($C_{3'b}$), 20.98 ($C_{3'b}$), 20.9 ($C_{3'b}$).[pp] $^{31}P\{^1H\}$ NMR (162 MHz, $CDCl_3$): δ 148.8 (s, 0.5P), 148.7 (s, 0.5P).[qq] $^{31}P\{^1H\}$ NMR (202 MHz, CD_3CN): δ 148.3 (s, 0.5P), 148.1 (s, 0.5P).[qq] HRMS-ESI (*m*/*z*): $[M+H]^+$ calc'd for $[C_{50}H_{57}ClN_8O_{11}P]^+$ 1011.3567; found 1011.3556 (error = 1.1 ppm).

[pp] Due to the similarity between the difference in shifts for each diastereomer and the magnitude of the J_{CP} coupling constants, J_{CP} values could not be confidently determined. As such, each peak is individually reported.

[qq] ^{31}P NMR spectroscopic data acquired in $CDCl_3$ for **25a** obtained in this work were in agreement with those previously reported.[12] Quantitative ^{31}P NMR spectroscopic data was acquired in CD_3CN to mitigate the possibility of trace acid in $CDCl_3$ causing dimethoxytrityl deprotection during data acquisition.

Synthesis of the N^2-NBOC protected O^6-(2-fluoroethyl) ether ***22b***.



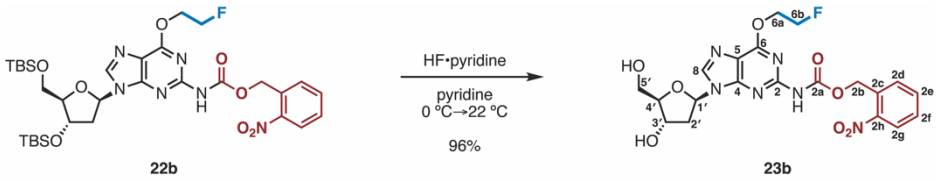
A solution of 1,4-diazabicyclo[2.2.2]octane (DABCO, 188 mg, 1.67 mmol, 1.50 equiv) in tetrahydrofuran (3.0 mL) was added in portions over 7 min to a solution of the sulfonate **21** (1.05 g, 1.11 mmol, 1 equiv) in tetrahydrofuran (7.0 mL) at 0 ºC. The transfer was quantitated with tetrahydrofuran (2 × 600 µL). The reaction mixture was stirred for 1.4 h at 0 ºC. The cooling bath was then removed, and the reaction mixture was allowed to warm to 22 °C. The warmed mixture was stirred for 1 h at 22 ºC, by which point LC/MS indicated consumption of the sulfonate **21** and formation of the N^2-NBOC DABCO adduct **S17**. 2-Fluoroethanol (640 µL, 10.9 mmol, 9.80 equiv) was added in one portion at 22 ºC. The reaction mixture was stirred for 4 min at 22 ºC. 1,8-Diazabicyclo(5.4.0)undec-7-ene (DBU, 500 µL, 3.35 mmol, 3.01 equiv) was then added in portions over 2 min. The reaction mixture was stirred for 1.9 h at 22 ºC. The product mixture was transferred to a separatory funnel and diluted with ethyl acetate (60 mL). The diluted product mixture was washed sequentially with saturated aqueous ammonium chloride solution–water (1:1 v/v, 5 × 20 mL) and saturated aqueous sodium chloride solution (20 mL). The washed organic layer was dried over sodium sulfate. The dried solution was filtered through a pad of Celite. The filtrate was concentrated. The residue obtained was purified by automated flash-column chromatography (eluting with 10% ethyl acetate–hexanes→100% ethyl acetate[rr]) to provide the 2-fluoroethyl ether **22b** as an off-white foam (673 mg, 84%).

R_f = 0.34 (50% ethyl acetate–hexanes; UV). ^{1}H NMR (600 MHz, DMSO-d_6): δ 10.65 (s, 1H, NH), 8.40 (s, 1H, H_8), 8.17 (dd, J = 8.1, 1.2 Hz, 1H, H_{2g}), 7.87 (dd, J = 8.0, 1.8 Hz, 1H, H_{2d}), 7.84 (app td, J = 7.4, 1.3 Hz, 1H, H_{2e}), 7.63 (ddd, J = 8.5, 7.0, 1.8 Hz, 1H, H_{2f}), 6.31 (t, J = 6.8 Hz, 1H, $H_{1'}$), 5.55 (s, 2H, H_{2b}), 4.89 – 4.71 (m, 4H, H_{6a} and H_{6b}), 4.66 (dt, J = 6.2, 3.3 Hz, 1H, $H_{3'}$), 3.85 – 3.78 (m, 2H, $H_{4'}$ and $H_{5'}$), 3.65 (dd, J = 10.5, 4.5 Hz, 1H, $H_{5'}$), 3.07 – 2.99 (m, 1H, $H_{2'}$), 2.27 (ddd, J = 13.2, 6.5, 3.7 Hz, 1H, $H_{2'}$), 0.88 (s, 9H, $H_{3'c}$ or $H_{5'c}$), 0.82 (s, 9H, $H_{3'c}$ or $H_{5'c}$), 0.09 (s, 3H, $H_{3'a}$ or $H_{5'a}$), 0.08 (s, 3H, $H_{3'a}$ or $H_{5'a}$), –0.026 (s, 3H, $H_{3'a}$ or $H_{5'a}$), –0.029 (s, 3H, $H_{3'a}$ or $H_{5'a}$). ^{13}C NMR (151 MHz, DMSO-d_6): δ 159.7 (C_6), 152.7 (C_4), 151.7 (C_2), 151.3 (C_{2a}), 146.8 (C_{2h}), 142.0 (C_8),

[rr] Detailed gradient: 10% B–A (395 mL), 10%→25% B–A (132 mL), 25% B–A (264 mL), 25%→60% B–A (396 mL), 60% B–A (1055 mL), 60% B–A→100% B (132 mL), and 100% B (404 mL), in that order; A = hexanes and B = ethyl acetate.

134.3 (C_{2e}), 132.6 (C_{2c}), 129.0 (C_{2f}), 128.5 (C_{2d}), 124.9 (C_{2g}), 117.6 (C_5), 87.5 ($C_{4'}$), 83.8 ($C_{1'}$), 81.7 (d, $^1J_{CF}$ = 166.0 Hz, C_{6b}), 72.3 ($C_{3'}$), 65.8 (d, $^2J_{CF}$ = 19.1 Hz, C_{6a}), 62.9 ($C_{5'}$), 62.5 (C_{2b}), 38.0 ($C_{2'}$), 25.72 ($C_{3'c}$ or $C_{5'c}$), 25.68 ($C_{3'c}$ or $C_{5'c}$), 17.9 ($C_{3'b}$ or $C_{5'b}$), 17.7 ($C_{3'b}$ or $C_{5'b}$), –4.8 ($C_{3'a}$ or $C_{5'a}$), –5.0 ($C_{3'a}$ or C5′a), –5.6 (2×CH_3, $C_{3'a}$ and/or $C_{5'a}$). ^{19}F NMR (376 MHz, DMSO-d_6): δ –223.16 (tt, J = 47.6, 30.1 Hz). IR (ATR-FTIR), cm^{-1}: 2953 (m), 2929 (m), 2858 (m), 1761 (m), 1611 (m), 1588 (m), 1527 (s), 1462 (m), 1344 (m), 1251 (m), 1191 (br s), 1097 (s), 1075 (s), 873 (s), 778 (s). HRMS-ESI (m/z): $[M+Na]^+$ calc'd for $[C_{32}H_{49}FN_6NaO_8Si_2]^+$ 743.3027; found 743.3021 (error = 0.8 ppm). $[a]_D^{20}$ = +13.41 (c = 0.362, $CHCl_3$).

*Synthesis of the O^6-(2-fluoroethyl) diol **23b**.*

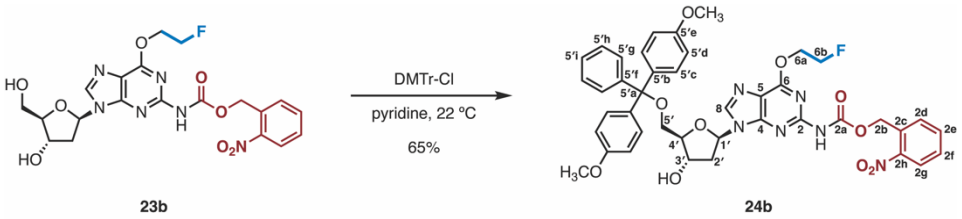


Hydrogen fluoride–pyridine complex (~70% HF, 2.40 mL, 18.6 mmol, 20.0 equiv) was added portionwise over 11 min to a 50-mL Falcon tube[ss] containing a solution of the N^2-NBOC protected O^6-(2-fluoroethyl) ether **22b** (671 mg, 931 µmol, 1 equiv) in pyridine (16 mL) at 0 ºC. The reaction mixture was stirred for 38 min at 0 ºC. The cooling bath was then removed, and the reaction mixture was allowed to warm to 22 °C. The warmed reaction mixture was stirred for 16.5 h at 22 °C. The product mixture was added over 2 min to a stirring mixture of sodium bicarbonate (3.80 g, 45.3 mmol, 48.6 equiv), methanol (30 mL), and water (15 mL). The transfer was quantitated with methanol (40 mL). The resulting mixture was stirred for 5 min at 0 ºC. The cooling bath was then removed, and the mixture was allowed to warm to 22 °C. The warmed mixture was stirred for 30 min at 22 °C. The product mixture was concentrated to remove most of the solvent. The partially concentrated reaction mixture was azeotroped with methanol (3 × 20 mL, 36 ºC) to remove water. The azeotroped product mixture was purified by automated flash-column chromatography (solid loading on Celite, eluting with 1%→10% methanol–dichloromethane[tt]) to provide the O^6-(2-fluoroethyl) diol **23b** as an off-white foam (440 mg, 96%).

R_f = 0.65 (15% methanol–dichloromethane; UV). ^{1}H NMR (600 MHz, DMSO-d_6): δ 10.63 (s, 1H, NH), 8.46 (s, 1H, H_8), 8.17 (dd, J = 8.2, 1.2 Hz, 1H, H_{2g}), 7.89 (dd, J = 7.8, 1.6 Hz, 1H, H_{2d}), 7.85 (td, J = 7.5, 1.3 Hz, 1H, H_{2e}), 7.64 (ddd, J = 8.4, 7.3, 1.6 Hz, 1H, H_{2f}), 6.33 (dd, J = 7.5, 6.2 Hz, 1H, $H_{1'}$), 5.54 (s, 2H, H_{2b}), 5.31 (d, J = 4.0 Hz, 1H, $C_{3'}$OH), 4.89 (t, J = 5.5 Hz, 1H, $C_{5'}$OH), 4.88 – 4.71 (m, 4H, H_{6a} and H_{6b}), 4.42 (dq, J = 6.5, 3.3 Hz, 1H, $H_{3'}$), 3.85 (td, J = 4.8, 2.8 Hz, 1H, $H_{4'}$), 3.60 (dt, J = 11.6, 5.3 Hz, 1H, $H_{5'}$), 3.52 (dt, J = 11.7, 5.1 Hz, 1H, $H_{5'}$), 2.73 (ddd, J = 13.2, 7.6, 5.8 Hz, 1H, $H_{2'}$), 2.28 (ddd, J = 13.2, 6.2, 3.3 Hz, 1H, $H_{2'}$). ^{13}C NMR (151 MHz, DMSO-d_6): δ 159.6 (C_6), 152.9 (C_4), 151.8 (C_2), 151.5 (C_{2a}), 147.0 (C_{2h}), 141.3 (C_8), 134.3 (C_{2e}), 132.4 (C_{2c}), 129.1 (C_{2f}), 128.9 (C_{2d}), 124.9 (C_{2g}), 117.2 (C_5), 87.9 ($C_{4'}$), 83.3 ($C_{1'}$), 81.7 (d, $^1J_{CF}$ = 166.0 Hz, C_{6b}), 70.7 ($C_{3'}$), 65.8 (d, $^2J_{CF}$ = 19.1 Hz, C_{6a}), 62.7 (C_{2b}), 61.6 ($C_{5'}$), 39.2 ($C_{2'}$)[uu]. ^{19}F NMR (376 MHz, DMSO-d_6): δ –223.19 (tt, J = 47.4, 30.1 Hz). IR (ATR-FTIR), cm^{-1}: 3245 (br w), 2951 (w), 1742 (m), 1609 (m), 1589 (m), 1521 (s), 1460 (m), 1441 (m), 1418 (m), 1393 (m), 1342 (m), 1308 (m), 1192 (s), 1094 (s), 1045 (s), 997 (m), 933 (m), 859 (m), 791 (m), 762 (m), 728 (s), 637 (m). HRMS-ESI (m/z): $[M + Na]^+$ calc'd for $[C_{20}H_{21}FN_6NaO_8]^+$ 515.1297; found 515.1281 (error = 3.1 ppm). $[a]_D^{20}$ = +13.33 (c = 0.350, CH_3OH).

[ss] The Falcon tube was sealed using an inverted 24/40 septum and secured with electrical tape.
[tt] Detailed gradient: 5% B–A (264 mL), 5%→15% B–A (264 mL), 15% B–A (528 mL), 15%→50% B–A (924 mL), and 50% B–A (385 mL), in that order; A = dichloromethane and B = 20% methanol–dichloromethane.
[uu] Due to overlap with the DMSO-d_6 septet in the 1-D ^{13}C{^{1}H} spectrum, this ^{13}C NMR shift was estimated from an ^{1}H-^{13}C HSQC{^{1}H}{^{13}C} spectrum (DMSO-d_6, 600/151 MHz).

*Synthesis of the O^6-(2-fluoroethyl) trityl ether **24b**.*



23b **24b**

The O^6-(2-fluoroethyl) diol **23b** (848 mg, 1.72 mmol, 1 equiv) was dried by sequential azeotropic distillation with pyridine (4 × 6.0 mL, 35 ºC) and benzene (2 × 6.0 mL, 35 ºC). The dried residue was dissolved in pyridine (4.4 mL), to provide a pale orange solution. 4,4′-Dimethoxytrityl chloride (877 mg, 2.59 mmol, 1.50 equiv) was added in one portion at 22 ºC. The inside walls of the flask were rinsed with pyridine (2.0 mL). Within 2 min, a dark orange solution was formed. The reaction mixture was stirred with protection from light for 4.6 h at 22 ºC. Absolute ethanol (2.5 mL) was then added in one portion. The diluted mixture was stirred for 45 min at 22 ºC. The product mixture concentrated to near dryness. The concentrated product mixture was dissolved in dichloromethane (30 mL) and Celite was added. The resulting suspension was concentrated, and the residue obtained was purified by automated flash-column chromatography (solid loading on Celite, eluting with 10% B–A→100% B; A = 100:1 v/v hexanes–triethylamine and B = 100:1 v/v ethyl acetate–triethylamine[vv]). Fractions containing the O^6-(2-fluoroethyl) trityl ether **24b** were combined and concentrated (30 ºC). Residual triethylamine[ww] was removed by sequential co-evaporation with hexanes–ethyl acetate (1:1 v/v, 5 × 100 mL) and dichloromethane (4 × 50 mL) to provide the O^6-(2-fluoroethyl) trityl ether **24b** as a light yellow foam (893 mg, 65%).

R_f = 0.60 (100% ethyl acetate; UV, stains orange with CAM). ^{1}H NMR (600 MHz, DMSO-d_6): δ 10.57 (s, 1H, NH), 8.35 (s, 1H, H_8), 8.18 (dd, J = 8.2, 1.3 Hz, 1H, H_{2g}), 7.89 (dd, J = 7.8, 1.4 Hz, 1H, H_{2d}), 7.84 (app td, J = 7.5, 1.3 Hz, 1H, H_{2e}), 7.64 (app td, J = 7.8, 1.5 Hz, 1H, H_{2f}), 7.28 (dd, J = 8.2, 1.7 Hz, 2H, $H_{5'g}$), 7.20 – 7.11 (m, 7H, $H_{5'c}$, $H_{5'h}$, and $H_{5'i}$), 6.75 (d, J = 8.9, 2H, $H_{5'd}$), 6.69 (d, J = 9.0, 2H, $H_{5'd}$), 6.38 (dd, J = 7.0, 5.7 Hz, 1H, $H_{1'}$), 5.53 (ABq, J = 15.1 Hz, Δν = 12.1 Hz, 2H, H_{2b}), 5.29 (d, J = 4.6 Hz, 1H, OH), 4.85 (dt, $^2J_{HF}$ = 47.8 Hz, $^3J_{HH}$ = 4.0 Hz, 2H, H_{6b}), 4.80 – 4.72 (m, $^3J_{HF}$ = 30.0 Hz, 2H, H_{6a})[xx], 4.55 (app dt, J = 9.2, 4.6 Hz, 1H, $H_{3'}$), 3.98 (app dt, J = 7.3, 3.7 Hz, 1H, $H_{4'}$), 3.693 (s, 3H, $H_{5'}OCH_3$), 3.688 (s, 3H, $H_{5'}OCH_3$), 3.37 – 3.30 (m, $H_{5'}$)[yy], 3.10 (dd, J = 10.4, 3.1 Hz, 1H, $H_{5'}$), 2.94 – 2.85 (m, 1H, $H_{2'}$), 2.34 (ddd, J = 13.5, 7.1, 4.9 Hz, 1H, $H_{2'}$). ^{13}C NMR (151 MHz, DMSO-d_6): δ 159.6 (C_6), 157.94 ($C_{5'e}$), 157.88 ($C_{5'e}$), 152.7 (C_4), 151.6 (C_2), 151.2 (C_{2a}), 146.8 (C_{2h}), 145.0 ($C_{5'f}$), 141.7 (C_8), 135.62 ($C_{5'b}$), 135.56 ($C_{5'b}$), 134.3 (C_{2e}), 132.5 (C_{2c}), 129.7 ($C_{5'c}$, $C_{5'h}$, or $C_{5'i}$), 129.6 ($C_{5'c}$, $C_{5'h}$, or $C_{5'i}$), 129.0 (C_{2f}), 128.7 (C_{2d}), 127.64 ($C_{5'c}$, $C_{5'g}$, $C_{5'h}$, or $C_{5'i}$), 127.57 ($C_{5'c}$, $C_{5'g}$, $C_{5'h}$, or $C_{5'i}$), 126.5 ($C_{5'c}$, $C_{5'h}$, or $C_{5'i}$), 124.9 (C_{2g}), 117.6 (C_5), 112.9 ($C_{5'd}$), 112.8 ($C_{5'd}$), 86.5 ($C_{4'}$), 85.3 ($C5'_a$), 83.7 ($C_{1'}$), 81.7 (d, $^1J_{CF}$ = 166.0 Hz, C_{6b}), 70.6 ($C_{3'}$), 65.8 (d, $^2J_{CF}$ = 19.1 Hz, C_{6a}), 64.6 ($C_{5'}$), 62.6 (C_{2b}), 54.92 ($C_{5'e}OCH_3$), 54.89 ($C_{5'e}OCH_3$), 38.7 ($C_{2'}$). ^{19}F

[vv] Detailed gradient: 10% B–A (396 mL), 10%→20% B–A (462 mL), 20% B–A (462 mL), 20%→50% B–A (396 mL), 50%→80% B–A (528 mL), 80% B–A (528 mL), 80% B–A→100% B (396 mL), and 100% B (2588 mL), in that order; A = 100:1 v/v hexanes–triethylamine and B = 100:1 v/v ethyl acetate–triethylamine.

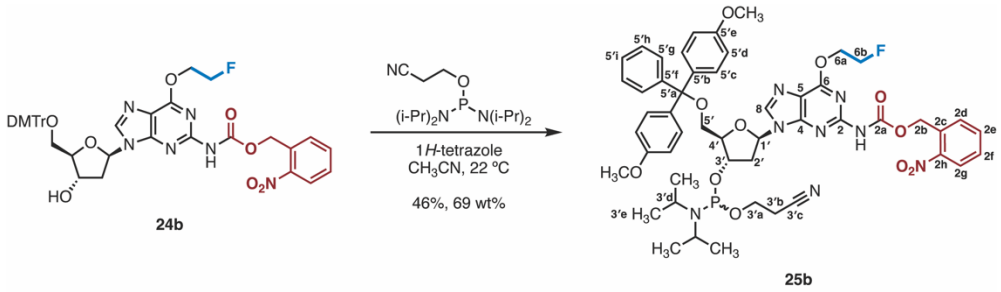
[ww] Removing triethylamine is critical as residual triethylamine will neutralize the 1*H*-tetrazole used in the conversion of the O^6-(2-fluoroethyl) trityl ether **24b** to the O^6-(2-fluoroethyl) phosphoramidite **25b**.

[xx] The $^3J_{HF}$ coupling constant was identifiable, but the J_{HH} coupling pattern is a multiplet.

[yy] Due to overlap with the water peak in the 1-D ^{1}H spectrum, it was not possible to obtain an accurate integration for this peak.

NMR (376 MHz, DMSO-d_6): δ –223.09 (tt, J = 47.8, 30.1 Hz). IR (ATR-FTIR), cm^{-1}: 3263 (br w), 2931 (w), 1754 (br m), 1608 (m), 1586 (m), 1524 (m), 1508 (s), 1459 (m), 1446 (m), 1417 (m), 1342 (m), 1303 (m), 1245 (s), 1175 (s), 1071 (m), 1029 (s), 827 (m), 791 (m), 726 (s), 701(m), 698 (m), 583 (s). HRMS-ESI (m/z): $[M + Na]^+$ calc'd for $[C_{41}H_{39}FN_6NaO_{10}]^+$ 795.2784; found 795.2787 (error = 0.4 ppm). $[a]_D^{20}$ = +27.27 (c = 0.427, CH_3OH).

*Synthesis of O^6-(2-fluoroethyl) phosphoramidite **25b**.*



The O^6-(2-fluoroethyl) trityl ether **24b** (455 mg, 573 µmol, 1 equiv) was dried by azeotropic distillation with acetonitrile (6 × 6.0 mL, 35 ºC) and then under vacuum for 14.3 h. To the dried starting material was added acetonitrile (3.6 mL) followed by 2-cyanoethyl *N*,*N*,*N*′,*N*′-tetraisopropylphosphorodiamidite (44 wt% by qNMR, 530 µL, 734 µmol, 1.28 equiv). The reaction mixture was stirred for 13 min at 22 ºC. 1*H*-Tetrazole (3 wt% in acetonitrile, 2.00 mL, 677 µmol, 1.18 equiv) was then added in one portion. Within 2 min, the reaction mixture transformed from a pale yellow solution to a white suspension. The reaction mixture was stirred for 7.9 h at 22 ºC. Methanol (1.5 mL) was then added in one portion. The diluted product mixture was stirred for 30 min and then concentrated (25 ºC). The residue obtained was purified by automated flash-column chromatography[zz] (20% B–A→100% B; A = 100:1 v/v hexanes–triethylamine and B = 100:1 v/v ethyl acetate–triethylamine[aaa]). Fractions containing both diastereomers of the O^6-(2-fluoroethyl) phosphoramidite **25b** were combined and concentrated (25 ºC). Residual triethylamine[bbb] was removed by sequential co-evaporations with hexanes–ethyl acetate (1:1 v/v, 5 × 100 mL, 25 ºC) and dichloromethane (4 × 50 mL, 25 ºC). The co-evaporated material was dried by azeotropic distillation with acetonitrile (4 × 3.0 mL, 30 ºC). The azeotroped product was co-evaporated with dry dichloromethane (3 × 3.0 mL, 30 ºC) to provide the O^6-(2-fluoroethyl) phosphoramidite **25b** as a white foam (380 mg @ 69 wt% by qNMR, 264 µmol, 46% yield).

R_f = 0.78 & 0.87 (90% ethyl acetate–hexanes; UV, stains orange with CAM[ccc]). ^{1}H NMR (500 MHz, CD_3CN): δ 8.48 – 8.42 (m, 1H, NH), 8.15 (br dt, J = 8.2, 1.6 Hz, 1H, H_{2g}), 8.01 (s, 1H, H_8), 7.82 – 7.72 (m, 2H, H_{2d} and H_{2e}), 7.60 – 7.54 (m, 1H, H_{2f}), 7.36 – 7.30 (m, 2H, $H_{5'g}$), 7.24 – 7.12 (m, 7H, $H_{5'c}$, $H_{5'h}$, and $H_{5'i}$), 6.75 – 6.69 (m, 2H, $H_{5'd}$), 6.69 – 6.63 (m, 2H, $H_{5'd}$), 6.36 – 6.32 (m, 1H, $H_{1'}$), 5.55 (ABq, J = 15.2 Hz, Δν = 10.0 Hz, 2H, H_{2b}), 4.95 – 4.71 (m, 5H, H_{6a}[ddd], H_{6b}[ddd], and

[zz] The silica gel used to run the column was dried in an oven for >48 h then cooled in a desiccator immediately prior to use. The solvent mixtures used to run the column were prepared immediately prior to use in the following manner: for 100:1 v/v hexanes–triethylamine, a new bottle (4 L) of hexanes was sparged with dinitrogen for ≥1 h. Then an appropriate amount was measured out and combined with triethylamine (distilled from CaH_2 under dinitrogen immediately before use) to give 100:1 v/v hexanes–triethylamine. The 100:1 v/v ethyl acetate–triethylamine mixture was prepared in parallel in the same manner.

[aaa] Detailed gradient: 20% B–A (132 mL), 20%→60% B–A (198 mL), 60% B–A (330 mL), 60%→75% B–A (132 mL), 75%→85% B–A (66 mL), 85% B–A (363 mL), 85% B–A→100% B (132 mL), and 100% B (1370 mL), in that order; A = 100:1 v/v hexanes–triethylamine and B = 100:1 v/v ethyl acetate–triethylamine.

[bbb] Removing triethylamine is critical as residual triethylamine will neutralize the tetrazole-based activator used in the solid phase oligonucleotide synthesis of deoxyoligonucleotide **26b**.

[ccc] Orange color likely results from deprotection of the dimethoxytrityl ether to give the dimethoxytrityl cation.

[ddd] $^2J_{HF}$ and $^3J_{HF}$ coupling constants could not be determined for H_{6b} and H_{6a}, respectively, from the 1-D ^{1}H spectrum due to spectral complexity. The magnitudes of these coupling constants as estimated from an ^{1}H-^{13}C HSQC{^{1}H}{^{13}C} spectrum (CD_3CN, 600/151 MHz) were $^2J_{HF}$ ~ 47 Hz for H_{6b} and $^3J_{HF}$ ~ 28 Hz for H_{6a}.

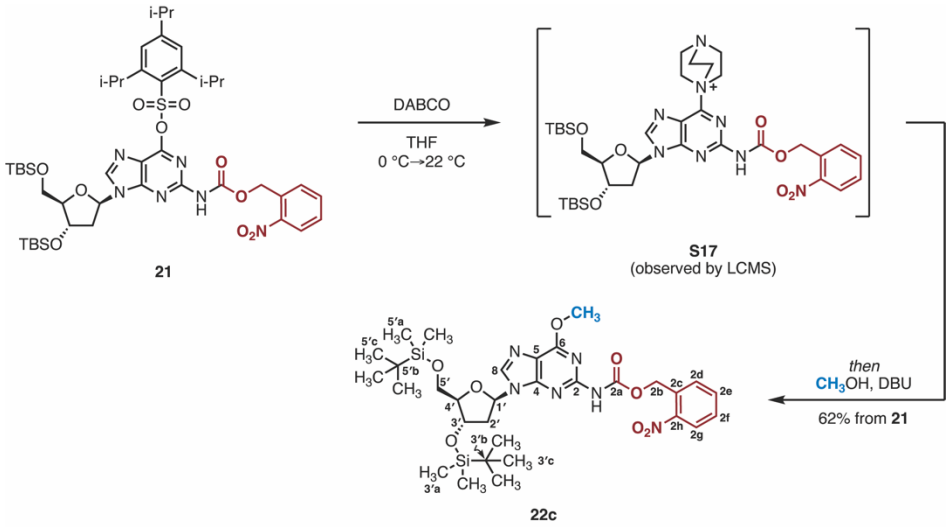
$H_{3'}$), 4.25 – 4.15 (m, 1H, $H_{4'}$), 3.80 – 3.63 (m, 8H, $H_{3'a}$ and $C_{5'e}OCH_3$), 3.62 – 3.44 (m, $H_{3'd}$ and $H_{5'}$)[eee], 3.28 – 3.19 (m, 1H, $H_{5'}$), 3.11 – 3.02 (m, 1H, $H_{2'}$), 2.66 – 2.51 (m, $H_{2'}$ and $H_{3'b}$)[eee], 2.49 (t, J = 6.0 Hz, 1H, $H_{3'b}$), 1.15 (d, J = 6.8 Hz, 3H, $H_{3'e}$), 1.13 (d, J = 6.8 Hz, 3H, $H_{3'e}$), 1.12 (d, J = 6.8 Hz, 3H, $H_{3'e}$), 1.04 (d, J = 6.8 Hz, 3H, $H_{3'e}$). ^{13}C NMR (151 MHz, CD_3CN): δ 161.2 (C_6), 159.54 ($C_{5'e}$), 159.53 ($C_{5'e}$), 159.46 ($C_{5'e}$), 159.45 ($C_{5'e}$), 153.77 (C_4), 153.76 (C_4), 152.5 (C_2), 152.4 (C_2), 151.8 (C_{2a}), 151.7 (C_{2a}), 148.34 (C_{2h}), 148.29 (C_{2h}), 146.25 ($C_{5'f}$), 146.22 ($C_{5'f}$), 142.5 (C_8), 136.9 ($C_{5'b}$), 136.84 ($C_{5'b}$), 136.8 ($C_{5'b}$), 136.78 ($C_{5'b}$), 135.2 (C_{2e}), 135.18 (C_{2e}), 133.6 (C_{2c}), 133.5 (C_{2c}), 131.1 ($C_{5'c}$), 131.0 ($C_{5'c}$), 130.9 ($C_{5'c}$), 129.93 (C_{2f}), 129.9 (C_{2f}), 129.7 (C_{2d}), 129.6 (C_{2d}), 128.94 ($C_{5'g}$), 128.87 ($C_{5'g}$), 128.6 ($C_{5'h}$ or $C_{5'i}$), 128.58 ($C_{5'h}$ or $C_{5'i}$), 127.7 ($C_{5'h}$ or $C_{5'i}$), 127.6 ($C_{5'h}$ or $C_{5'i}$), 126.0 (C_{2g}), 119.57 (C_5 or $C_{3'c}$)[fff], 119.55 (C_5 or $C_{3'c}$)[fff], 119.52 (C_5 or $C_{3'c}$)[fff], 119.4 (C_5 or $C_{3'c}$)[fff], 113.8 ($C_{5'd}$), 113.75 ($C_{5'd}$), 113.73 ($C_{5'd}$), 87.24 ($C_{4'}$), 87.21 ($C_{4'}$), 87.0 ($C_{4'}$), 86.94 ($C_{4'}$), 86.86 ($C_{5'a}$), 86.8 ($C_{5'a}$), 85.84 ($C_{1'}$), 85.81 ($C_{1'}$), 82.9 (d, $^1J_{CF}$ = 166.5 Hz, C_{6b}), 74.8 ($C_{3'}$), 74.7 ($C_{3'}$), 74.6 ($C_{3'}$), 74.5 ($C_{3'}$), 67.0 (d, $^2J_{CF}$ = 19.5 Hz, C_{6a}), 65.3 ($C_{5'}$), 65.2 ($C_{5'}$), 64.23 (C_{2b}), 64.18 (C_{2b}), 59.5 ($C_{3'a}$), 59.43 ($C_{3'a}$), 59.35 ($C_{3'a}$), 59.3 ($C_{3'a}$), 55.82 ($C_{5'e}OCH_3$), 55.80 ($C_{5'e}OCH_3$), 55.79 ($C_{5'e}OCH_3$), 44.1 ($C_{3'd}$), 44.02 ($C_{3'd}$), 43.97 ($C_{3'd}$), 43.9 ($C_{3'd}$), 39.28 ($C_{2'}$), 39.26 ($C_{2'}$), 39.04 ($C_{2'}$), 39.01 ($C_{2'}$), 24.9 ($C_{3'e}$), 24.88 ($C_{3'e}$), 24.86 ($C_{3'e}$), 24.84 ($C_{3'e}$), 24.81 ($C_{3'e}$), 24.79 ($C_{3'e}$), 21.03 ($C_{3'b}$), 20.98 ($C_{3'b}$), 20.9 ($C_{3'b}$).[ggg] ^{19}F NMR (471 MHz, CD_3CN): δ –225.10 (tt, J = 47.4, 29.7 Hz). ^{31}P{^{1}H} NMR (202 MHz, CD_3CN): δ 148.3 (s, 0.5P), 148.1 (s, 0.5P). HRMS-ESI (m/z): $[M + Na]^+$ calc'd for $[C_{50}H_{56}FN_8NaO_{11}P]^+$ 1017.3682; found 1017.3683 (error = 0.1 ppm).

[eee] Due to overlap with impurities in the 1-D ^{1}H spectrum, it was not possible to obtain an accurate integration for these multiplets.

[fff] Of the peaks δ_C 119.6, 119.55, 119.52, and 119.4, one may or may not be assignable to C_5 or $C_{3'c}$.

[ggg] Due to the similarity between the difference in shifts for each diastereomer and the magnitude of the J_{CP} coupling constants, J_{CP} values could not be confidently determined. As such, each peak is individually reported.

*Synthesis of the N^2-NBOC protected O^6-methyl ether **22c**.*



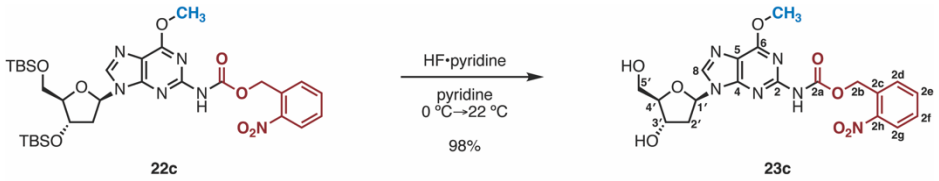
A solution of 1,4-diazabicyclo[2.2.2]octane (DABCO, 106 mg, 945 µmol, 1.76 equiv) in tetrahydrofuran (0.80 mL) was added in portions over 5 min to a solution of the sulfonate **21** (506 mg, 538 µmol, 1 equiv) in tetrahydrofuran (4.0 mL) at 0 ºC. The transfer was quantitated using tetrahydrofuran (2 × 200 µL). The reaction mixture was stirred for 2 h at 0 ºC. The cooling bath was removed, and the reaction mixture was allowed to warm to 22 °C. The warmed mixture was stirred for 27 min at 22 ºC, by which point LC/MS analysis indicated consumption of the sulfonate **4** and formation of the N^2-NBOC DABCO adduct **S17**. Methanol (330 µL, 8.15 mmol, 15.2 equiv) was then added. The reaction mixture was stirred for 2 min at 22 ºC. 1,8-Diazabicyclo(5.4.0)undec-7-ene (DBU, 240 µL, 1.61 mmol, 2.99 equiv) was then added in portions over 8 min at 22 ºC. The reaction mixture was stirred for 2.5 h at 22 °C. The product mixture was transferred to a separatory funnel and diluted with ethyl acetate (35 mL). The diluted product mixture was washed sequentially with saturated aqueous ammonium chloride solution–water (1:1 v/v, 4 × 10 mL) and saturated aqueous sodium chloride solution (10 mL). The washed organic layer was dried over sodium sulfate. The dried solution was filtered through a pad of sodium sulfate. The filtrate was concentrated. The resulting residue was purified by automated flash-column chromatography (solid loading on Celite, eluting with 10% ethyl acetate–hexanes→100% ethyl acetate[hhh]) to provide the N^2-NBOC protected O^6-methyl ether **22c** as an off-white foam (229 mg, 62%).

R_f = 0.28 (50% ethyl acetate–hexanes; UV). ^{1}H NMR (600 MHz, $CDCl_3$): δ 8.22 (s, 1H, H_8), 8.14 (dd, J = 8.2, 1.4 Hz, 1H, H_{2g}), 7.73 (d, J = 7.8 Hz, 1H, H_{2d}), 7.67 (td, J = 7.6, 1.4 Hz, 1H, H_{2e}), 7.58 (s, 1H, NH), 7.53 – 7.48 (m, 1H, H_{2f}), 6.41 (t, J = 6.4 Hz, 1H, $H_{1'}$), 5.69 (s, 2H, H_{2b}), 4.62 (dt, J = 5.6, 3.8 Hz, 1H, $H_{3'}$), 4.14 (s, 3H, C_6OCH_3), 3.99 (q, J = 3.7 Hz, 1H, $H_{4'}$), 3.87 (ABXdd, $^2J_{AB}$ = 11.2 Hz, $^3J_{AX}$ = 4.2 Hz, Δν = 59.2 Hz, 1H, $H_{5'}$), 3.77 (ABXdd, $^2J_{AB}$ = 11.2 Hz, $^3J_{AX}$ = 3.3 Hz, Δν = 59.2 Hz, 1H, $H_{5'}$), 2.65 (dt, J = 12.8, 6.2 Hz, 1H, $H_{2'}$), 2.42 (ddd, J = 13.1, 6.3, 4.0 Hz, 1H, $H_{2'}$), 0.899 (s, 9H, $H_{3'c}$ or $H_{5'c}$), 0.896 (s, 9H, $H_{3'c}$ or $H_{5'c}$), 0.09 (s, 3H, $H_{3'a}$ or $H_{5'a}$), 0.083 (s, 3H, $H_{3'a}$ or $H_{5'a}$), 0.077 (s, 6H, $H_{3'a}$ and/or $H_{5'a}$). ^{13}C NMR (151 MHz, $CDCl_3$): δ 161.3 (C_6), 152.2 (C_4),

[hhh] Detailed gradient: 10% B–A (99 mL), 10%→25% B–A (33 mL), 25% B–A (66 mL), 25%→60% B–A (49 mL), 60% B–A (99 mL), 60%→63% B–A (14 mL), 63% B–A (70 mL), 63%→80% B–A (84 mL), 80% B–A →100% B (33 mL), and 100% B (70 mL), in that order; A = hexanes and B = ethyl acetate.

151.9 (C_2), 150.9 (C_{2a}), 147.5 (C_{2h}), 140.1 (C_8), 134.0 (C_{2e}), 132.5 (C_{2c}), 129.1 (C_{2d}), 129.0 (C_{2f}), 125.3 (C_{2g}), 117.8 (C_5), 88.2 ($C_{4'}$), 84.7 ($C_{1'}$), 71.9 ($C_{3'}$), 64.0 (C_{2b}), 62.9 ($C_{5'}$), 54.7 (C_6OCH_3), 41.3 ($C_{2'}$), 26.1 ($C_{3'c}$ or $C_{5'c}$), 25.9 ($C_{3'c}$ or $C_{5'c}$), 18.6 ($C_{3'b}$ or $C_{5'b}$), 18.1 ($C_{3'b}$ or $C_{5'b}$), –4.5 ($C_{3'a}$ or $C_{5'a}$), –4.7 ($C_{3'a}$ or $C_{5'a}$), –5.2 ($C_{3'a}$ or $C_{5'a}$), –5.3 ($C_{3'a}$ or $C_{5'a}$). IR (ATR-FTIR), cm^{-1}: 2953 (m), 2929 (m), 2858 (m), 1763 (m), 1610 (s), 1589 (m), 1528 (s), 1468 (s), 1248 (m), 1194 (s), 1097 (s), 1076 (s), 1071 (s), 837 (s), 778 (m). HRMS-ESI (m/z): $[M + Na]^+$ calc'd for $[C_{31}H_{48}N_6NaO_8Si_2]^+$ 711.2964; found 711.2958 (error = 0.8 ppm). $[\alpha]_D^{20}$ = +16.86, (c = 0.327, $CHCl_3$).

*Synthesis of O^6-methyl diol **23c**.*

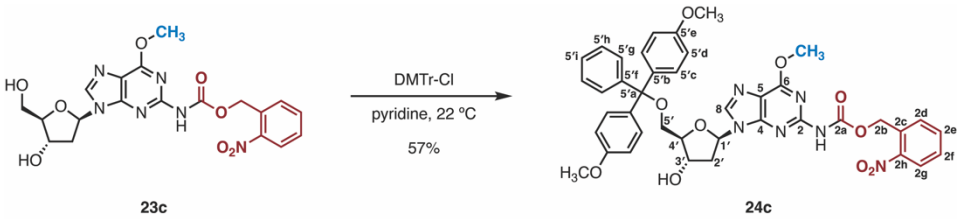


Hydrogen fluoride–pyridine complex (~70% HF, 4.00 mL, 31.0 mmol, 20.0 equiv) was added dropwise via syringe pump over 30 min to a 50-mL Falcon tube[iii] containing a solution of the N^2-NBOC protected O^6-methyl ether **22c** (1.07 g, 1.55 mmol, 1 equiv) in pyridine (26 mL) at 0 ºC. The cooling bath was then removed, and the reaction mixture was allowed to warm to 22 ºC. The reaction mixture was stirred for 17.5 h at 22 °C. The product mixture was slowly added to a stirred mixture of sodium bicarbonate (6.68 g, 67.4 mmol, 43.4 equiv), methanol (70 mL), and water (35 mL) at 0 ºC. The transfer was quantitated using methanol (25 mL). The cooling bath was removed, and the product mixture was allowed to warm to 22 °C. The warmed product mixture was stirred for 45 min at 22 °C, open to air. The mixture was filtered, and the filtrate was concentrated. The residue obtained was dried by sequential azeotropic distillation with methanol (2 × 40 mL) and toluene (2 × 40 mL). The azeotroped product mixture was purified by automated flash-column chromatography (solid loading on Celite, eluting with 1%→20% methanol–dichloromethane[iii]) to provide the O^6-methyl diol **23c** as a white powder (700 mg, 98%).

R_f = 0.58 (20% methanol–dichloromethane; UV). ^{1}H NMR (600 MHz, DMSO-d_6): δ 10.59 (s, 1H, NH), 8.43 (s, 1H, H_8), 8.17 (dd, J = 8.2, 1.2 Hz, 1H, H_{2g}), 7.90 (dd, J = 7.9, 1.4 Hz, 1H, H_{2d}), 7.85 (td, J = 7.6, 1.3 Hz, 1H, H_{2e}), 7.66 – 7.61 (m, 1H, H_{2f}), 6.33 (dd, J = 7.5, 6.2 Hz, 1H, $H_{1'}$), 5.54 (s, 2H, H_{2b}), 5.30 (d, J = 3.9 Hz, 1H, $C_{3'}OH$), 4.88 (t, J = 5.5 Hz, 1H, $C_{5'}OH$), 4.42 (dq, J = 6.4, 3.3 Hz, 1H, $H_{3'}$), 4.06 (s, 3H, C_6OCH_3), 3.85 (td, J = 4.8, 2.8 Hz, 1H, $H_{4'}$), 3.59 (dt, J = 10.8, 5.3 Hz, 1H, $H_{5'}$), 3.51 (dt, J = 11.9, 5.1 Hz, 1H, $H_{5'}$), 2.73 (ddd, J = 13.3, 7.6, 5.8 Hz, 1H, $H_{2'}$), 2.27 (ddd, J = 13.2, 6.2, 3.2 Hz, 1H, $H_{2'}$). ^{13}C NMR (151 MHz, DMSO-d_6): δ 160.4 (C_6), 152.6 (C_4), 151.9 (C_2), 151.5 (C_{2a}), 146.9 (C_{2h}), 141.1 (C_8), 134.3 (C_{2e}), 132.5 (C_{2c}), 129.0 (C_{2f}), 128.9 (C_{2d}), 124.9 (C_{2g}), 117.4 (C_5), 87.9 ($C_{4'}$), 83.2 ($C_{1'}$), 70.7 ($C_{3'}$), 62.7 (C_{2b}), 61.7 ($C_{5'}$), 53.9 (C_6OCH_3), 39.2 ($C_{2'}$). IR (ATR-FTIR), cm^{-1}: 1735 (m), 1617 (m), 1578 (m), 1521 (m), 1505 (m), 1470 (m), 1396 (m), 1382 (m), 1338 (m), 1310 (m), 1225 (s), 1206 (s), 1096 (s), 1067 (m), 1051 (m), 1031 (m), 989 (m), 955 (m), 930 (m), 858 (m), 792 (s), 760 (m), 726 (s). HRMS-ESI (m/z): $[M + Na]^+$ calc'd for $[C_{19}H_{20}N_6NaO_8]^+$ 483.1235; found 483.1220 (error = 3.1 ppm). $[a]_D^{20}$ = +16.62 (c = 0.225, CH_3OH).

[iii] The Falcon tube was sealed using an inverted 24/40 septum and secured with electrical tape.
[jjj] Detailed gradient: 5% B–A (165 mL), 5%→15% B–A (165 mL), 15% B–A (396 mL), 15%→18% B–A (48 mL), 18% B–A (82 mL), 18%→36% B–A (329 mL), 36% B–A (72 mL), 36%→37% B–A (20 mL), 37% B–A (52 mL), 50% B–A (336 mL), 50% B–A→100% B (66 mL), and 100% B (174 mL), in that order; A = dichloromethane and B = 20% methanol–dichloromethane.

*Synthesis of O⁶-methyl trityl ether **24c**.*



The O⁶-methyl diol **23c** (700 mg, 1.52 mmol, 1 equiv) was dried by sequential azeotropic distillation with pyridine (4 × 5.0 mL) and benzene (2 × 10 mL) and then under vacuum overnight. 4,4′-Dimethoxytrityl chloride (779 mg, 2.29 mmol, 1.51 equiv) and pyridine (5.7 mL) were then added in sequence. The reaction mixture was stirred for 1 h at 22 ºC with protection from light. Absolute ethanol (2.5 mL) was then added, and the resulting mixture was stirred for 1.5 h at 22 ºC. The product mixture was concentrated, and the residue obtained was co-evaporated with dichloromethane (2 × 10 mL). The resulting residue was purified by automated flash-column chromatography (eluting with 10% B–A→100% B→98% C–B; A = 100:1 v/v hexanes–triethylamine , B = 100:1 v/v ethyl acetate–triethylamine , and C = methanol[kkk]). Fractions containing the O⁶-methyl trityl ether **24c** were combined and concentrated. Residual triethylamine[lll] was removed by sequential co-evaporation with hexanes–ethyl acetate (1:1 v/v, 5 × 100 mL) and dichloromethane (4 × 50 mL) to provide the O⁶-methyl trityl ether **24c** as a white solid (665 mg, 57%).

R_f = 0.51 (100% ethyl acetate; UV, stains orange with CAM[mmm]). ^{1}H NMR (600 MHz, DMSO-d_6): δ 10.54 (s, 1H, NH), 8.32 (s, 1H, H_8), 8.19 (dd, J = 8.3, 1.3 Hz, 1H, H_{2g}), 7.90 (dd, J = 7.8, 1.4 Hz, 1H, H_{2d}), 7.84 (td, J = 7.6, 1.3 Hz, 1H, H_{2e}), 7.64 (ddd, J = 8.6, 7.4, 1.5, 1H, H_{2f}), 7.30 – 7.25 (m, 2H, $H_{5'g}$), 7.18 – 7.11 (m, 7H, $H_{5'c}$, $H_{5'h}$, and $H_{5'i}$), 6.77 – 6.72 (m, 2H, $H_{5'd}$), 6.70 – 6.66 (m, 2H, $H_{5'd}$), 6.37 (dd, J = 7.0, 5.7 Hz, 1H, $H_{1'}$), 5.52 (ABq, J = 15.2 Hz, Δν = 14.2 Hz, 2H, H_{2b}), 5.28 (d, J = 4.6 Hz, 1H, $C_{3'}OH$), 4.55 (dq, J = 6.7, 4.7 Hz, 1H, $H_{3'}$), 4.06 (s, 3H, C_6OCH_3), 3.96 (dt, J = 7.3, 3.7 Hz, 1H, $H_{4'}$), 3.692 (s, 3H, $C_{5'}OCH_3$), 3.687 (s, 3H, $C_{5'}OCH_3$), 3.33 ($H_{5'}$)[nnn], 3.08 (dd, J = 10.4, 3.1 Hz, 1H, $H_{5'}$), 2.90 (dt, J = 13.4, 6.2 Hz, 1H, $H_{2'}$), 2.32 (ddd, J = 13.4, 7.1, 4.9 Hz, 1H, $H_{2'}$). ^{13}C NMR (151 MHz, DMSO-d_6): δ 160.4 (C_6), 157.9 ($C_{5'e}$), 157.8 ($C_{5'e}$), 152.3 (C_4), 151.8 (C_2), 151.2 (C_{2a}), 146.8 (C_{2h}), 145.0 ($C_{5'f}$), 141.5 (C_8), 135.59 ($C_{5'b}$), 135.57 ($C_{5'b}$), 134.4 (C_{2e}), 132.6 (C_{2c}), 129.7 ($C_{5'c}$, $C_{5'h}$, or $C_{5'i}$), 129.5 ($C_{5'c}$, $C_{5'h}$, or $C_{5'i}$), 129.0 (C_{2f}), 128.7 (C_{2d}), 127.63 ($C_{5'c}$, $C_{5'g}$, $C_{5'h}$, or $C_{5'i}$), 127.56 ($C_{5'c}$, $C_{5'g}$, $C_{5'h}$, or $C_{5'i}$), 126.5 ($C_{5'h}$ or $C_{5'i}$), 124.9 (C_{2g}), 117.7 (C_5), 112.9 ($C_{5'd}$), 112.8 ($C_{5'd}$), 86.5 ($C_{4'}$), 85.3 ($C_{5'a}$), 83.7 ($C_{1'}$), 70.6 ($C_{3'}$), 64.5 ($C_{5'}$), 62.6 (C_{2b}), 54.92 ($C_{5'e}OCH_3$), 54.88 ($C_{5'e}OCH_3$), 53.9 (C_6OCH_3), 38.6 ($C_{2'}$). IR (ATR-FTIR), cm^{-1}: 3259 (br w), 2952 (w), 2948 (w), 2933 (w), 1757 (w), 1607 (m), 1588 (m), 1524 (m), 1507 (s), 1465 (m), 1446 (m), 1393 (m), 1375 (m), 1342 (m), 1303 (m), 1244 (s), 1174 (s), 1064 (s), 1030

[kkk] Detailed gradient: 10% B–A (528 mL), 10%→36% B–A (330 mL), 36% B–A (277 mL), 36%→50% B–A (184 mL), 50% B–A (620 mL), 50%→72% B–A (290 mL), 72% B–A (330 mL), 72%→80% B–A (102 mL), 80%→99% B–A (132 mL), 100% B (2571 mL), 0%→98% C–B (198 mL), and 98% C–B (158 mL), in that order; A = 100:1 v/v hexanes–triethylamine, B = 100:1 v/v ethyl acetate–triethylamine, and C = methanol (no triethylamine).

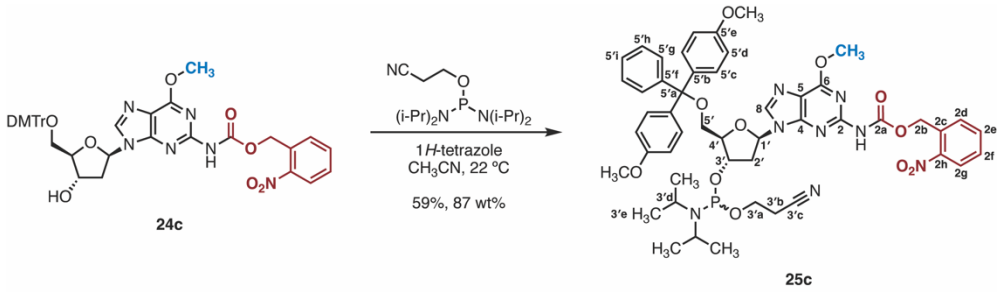
[lll] Removing triethylamine is critical as residual triethylamine will neutralize the 1*H*-tetrazole used in the conversion of the O⁶-methyl trityl ether **24c** to the O⁶-methyl phosphoramidite **25c**.

[mmm] Orange color likely results from deprotection of the dimethoxytrityl ether to give the dimethoxytrityl cation.

[nnn] Due to overlap with the water peak in the 1-D ^{1}H spectrum, this ^{1}H shift was estimated from an ^{1}H-^{13}C HSQC{^{1}H}{^{13}C} spectrum (DMSO-d_6, 600/151 MHz) and neither multiplicity nor integration are given.

(s), 827 (m) 791 (m), 757 (m), 727 (m), 701 (m), 583 (m). HRMS-ESI (*m*/*z*): $[M + H]^+$ calc'd for $[C_{40}H_{39}N_6O_{10}]^+$ 763.2722; found 763.2701 (error = 2.8 ppm). $[a]_D^{20} = +29.93$ ($c = 0.143$, CH_3OH).

Synthesis of O^6-methyl phosphoramidite ***25c****.*



The O^6-methyl trityl ether **24c** (258 mg, 339 µmol, 1 equiv) was dried by azeotropic distillation with acetonitrile (6 × 4.0 mL, 35 ºC) then under vacuum overnight. Acetonitrile (2.1 mL) and 2-cyanoethyl *N*,*N*,*N*′,*N*′-tetraisopropylphosphorodiamidite (35 wt% by qNMR, 450 µL, 493 µmol, 1.46 equiv) were then added in sequence. The reaction mixture was stirred for 4 min at 22 ºC. 1*H*-Tetrazole (3 wt% in acetonitrile, 1.20 mL, 406 µmol, 1.20 equiv) was then added at 22 ºC. The reaction mixture was stirred for 3 h at 22 ºC with protection from light. Methanol (1.0 mL) was then added, and the resulting mixture was stirred for 10 min at 22 °C. The solution was concentrated, and the residue obtained was purified by automated flash-column chromatography[ooo] (eluting with 15% B–A→100% B; A = 100:1 v/v hexanes–triethylamine and B = 100:1 v/v ethyl acetate–triethylamine [ppp]). Fractions containing both diastereomers of the O^6-methyl phosphoramidite **25c** were combined and concentrated. The concentrated product was co-evaporated sequentially with hexanes–ethyl acetate (1:1 v/v, 5 × 100 mL) and dichloromethane (5 × 100 mL) to remove triethylamine.[qqq] The resulting residue was dried by azeotropic distillation with acetonitrile (4 × 3.0 mL) and then co-evaporated with dichloromethane (5 × 3.0 mL) to provide the O^6-methyl phosphoramidite **25c** as an off-white solid (221 mg @ 87 wt% by qNMR, 59% yield).

R_f = 0.59 & 0.67 (90% ethyl acetate–hexanes; UV, stains orange with CAM[rrr]). ^{1}H NMR (600 MHz, CD_3CN): δ 8.46 – 8.41 (m, 1H, NH), 8.18 – 8.12 (m, 1H, H_{2g}), 7.99 (s, 1H, H_8), 7.82 – 7.77 (m, 1H, H_{2d}), 7.77 – 7.73 (m, 1H, H_{2e}), 7.60 – 7.55 (m, 1H, H_{2f}), 7.35 – 7.29 (m, 2H, $H_{5'g}$), 7.23 – 7.11 (m, 7H, $H_{5'c}$, $H_{5'h}$, and $H_{5'i}$), 6.74 – 6.69 (m, 2H, $H_{5'd}$), 6.69 – 6.62 (m, 2H, $H_{5'd}$), 6.35 – 6.31 (m, 1H, $H_{1'}$), 5.55 (ABq, J = 15.2 Hz, Δν, = 12.7 Hz, 2H, H_{2b}), 4.97 – 4.87 (m, 1H, $H_{3'}$), 4.23 – 4.14 (m, 1H, $H_{4'}$), 4.11 – 4.08 (m, 3H, C_6OCH_3), 3.80 – 3.62 (m, 8H, $C_{5'}OCH_3$ and $H_{3'a}$), 3.62 – 3.41 (m, $H_{3'd}$ and $H_{5'}$)[sss], 3.27 – 3.19 (m, 1H, $H_{5'}$), 3.13 – 2.98 (m, 1H, $H_{2'}$), 2.62 – 2.59 (m, 1H,

[ooo] The silica gel used to run the column was dried in an oven for >48 h then cooled in a desiccator immediately prior to use. The solvent mixtures used to run the column were prepared immediately prior to use in the following manner: for 100:1 v/v hexanes–triethylamine, a new bottle (4 L) of hexanes was sparged with dinitrogen for ≥ 1 h. Then an appropriate amount was measured out and combined with triethylamine (distilled from CaH_2 under dinitrogen immediately before use) to give the 100:1 v/v hexanes–triethylamine used to run the column. The 100:1 v/v ethyl acetate–triethylamine mixture was prepared in parallel in the same manner.

[ppp] Detailed gradient: 15%→20% B–A (66 mL), 20% B–A (198 mL), 20%→60% B–A (165 mL), 60% B–A (593 mL), 60%→75% B–A (85 mL), 75% B–A (330 mL), 75%→85% B–A (66 mL), 85% B–A (297 mL), 85% B–A →100% B (99 mL), and 100% B (198 mL), in that order; where A = 100:1 hexanes–triethylamine (v/v) and B = 100:1 ethyl acetate–triethylamine (v/v).

[qqq] Removing triethylamine is critical as residual triethylamine will neutralize the tetrazole-based activator used in the solid phase oligonucleotide synthesis deoxyoligonucleotide **26e**.

[rrr] Orange color likely results from deprotection of the dimethoxytrityl ether to give the dimethoxytrityl cation.

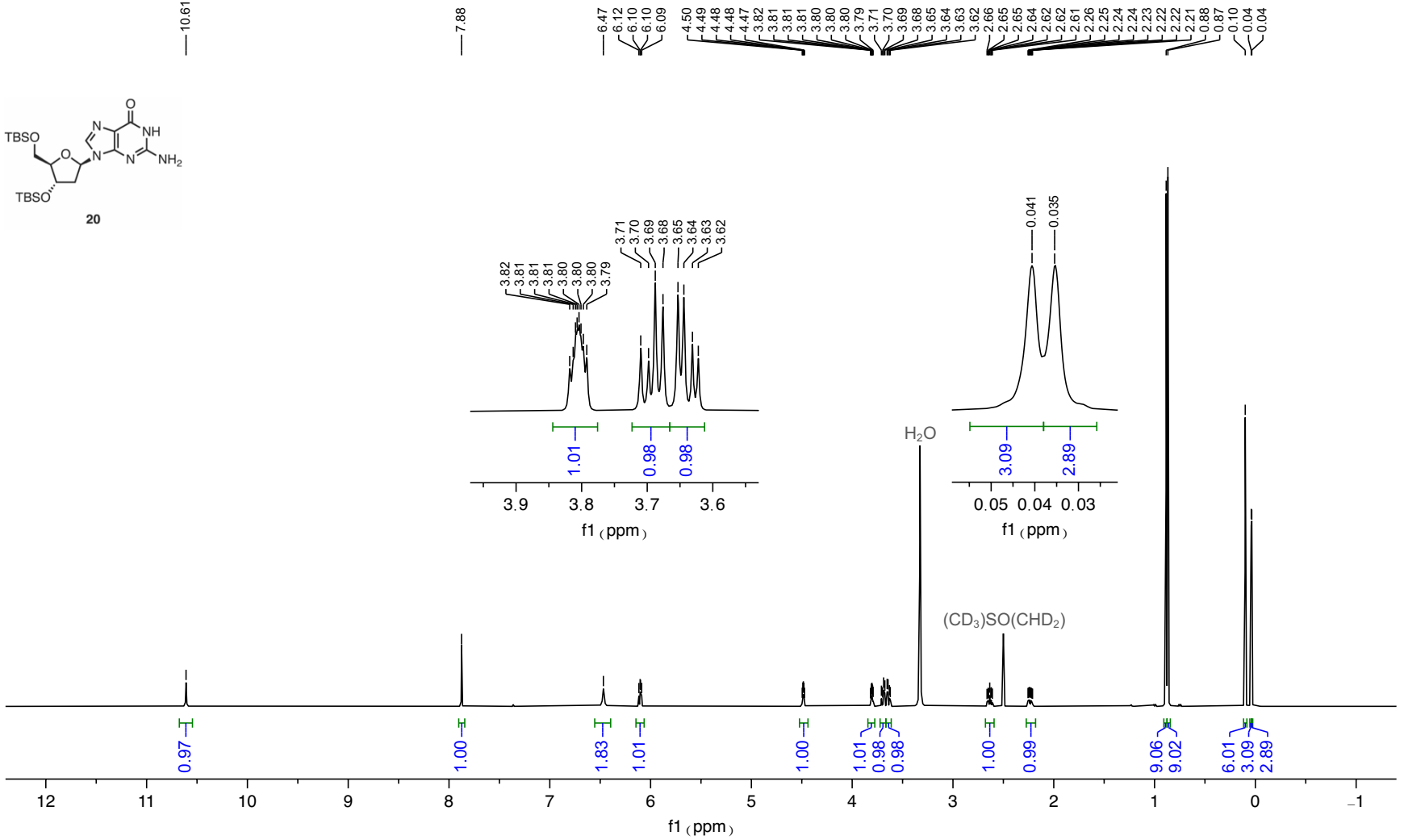
[sss] Due to overlap with impurities in the 1-D ^{1}H spectrum, it was not possible to obtain an accurate integration for this multiplet.

$H_{3'b}$), 2.59 – 2.51 (m, 1H, $H_{2'}$), 2.50 – 2.46 (m, 1H, $H_{3'b}$), 1.15 (d, J = 6.8 Hz, 3H, $H_{3'e}$), 1.13 (d, J = 6.8 Hz, 3H, $H_{3'e}$), 1.12 (d, J = 6.8, 3H, $H_{3'e}$), 1.03 (d, J = 6.7 Hz, 3H, $H_{3'e}$). $^{13}C\{^1H\}$ NMR (151 MHz, CD_3CN): δ 162.1 (C_6), 159.54 ($C_{5'e}$), 159.53 ($C_{5'e}$), 159.45 ($C_{5'e}$), 159.4 ($C_{5'e}$), 153.43 (C_4), 153.41 (C_4), 152.6 (C_2), 151.82 (C_{2a}), 151.75 (C_{2a}), 148.33 (C_{2h}), 148.28 (C_{2h}), 146.3 ($C_{5'f}$), 146.2 ($C_{5'f}$), 142.3 (C_8), 136.94 ($C_{5'b}$), 136.87 ($C_{5'b}$), 136.78 ($C_{5'b}$), 136.75 ($C_{5'b}$), 135.20 (C_{2e}), 135.18 (C_{2e}), 133.7 (C_{2c}), 133.6 (C_{2c}), 131.1 ($C_{5'c}$, $C_{5'h}$, or $C_{5'i}$), 131.0 ($C_{5'c}$, $C_{5'h}$, or $C_{5'i}$), 130.8 ($C_{5'c}$, $C_{5'h}$, or $C_{5'i}$), 129.91 (C_{2f}), 129.87 (C_{2f}), 129.7 (C_{2d}), 129.6 (C_{2d}), 128.94 ($C_{5'g}$), 128.88 ($C_{5'g}$), 128.60 ($C_{5'c}$, $C_{5'h}$, or $C_{5'i}$), 128.58 ($C_{5'c}$, $C_{5'h}$, or $C_{5'i}$), 127.7 ($C_{5'h}$ or $C_{5'i}$), 127.6 ($C_{5'h}$ or $C_{5'i}$), 126.0 (C_{2g}), 119.71 (C_5), 119.69 (C_5), 119.6 ($C_{3'c}$), 119.4 ($C_{3'c}$), 113.8 ($C_{5'd}$), 113.74 ($C_{5'd}$), 113.72 ($C_{5'd}$), 87.23 ($C_{4'}$), 87.20 ($C_{4'}$), 87.0 ($C_{4'}$), 86.9 ($C_{4'}$), 86.84 ($C_{5'a}$), 86.81 ($C_{5'a}$), 85.82 ($C_{1'}$), 85.79 ($C_{1'}$), 74.8 ($C_{3'}$), 74.7 ($C_{3'}$), 74.6 ($C_{3'}$), 74.5 ($C_{3'}$), 65.28 ($C_{5'}$), 65.26 ($C_{5'}$), 64.20 (C_{2b}), 64.15 (C_{2b}), 59.5 ($C_{3'a}$), 59.43 ($C_{3'a}$), 59.36 ($C_{3'a}$), 59.3 ($C_{3'a}$), 55.8 ($C_{5'e}OCH_3$), 54.8 (C_6OCH_3), 44.1 ($C_{3'd}$), 44.02 ($C_{3'd}$), 43.97 ($C_{3'd}$), 43.9 ($C_{3'd}$), 39.27 ($C_{2'}$), 39.25 ($C_{2'}$), 39.01 ($C_{2'}$), 38.98 ($C_{2'}$), 24.90 ($C_{3'e}$), 24.88 ($C_{3'e}$), 24.85 ($C_{3'e}$), 24.83 ($C_{3'e}$), 24.80 ($C_{3'e}$), 24.79 ($C_{3'e}$), 21.03 ($C_{3'b}$), 20.98 ($C_{3'b}$), 20.9 ($C_{3'b}$).[ttt] $^{31}P\{^1H\}$ NMR (202 MHz, CD_3CN): δ 148.3 (s, 0.5P), 148.1 (s, 0.5P). HRMS-ESI (m/z): $[M+Na]^+$ calc'd for $[C_{49}H_{55}N_8NaO_{11}P]^+$ 985.3620; found 985.3627 (error = 0.7 ppm).

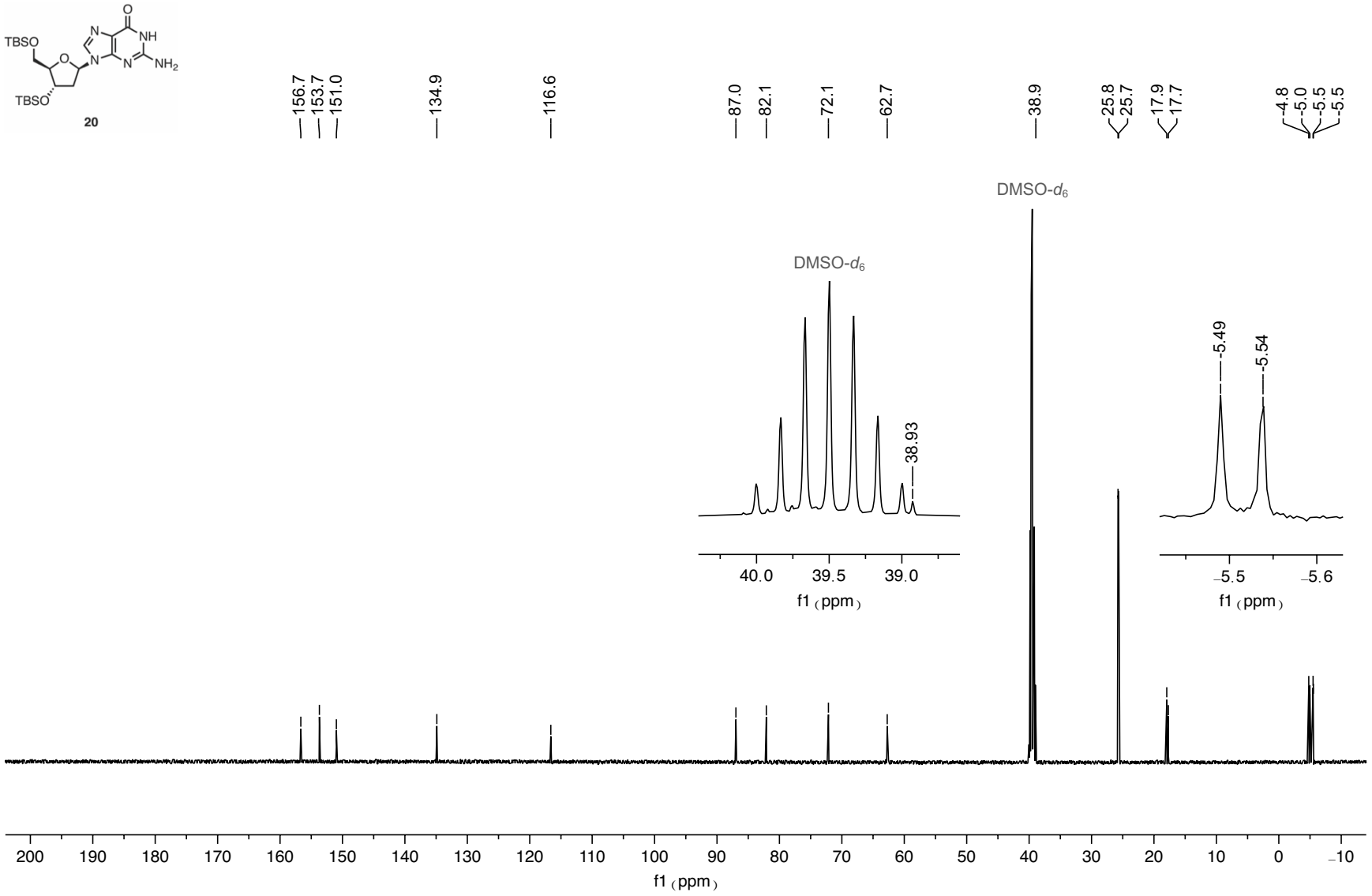
[ttt] Due to the similarity between the difference in shifts for each diastereomer and the magnitude of the J_{CP} coupling constants, J_{CP} values could not be confidently determined. As such, each peak is individually reported.

Catalog of Nuclear Magnetic Resonance Spectra.

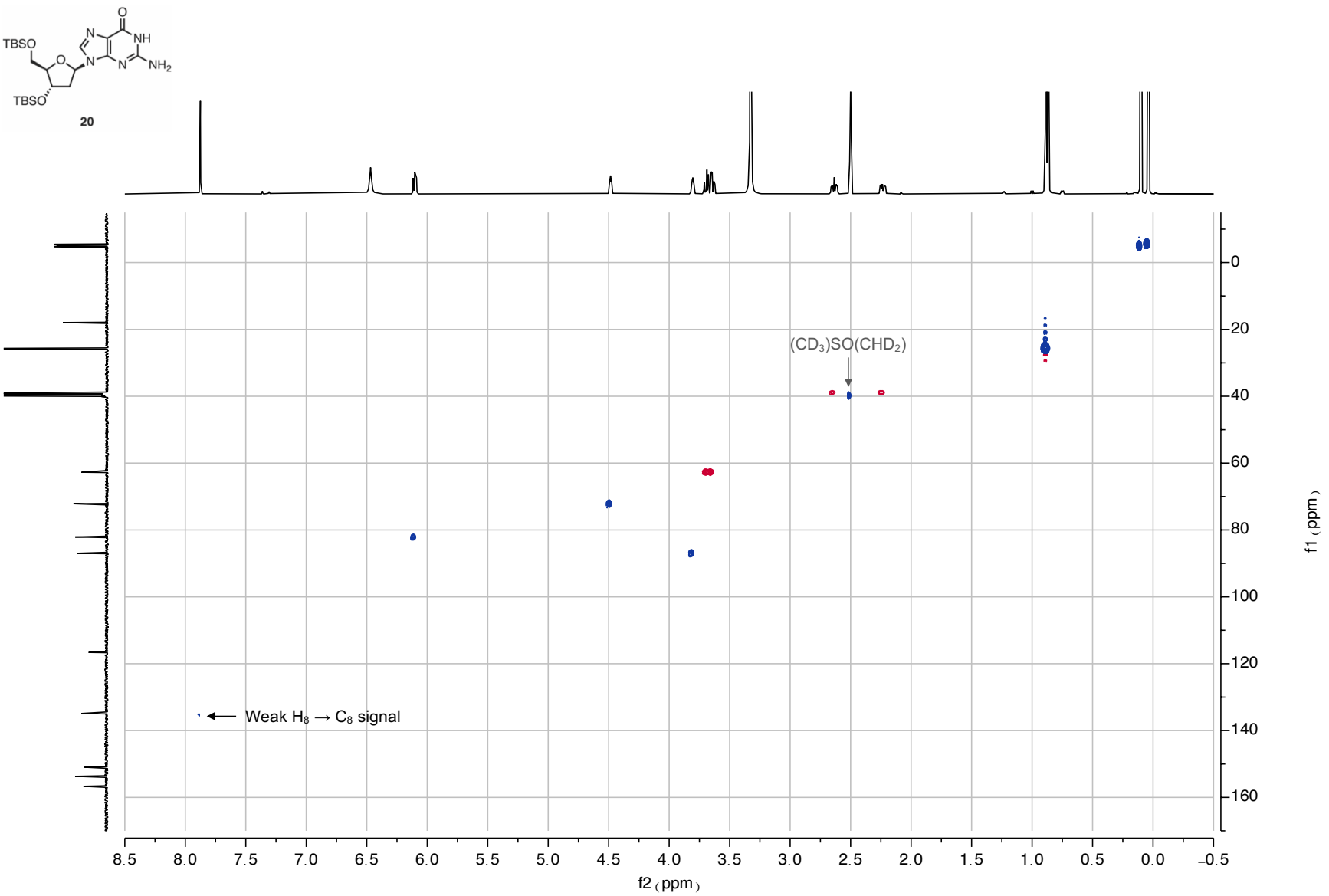
^{1}H NMR, 500 MHz, DMSO-d_6



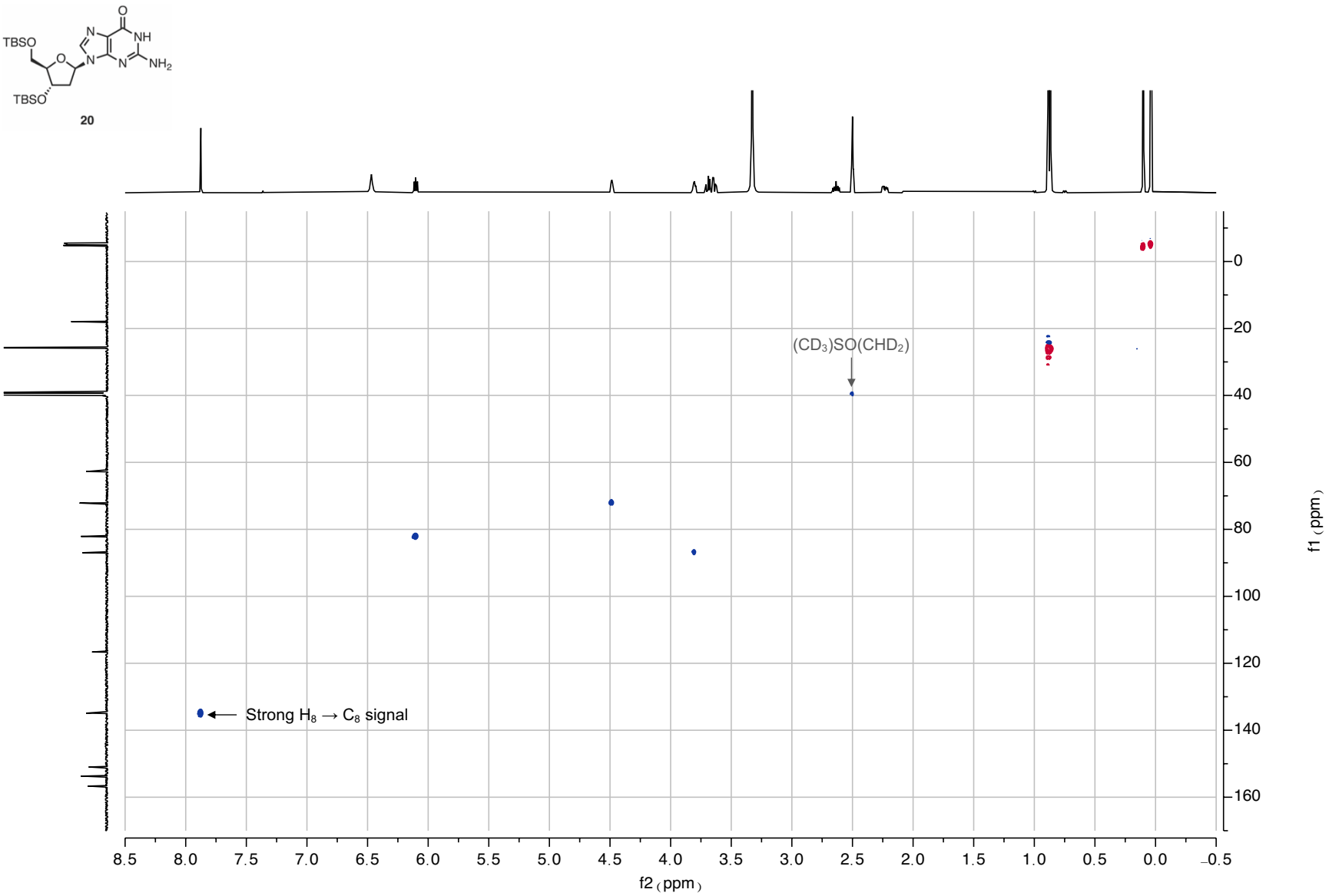
$^{13}C\{^{1}H\}$ NMR, 126 MHz, DMSO-d_6



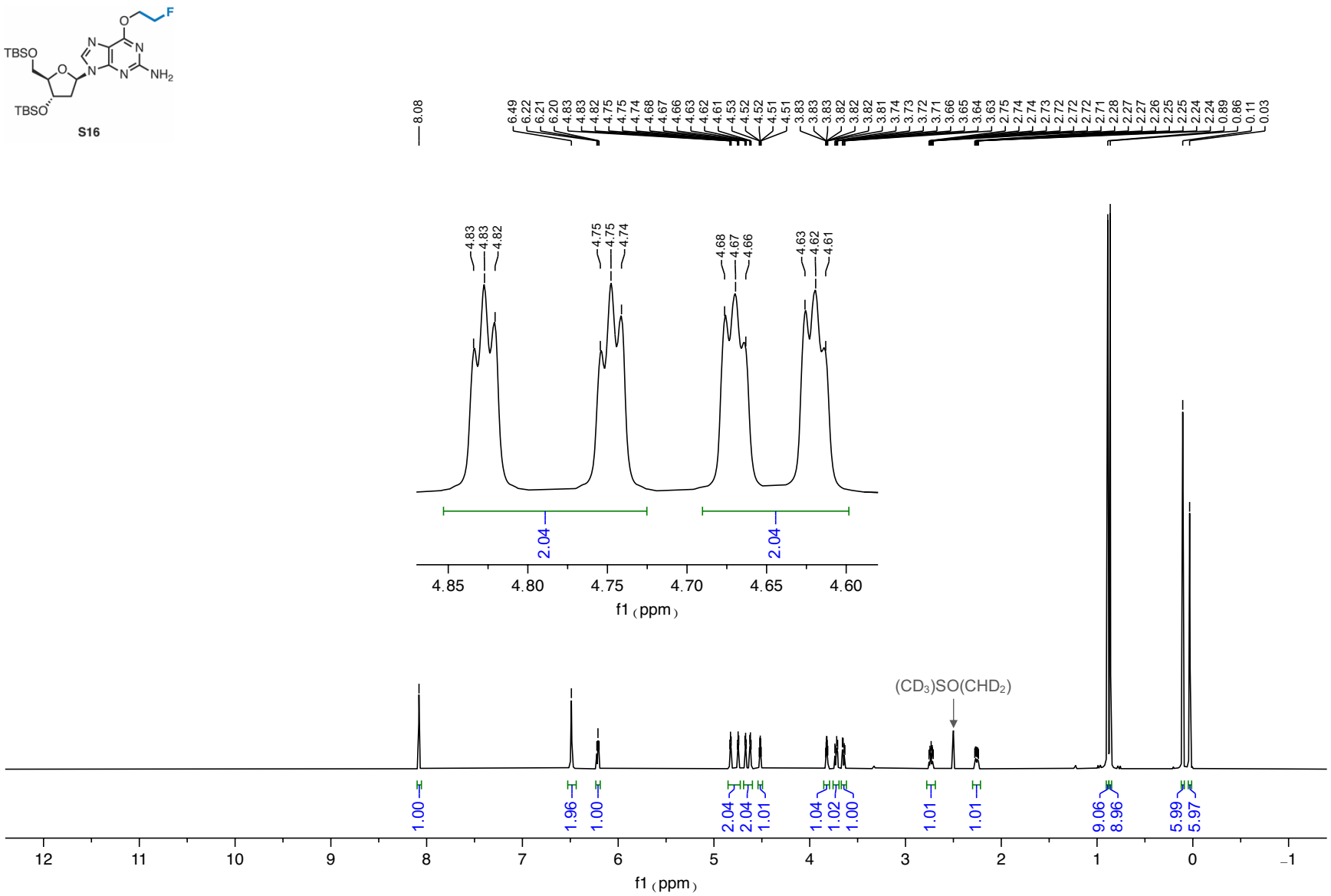
^{1}H–^{13}C HSQC{^{1}H}{^{13}C} NMR, 500/126 MHz, DMSO-d_6, $^{1}J_{CH}$ = 146 Hz, Blue = CH/CH_3, Red = CH_2



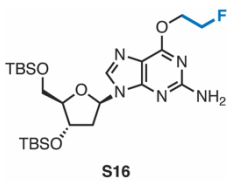
^{1}H–^{13}C HSQC{^{1}H}{^{13}C} NMR, 500/126 MHz, DMSO-d_6, $^{1}J_{CH}$ = 210 Hz, Blue = CH/CH_3, Red = CH_2

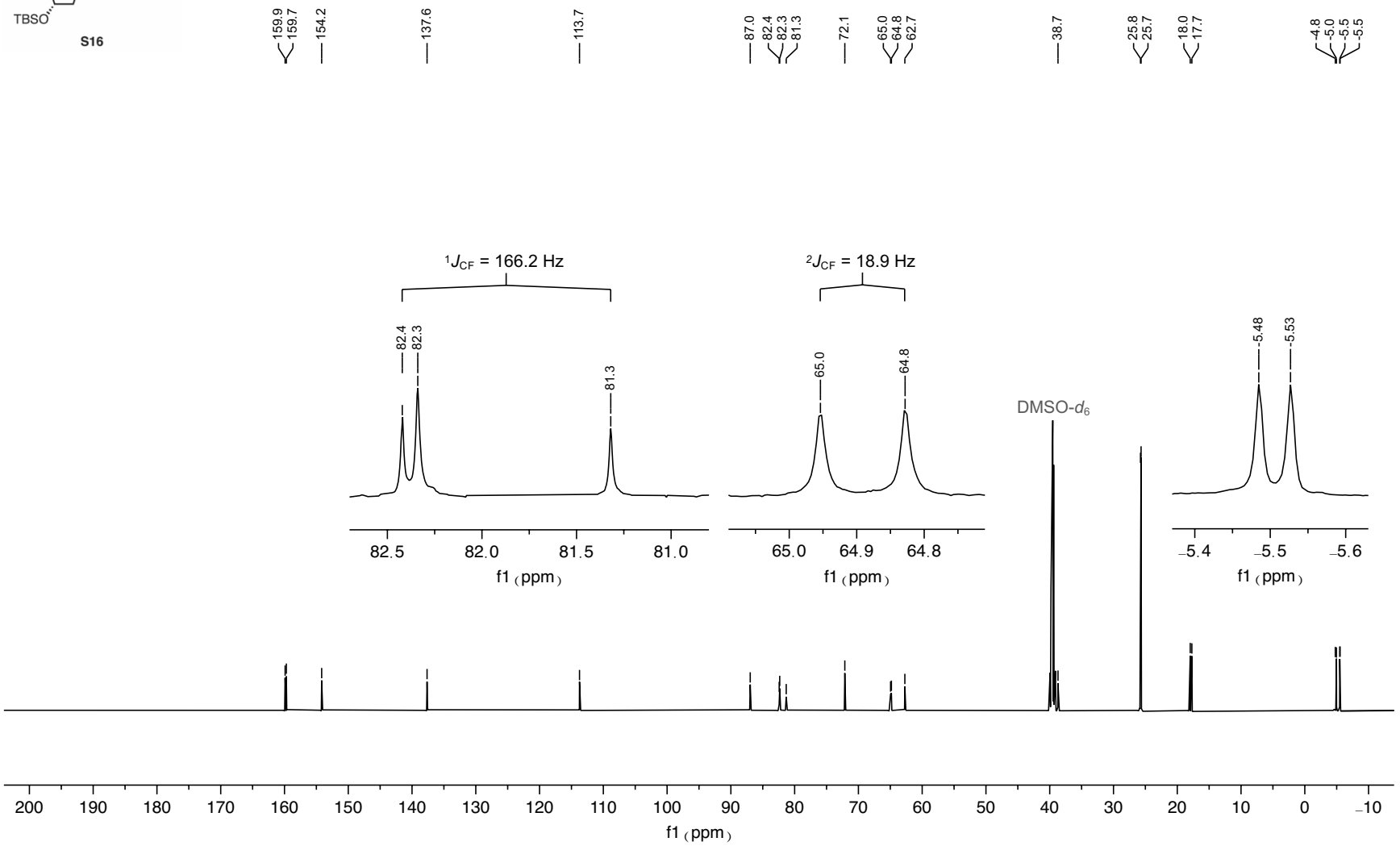


^{1}H NMR, 600 MHz, DMSO-d_6

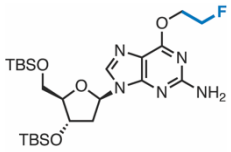


$^{13}C\{^{1}H\}$ NMR, 151 MHz, DMSO-d_6

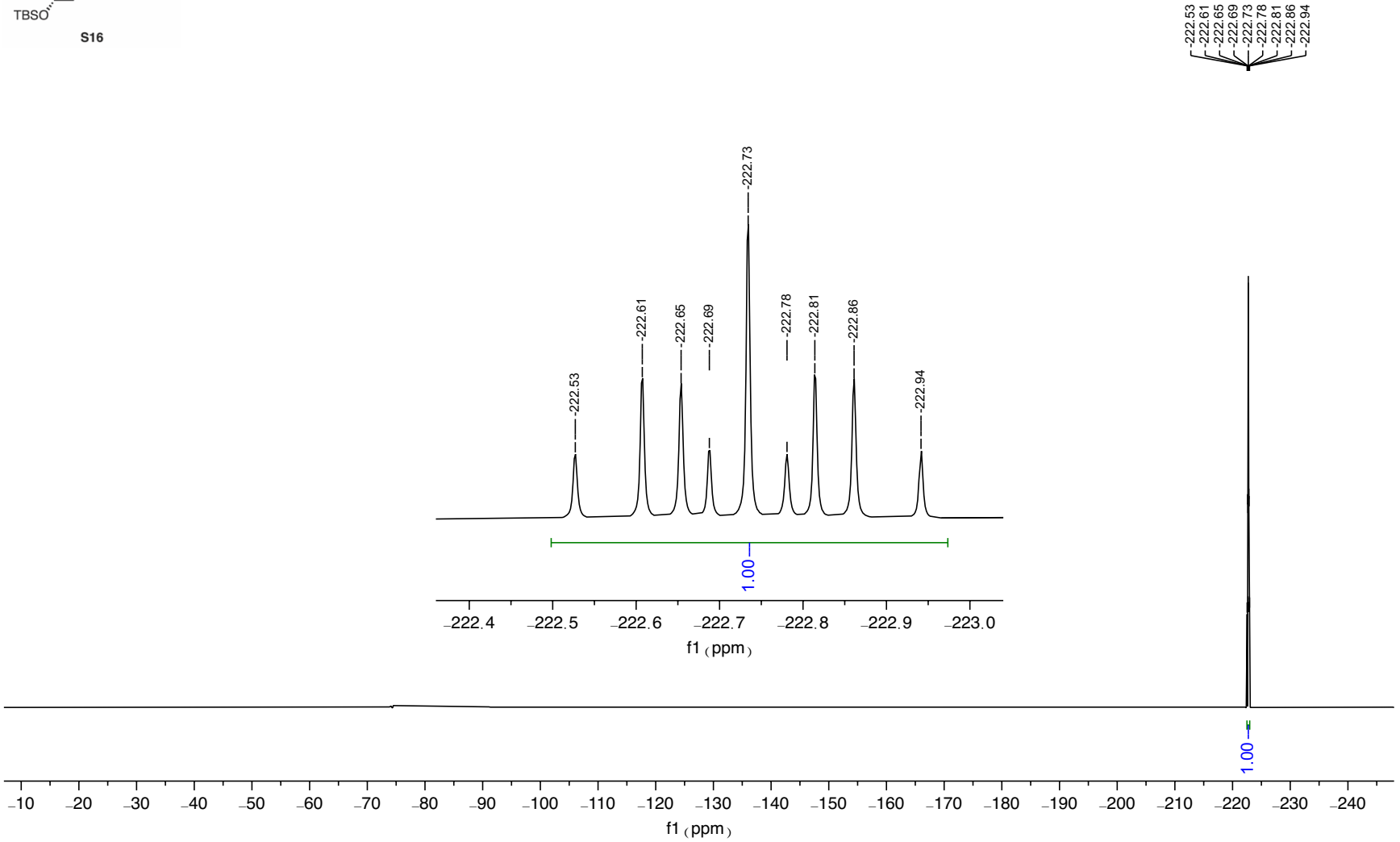




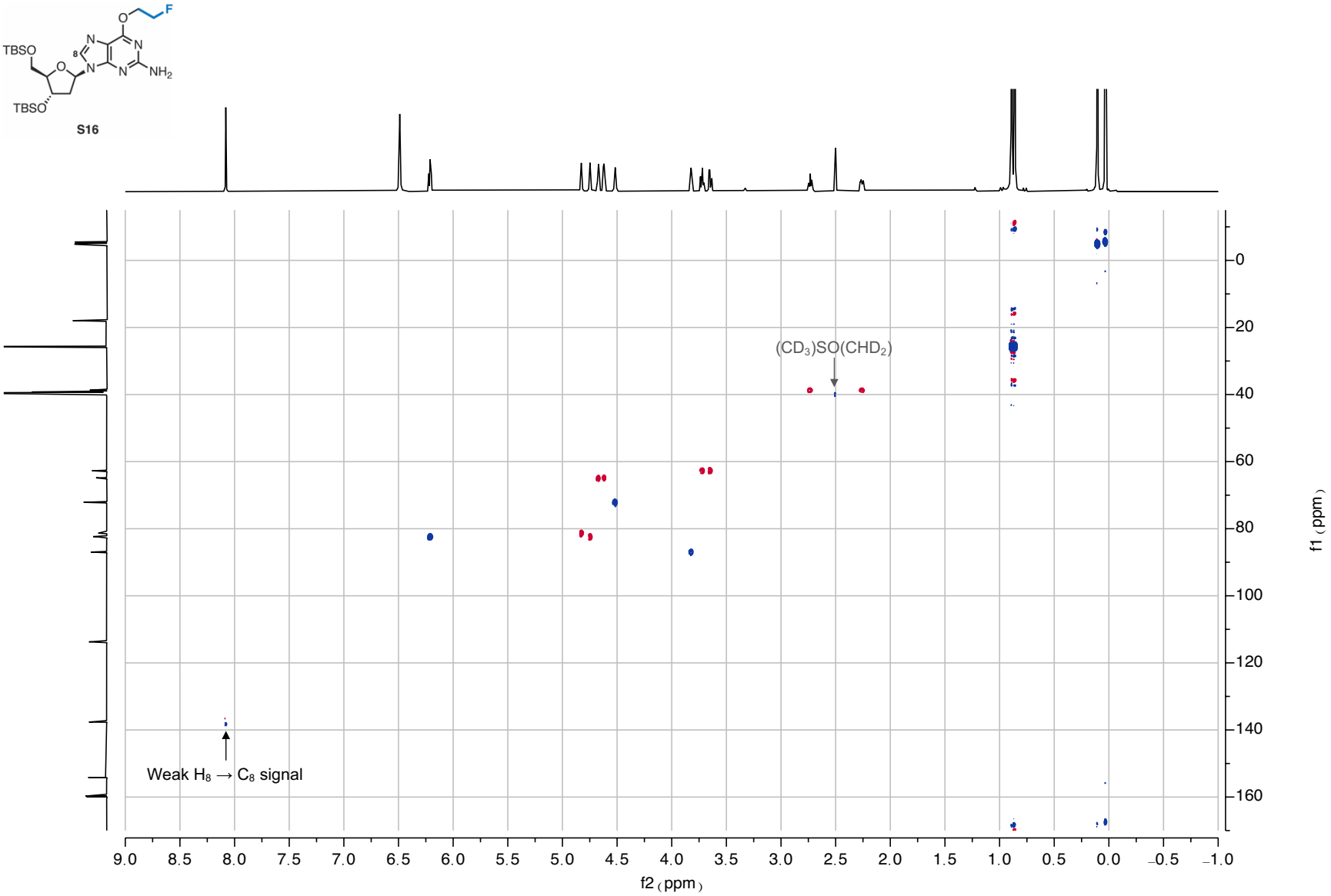
^{19}F NMR, 376 MHz, DMSO-d_6



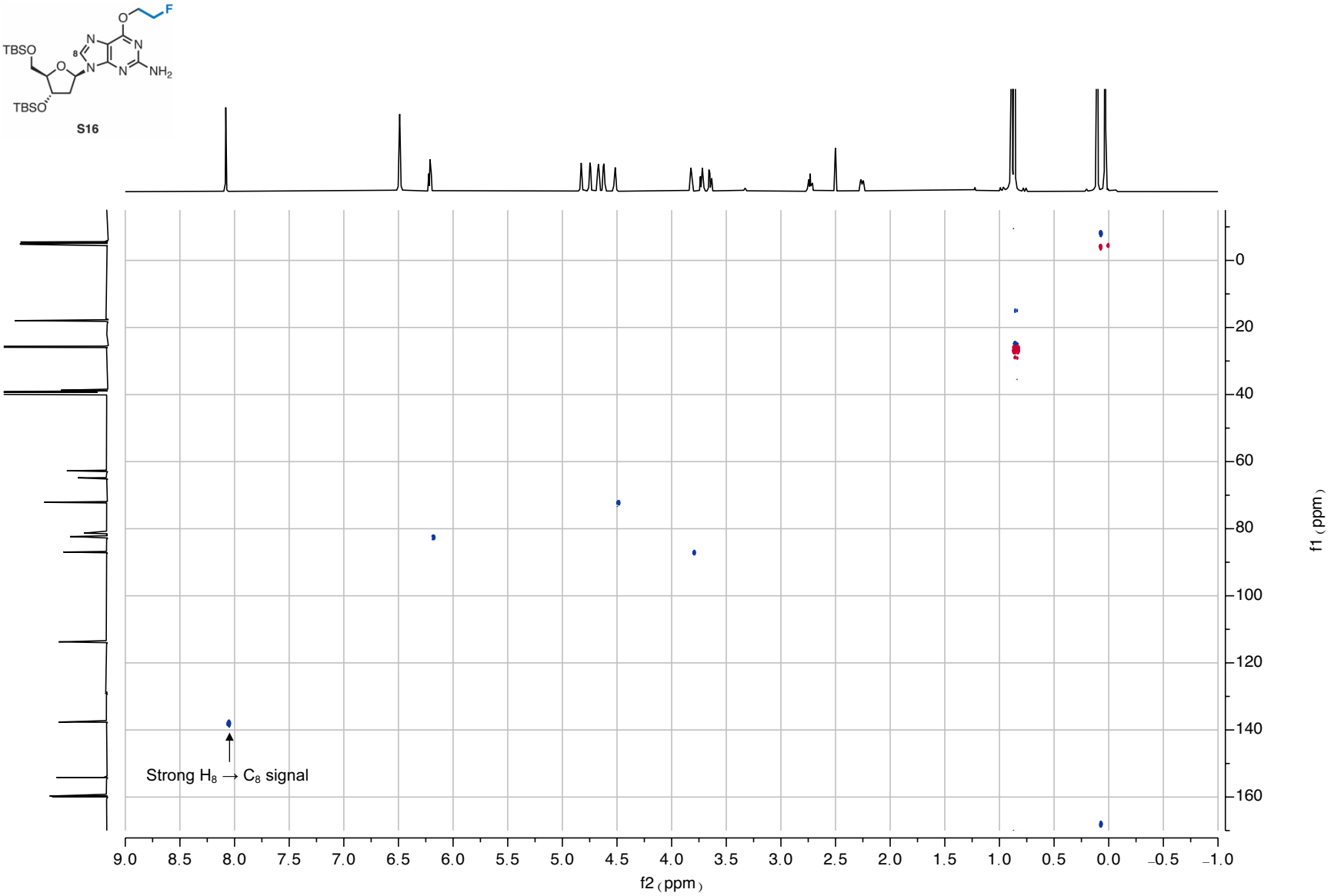
S16



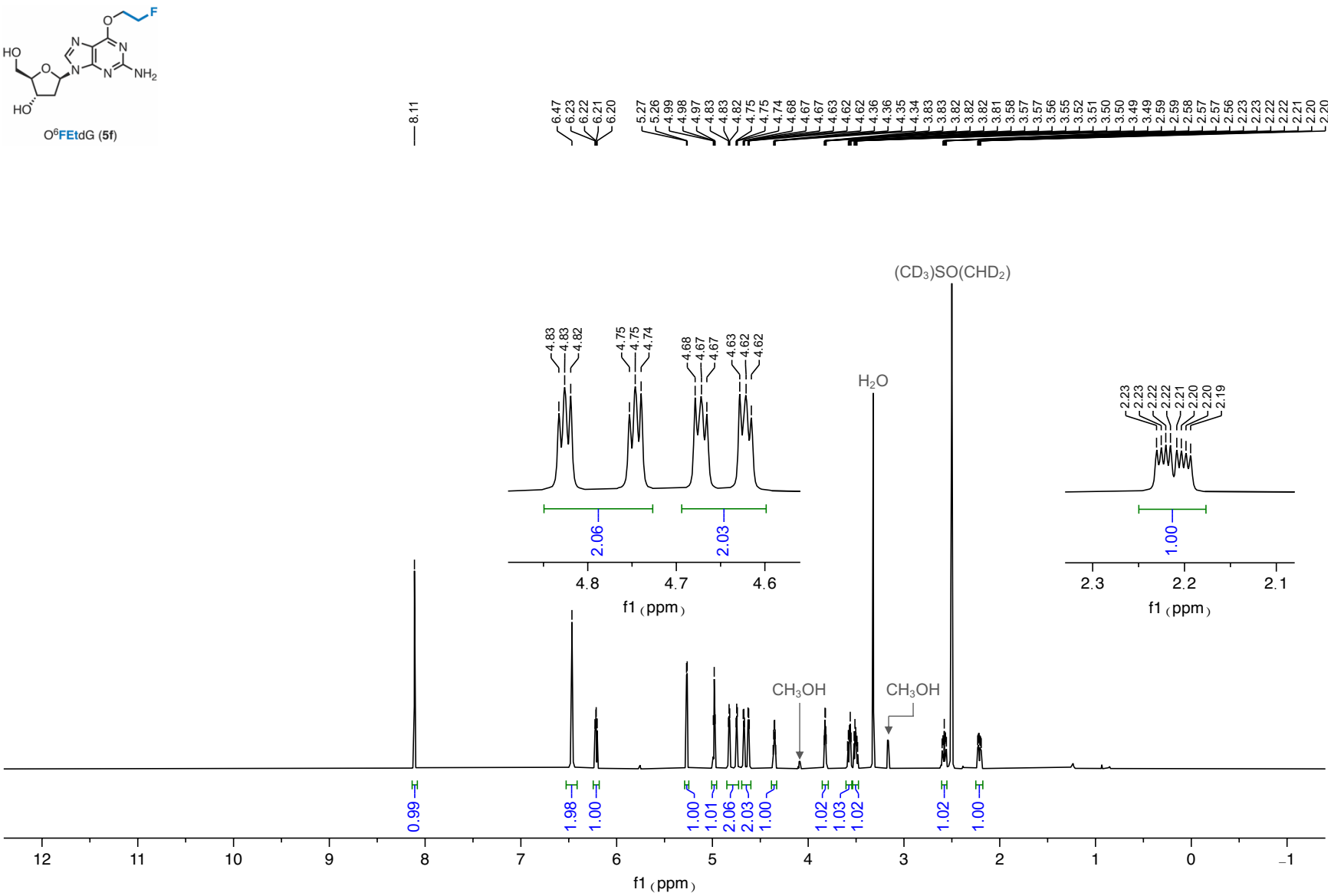
^{1}H–^{13}C HSQC{^{1}H}{^{13}C} NMR, 600/151 MHz, DMSO-d_6, $^{1}J_{CH}$ = 146 Hz, Blue = CH/CH_3, Red = CH_2



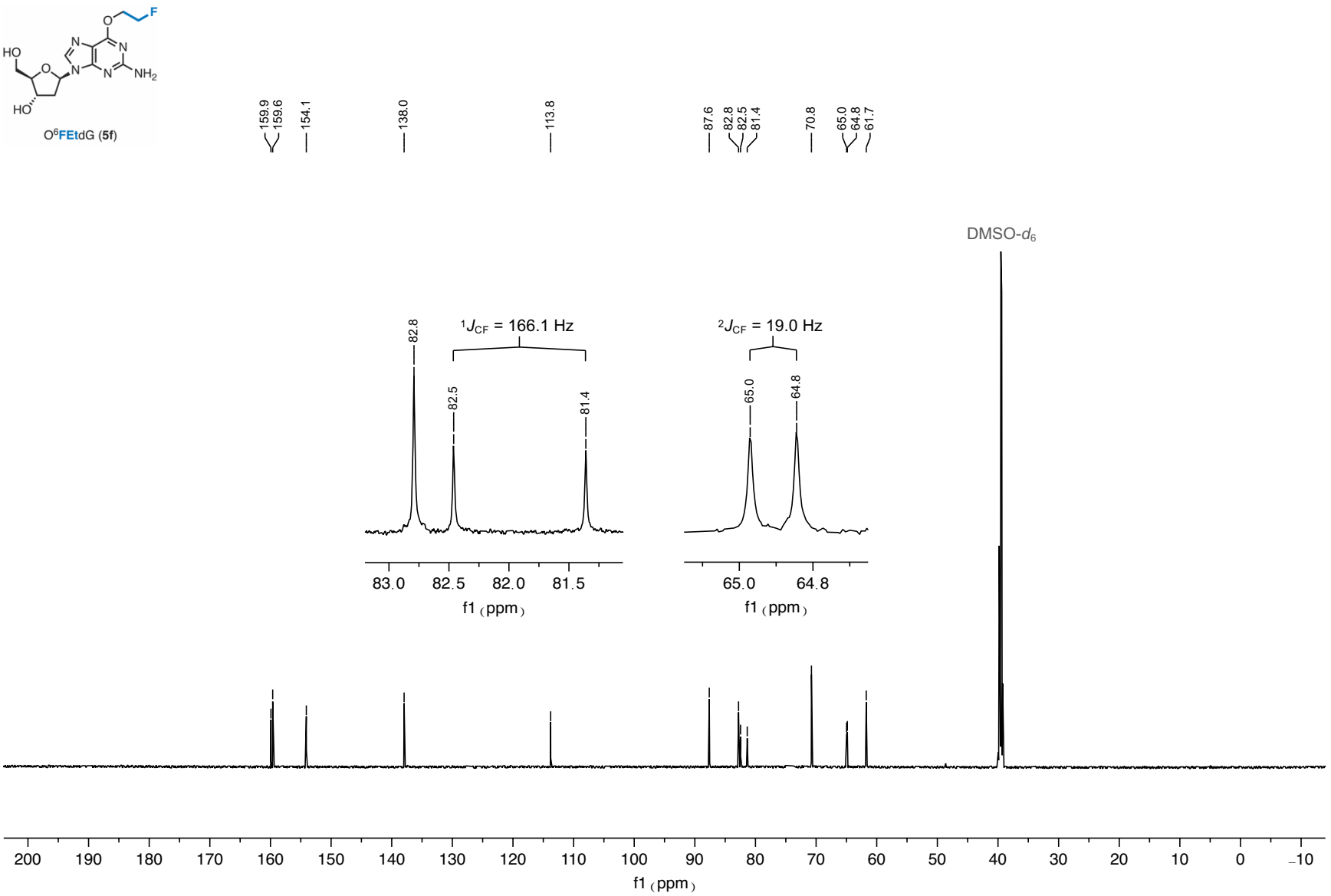
^{1}H–^{13}C HSQC{^{1}H}{^{13}C} NMR, 600/151 MHz, DMSO-d_6, $^{1}J_{CH}$ = 210 Hz, Blue = CH/CH_3, Red = CH_2



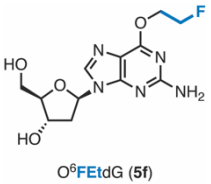
^{1}H NMR, 600 MHz, DMSO-d_6



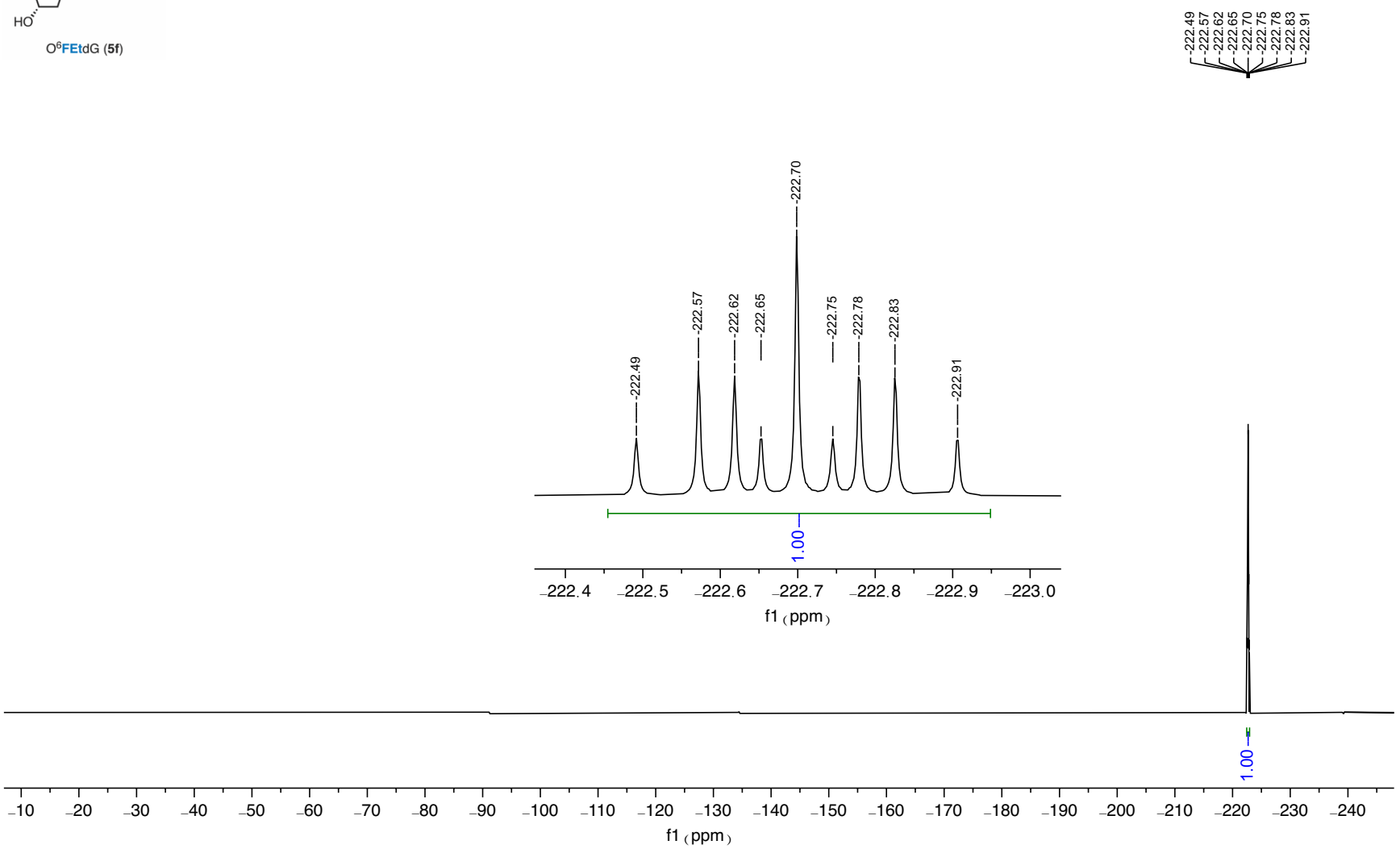
$^{13}C\{^{1}H\}$ NMR, 151 MHz, DMSO-d_6



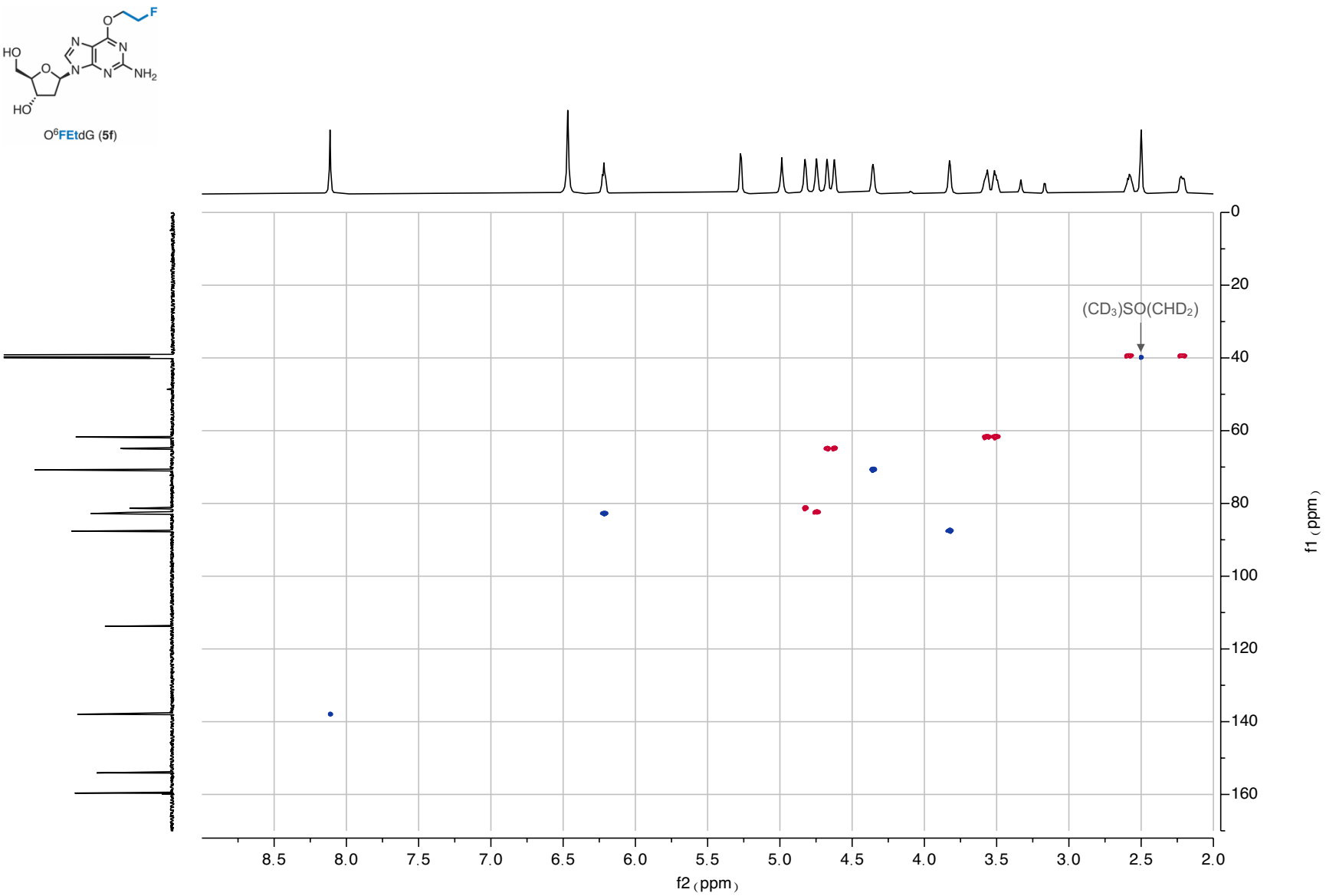
^{19}F NMR, 376 MHz, DMSO-d_6



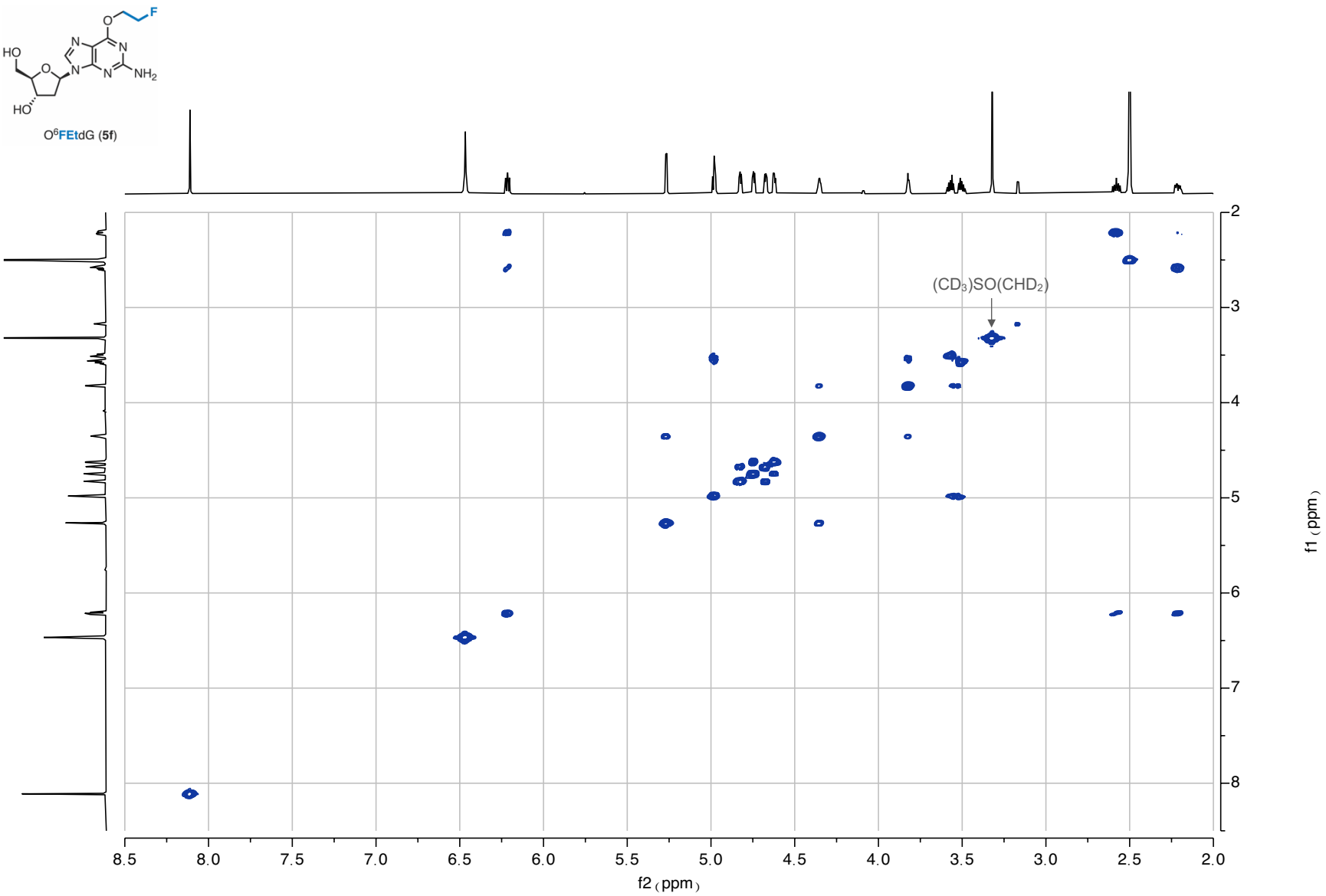
O^{6}FEtdG (**5f**)



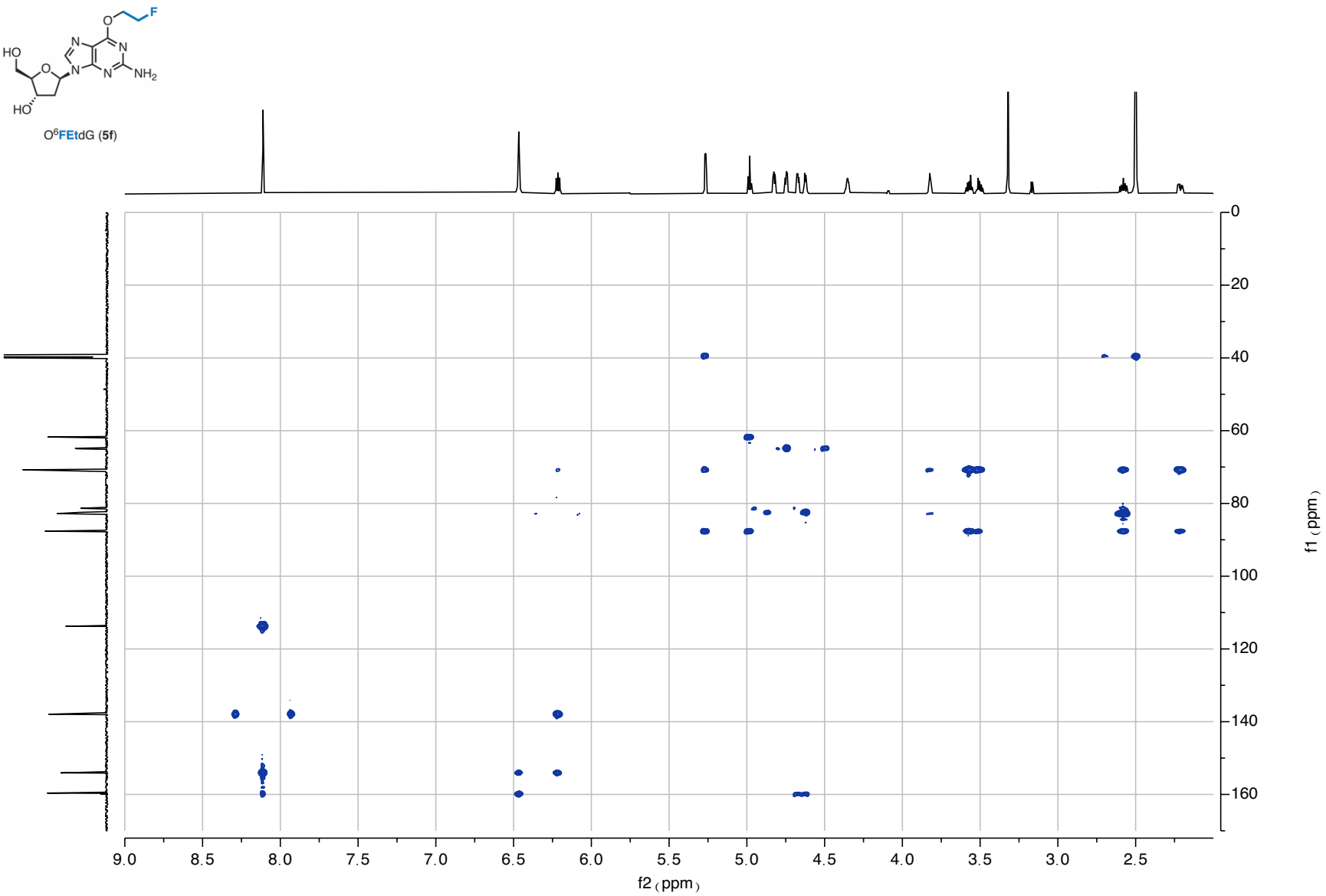
^{1}H–^{13}C HSQC{^{1}H}{^{13}C} NMR, 600/151 MHz, DMSO-d_6, $^{1}J_{CH}$ = 146 Hz, Blue = CH/CH_3, Red = CH_2



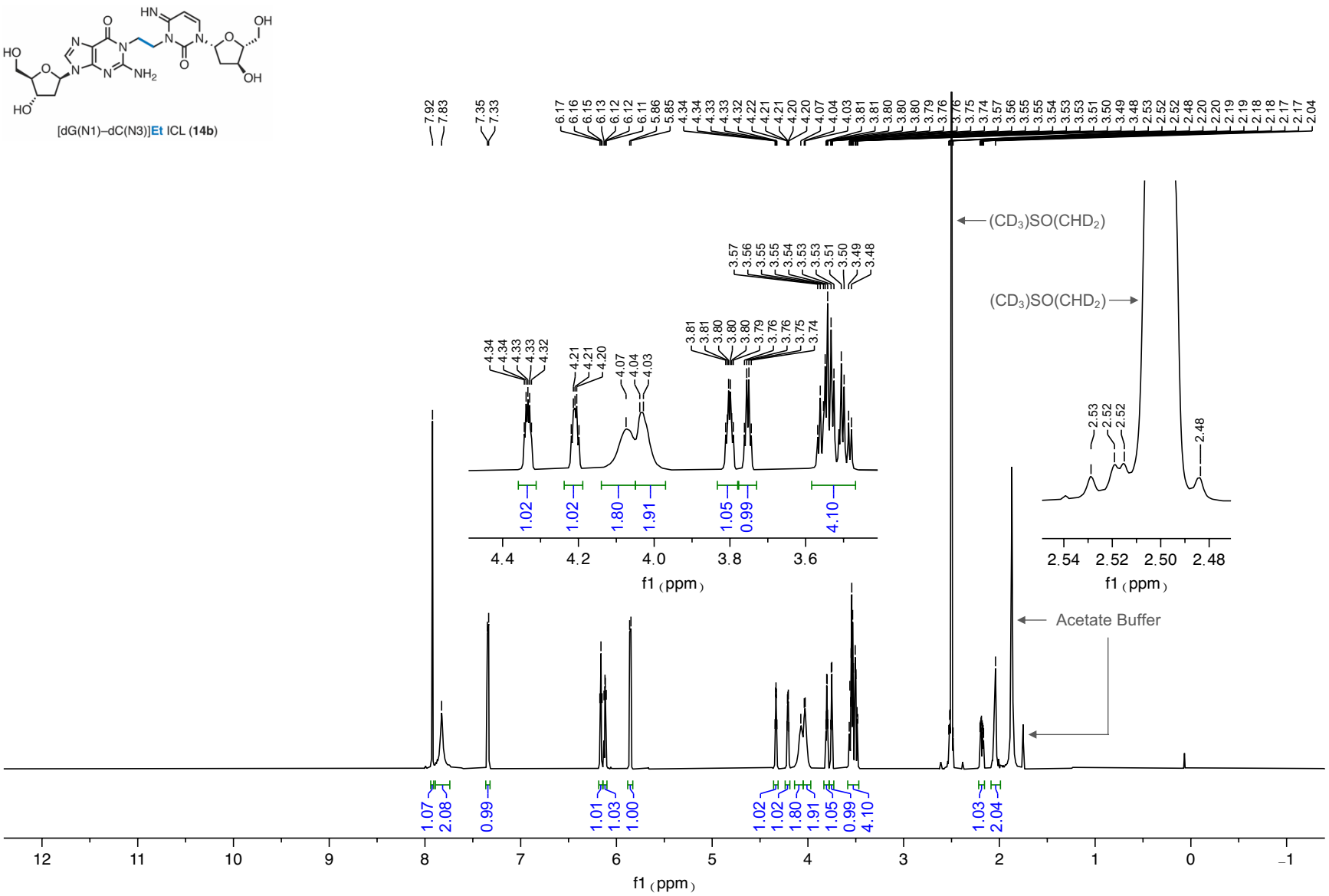
^{1}H–^{1}H gCOSY NMR, 600 MHz, DMSO-d_6



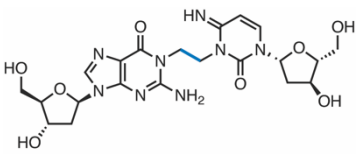
^{1}H–^{13}C gHMBC NMR, 600/151 MHz, DMSO-d_6



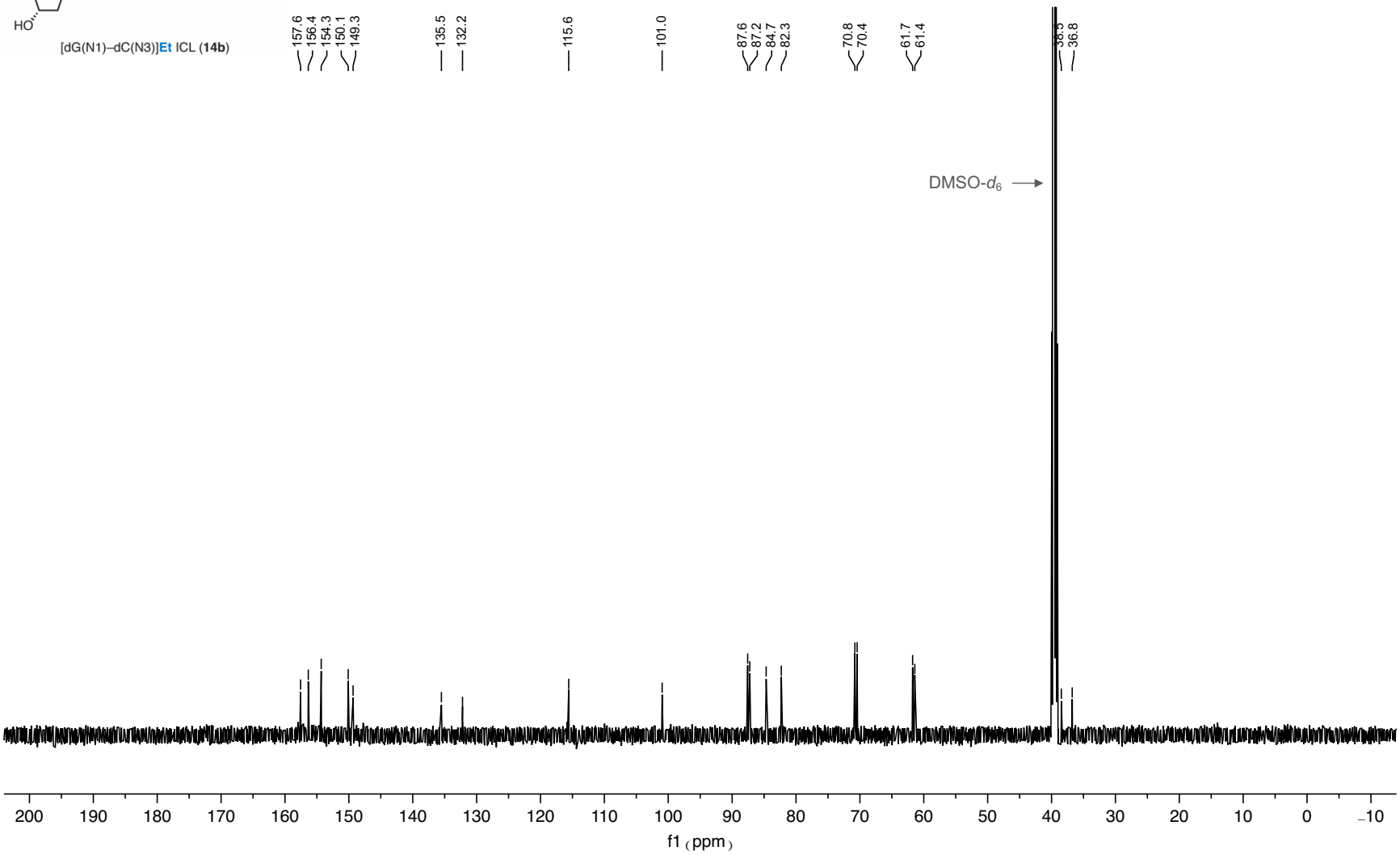
^{1}H NMR, 600 MHz, DMSO-d_6



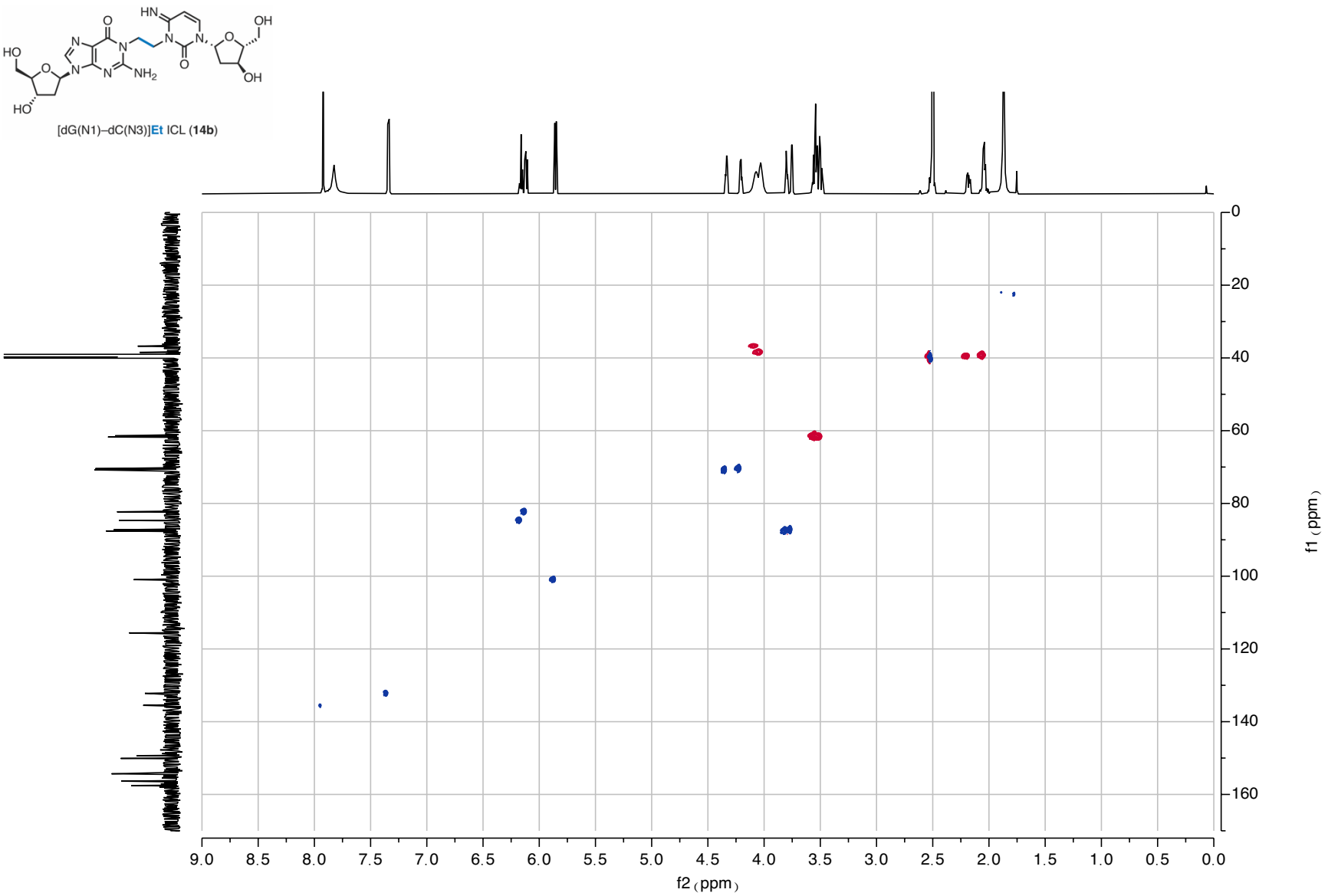
$^{13}C\{^{1}H\}$ NMR, 151 MHz, DMSO-d_6



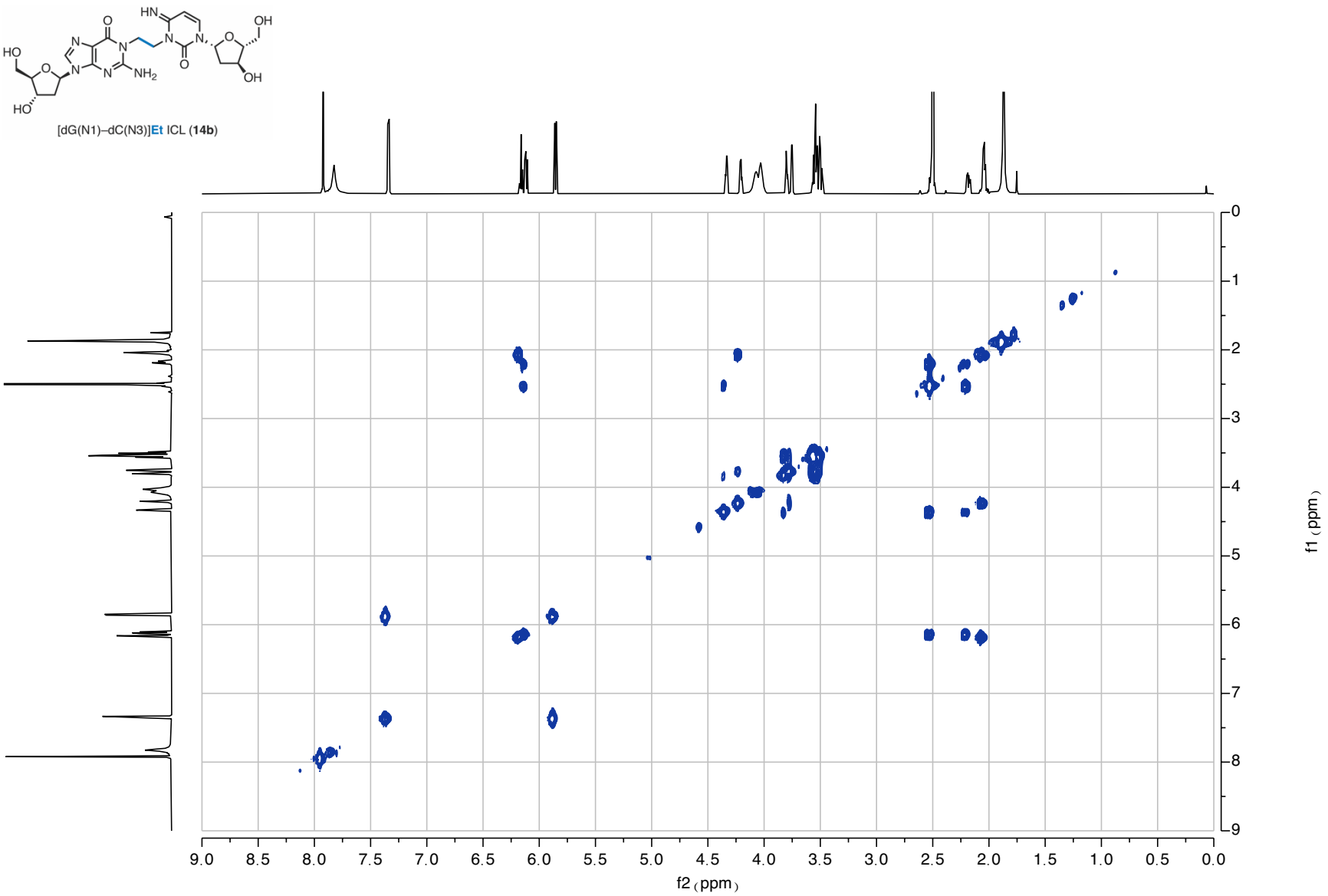
[dG(N1)–dC(N3)]Et ICL (**14b**)



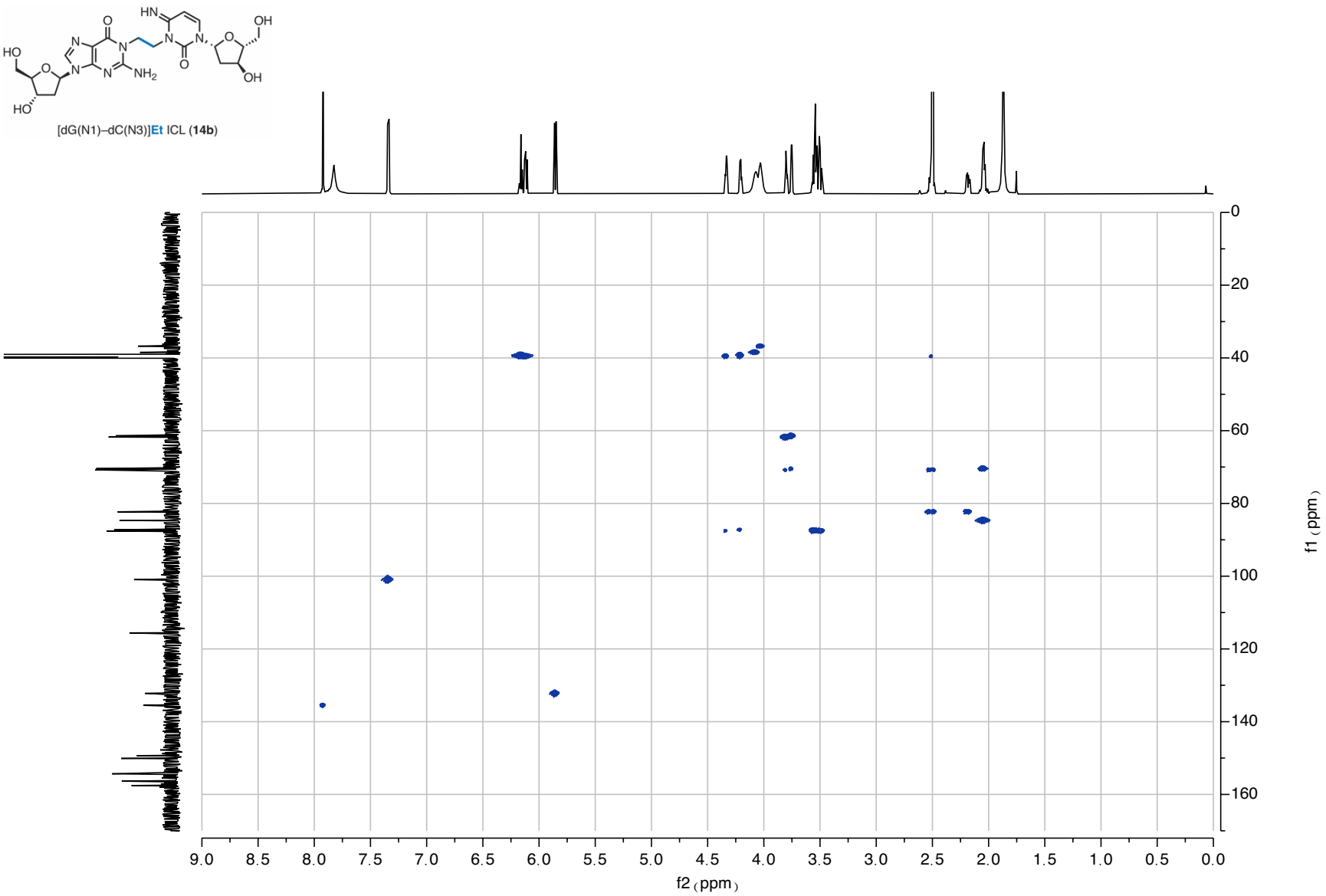
^{1}H–^{13}C HSQC{^{1}H}{^{13}C} NMR, 600/151 MHz, DMSO-d_6, $^{1}J_{CH}$ = 146 Hz, Blue = CH/CH_3, Red = CH_2



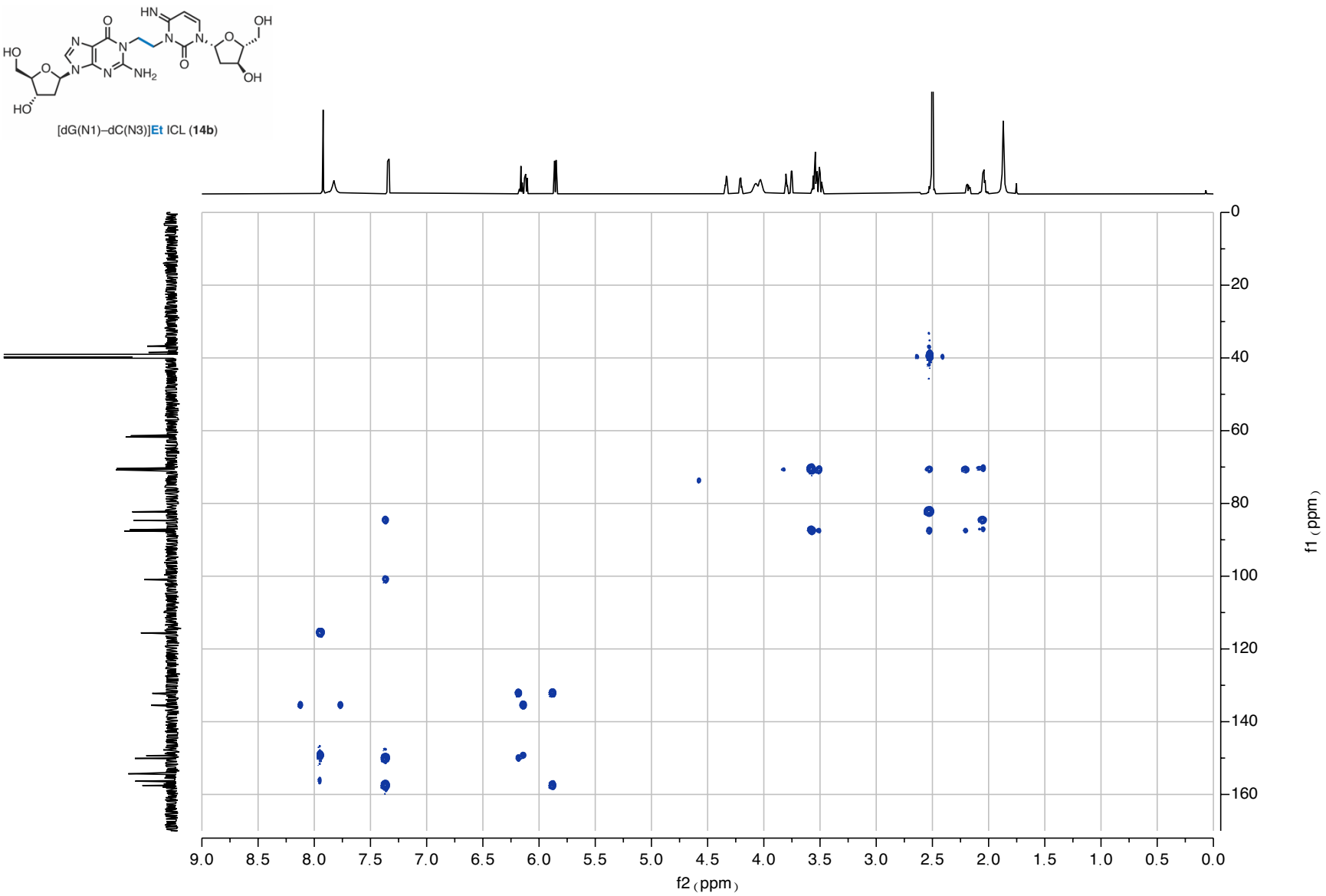
$^{1}H–^{13}C$ gCOSY NMR, 600 MHz, DMSO-d_6



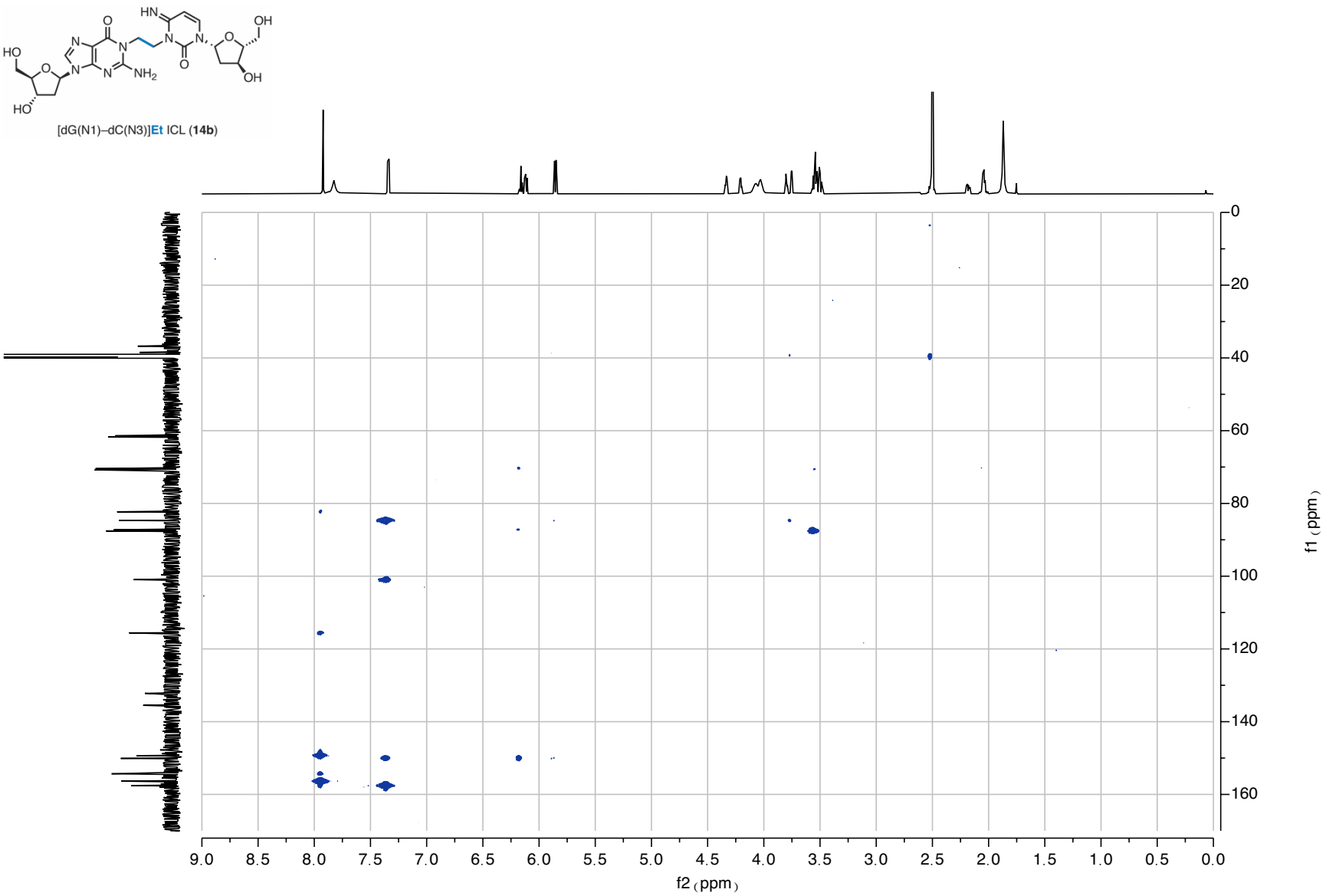
$^{1}H–^{13}C$ H2BC NMR, 600/151 MHz, DMSO-d_6



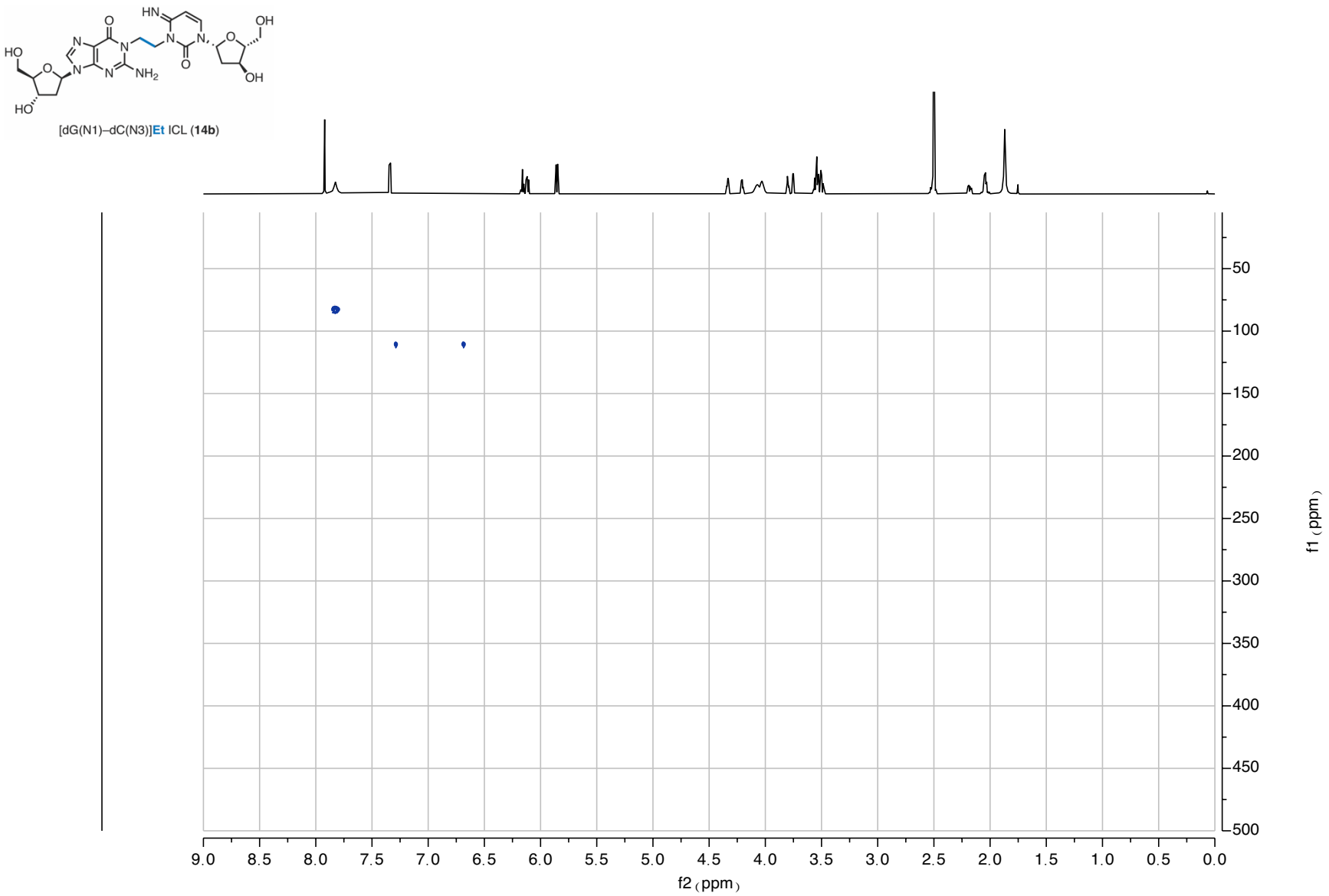
^{1}H–^{13}C gHMBC NMR, 600/151 MHz, DMSO-d_6



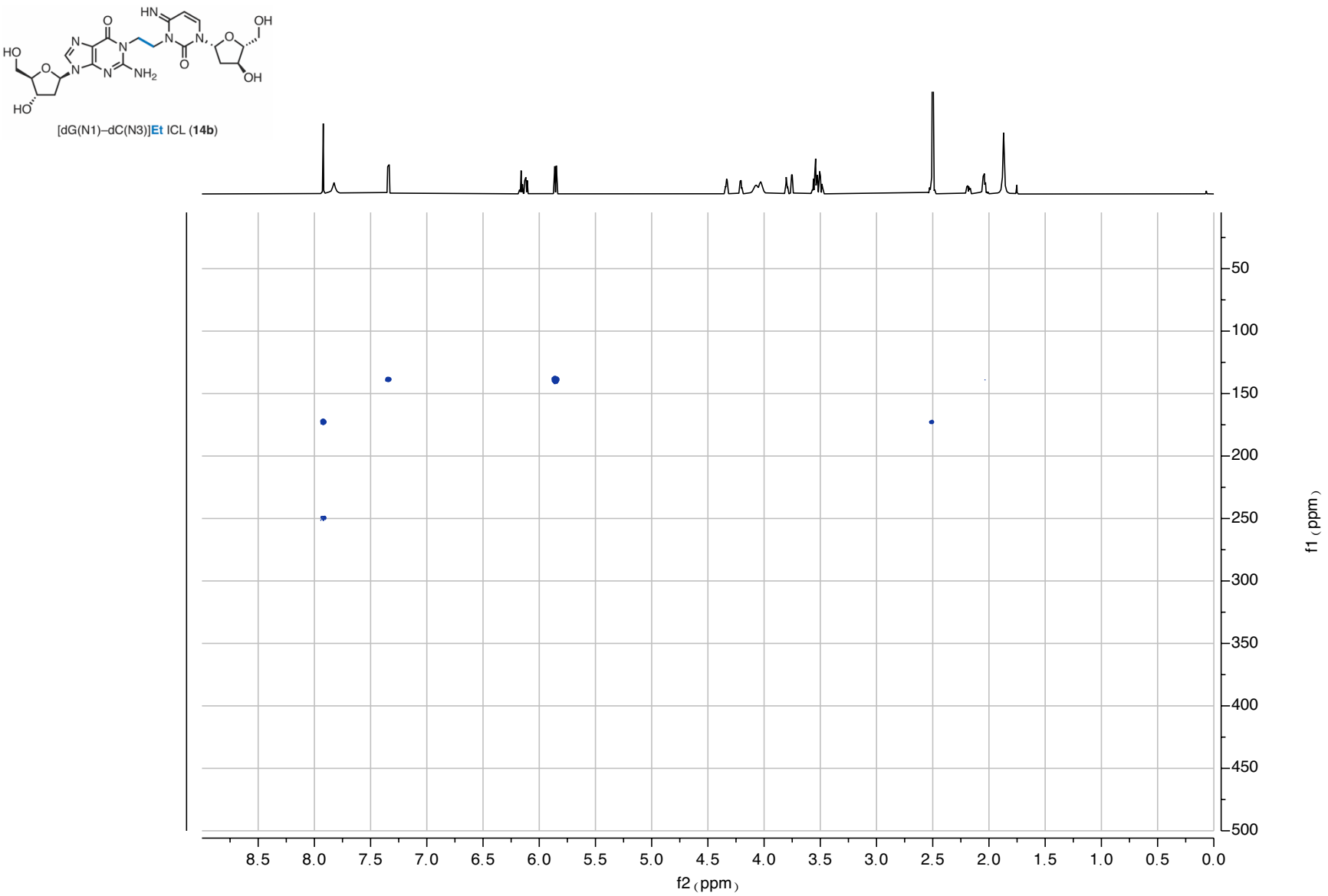
^{1}H–^{13}C LR-HSQMBC NMR, 600/151 MHz, DMSO-d_6



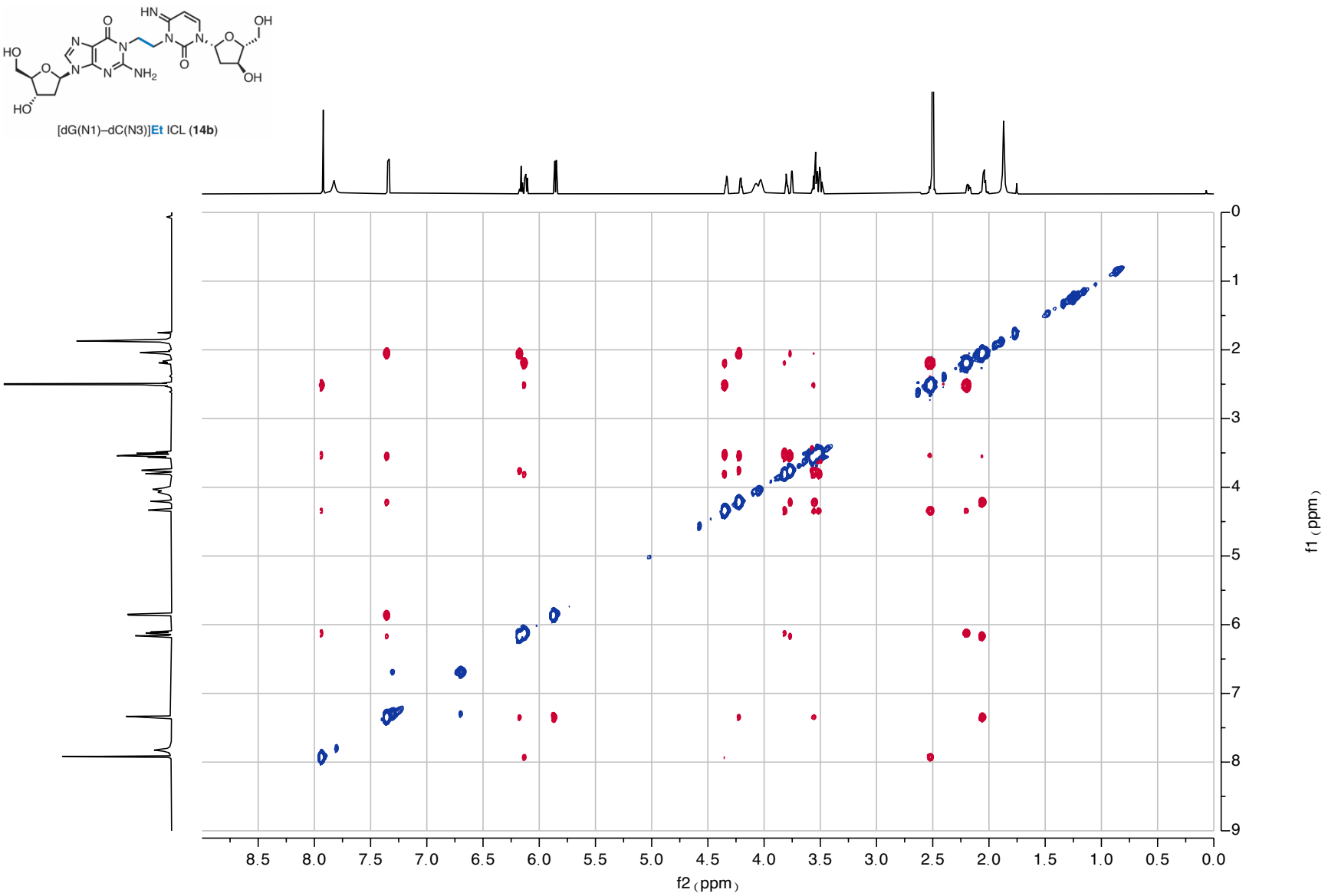
$^{1}H–^{15}N$ HSQC NMR, 600/61 MHz, DMSO-d_6



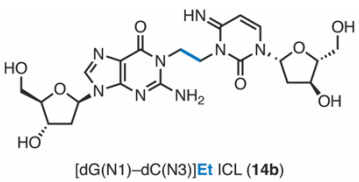
$^{1}H–^{15}N$ gHMBC NMR, 600/61 MHz, DMSO-d_6



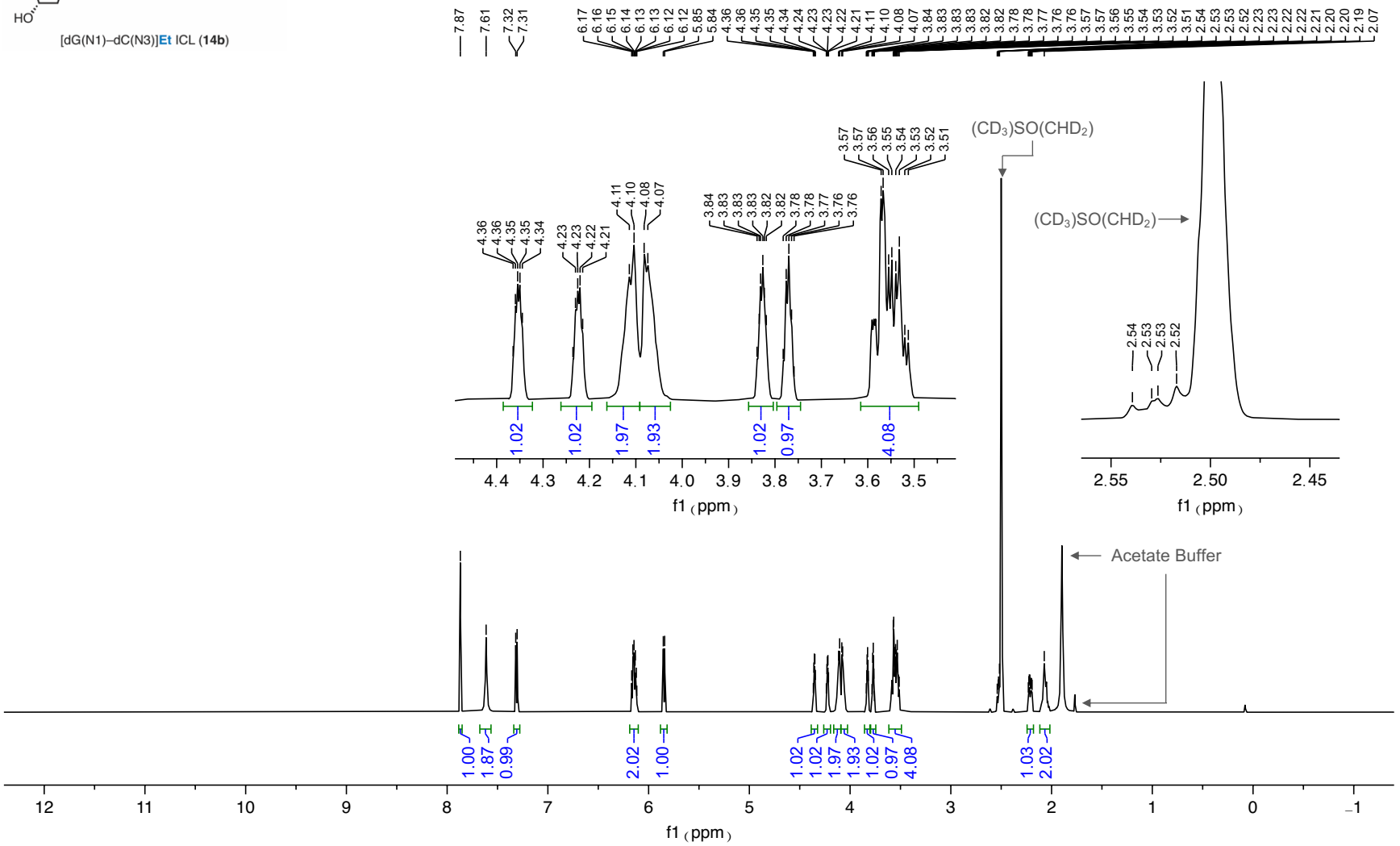
^{1}H–^{1}H ROESY NMR, 600 MHz, DMSO-d_6



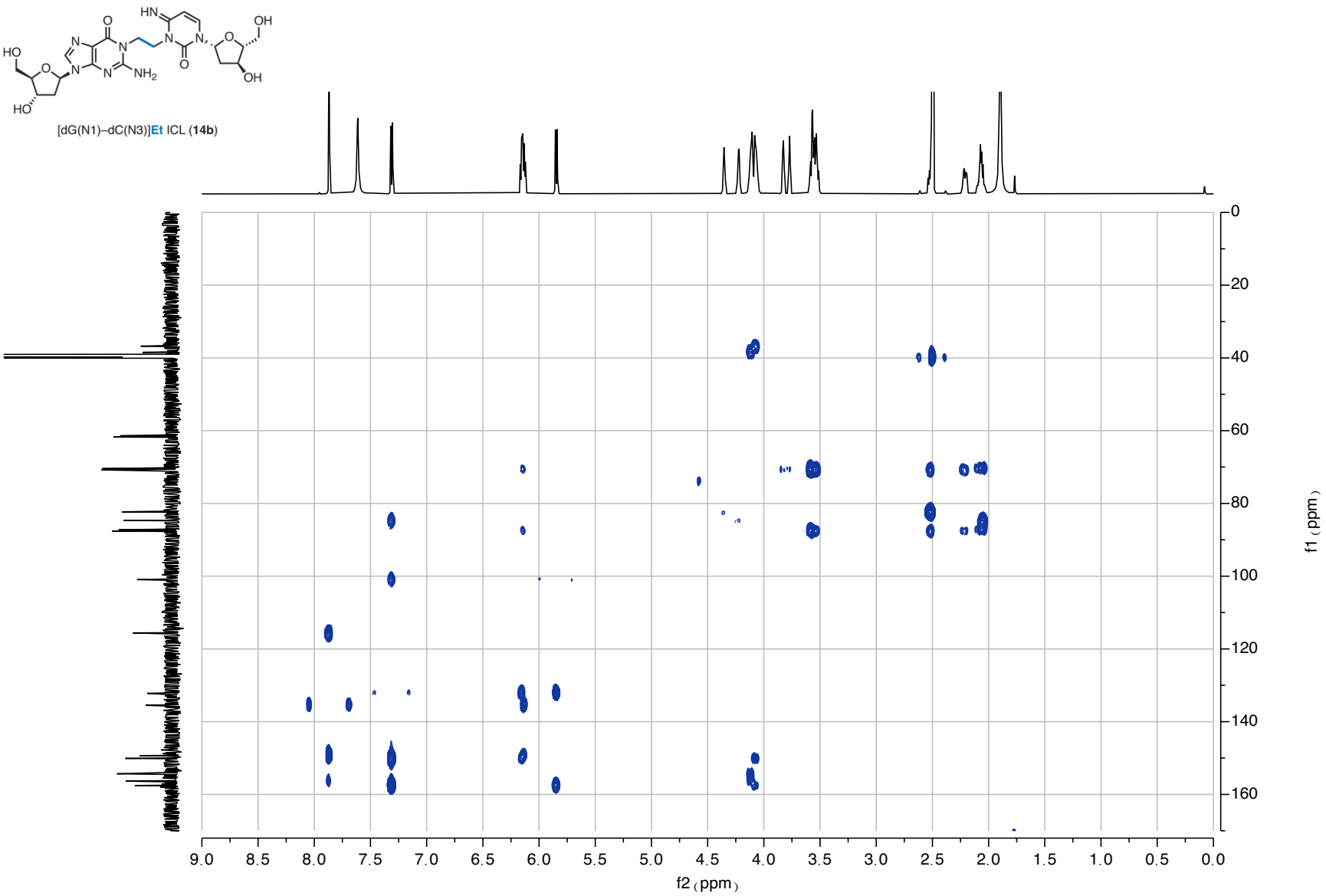
^{1}H NMR, 600 MHz, DMSO-d_6 at 65 ºC



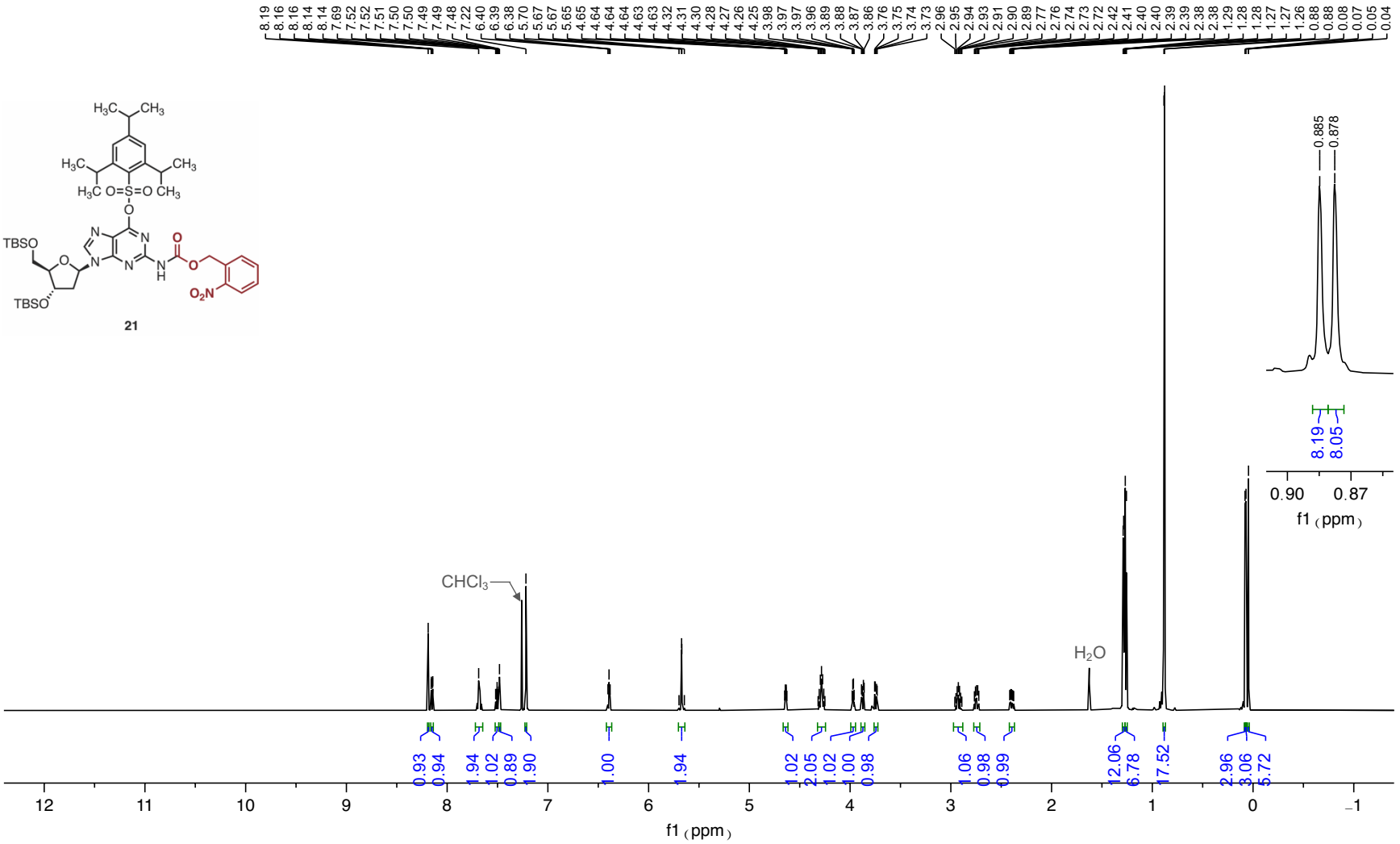
[dG(N1)–dC(N3)]Et ICL (**14b**)



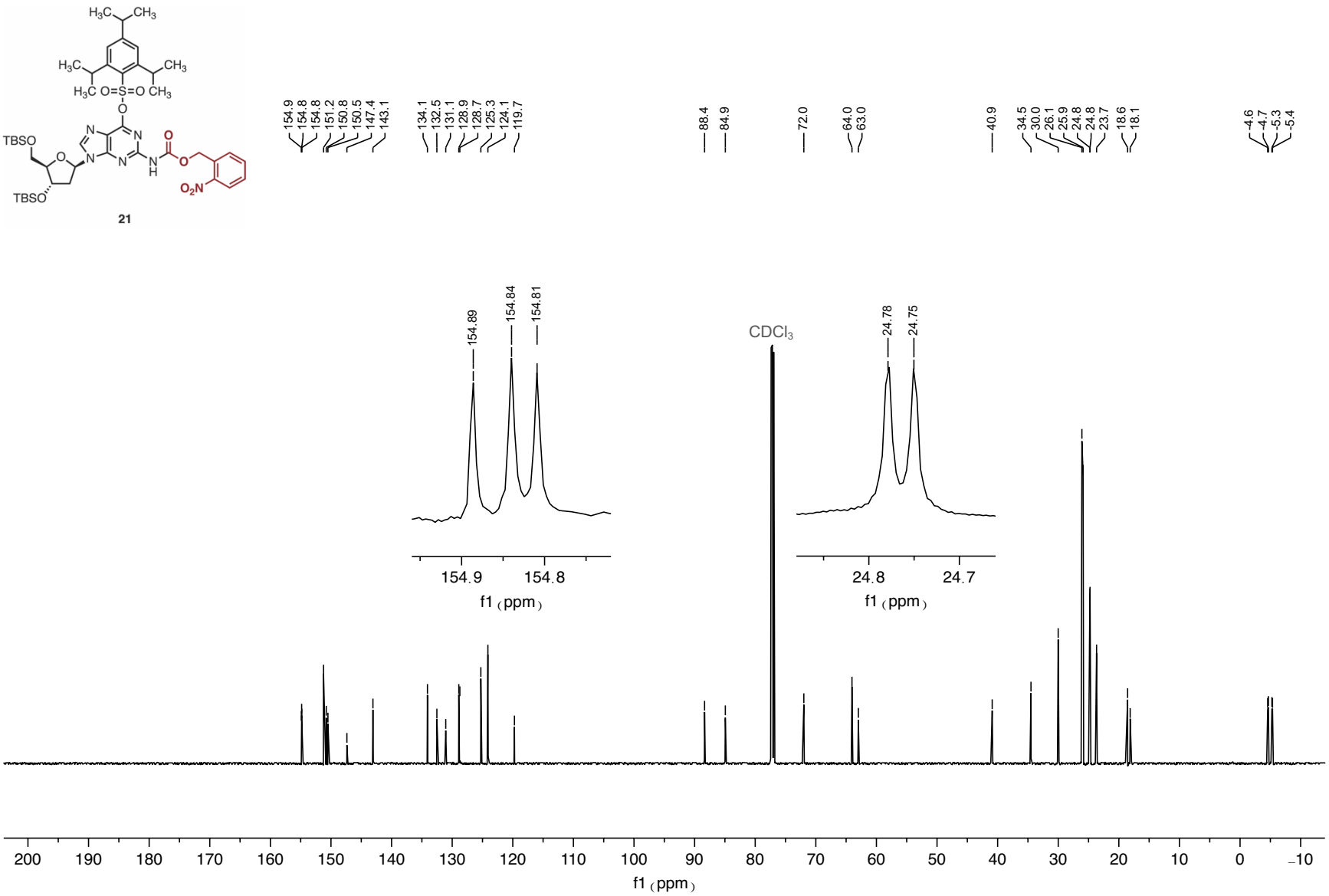
^{1}H–^{13}C gHMBC NMR, 600/151 MHz, DMSO-d_6 at 65 ºC



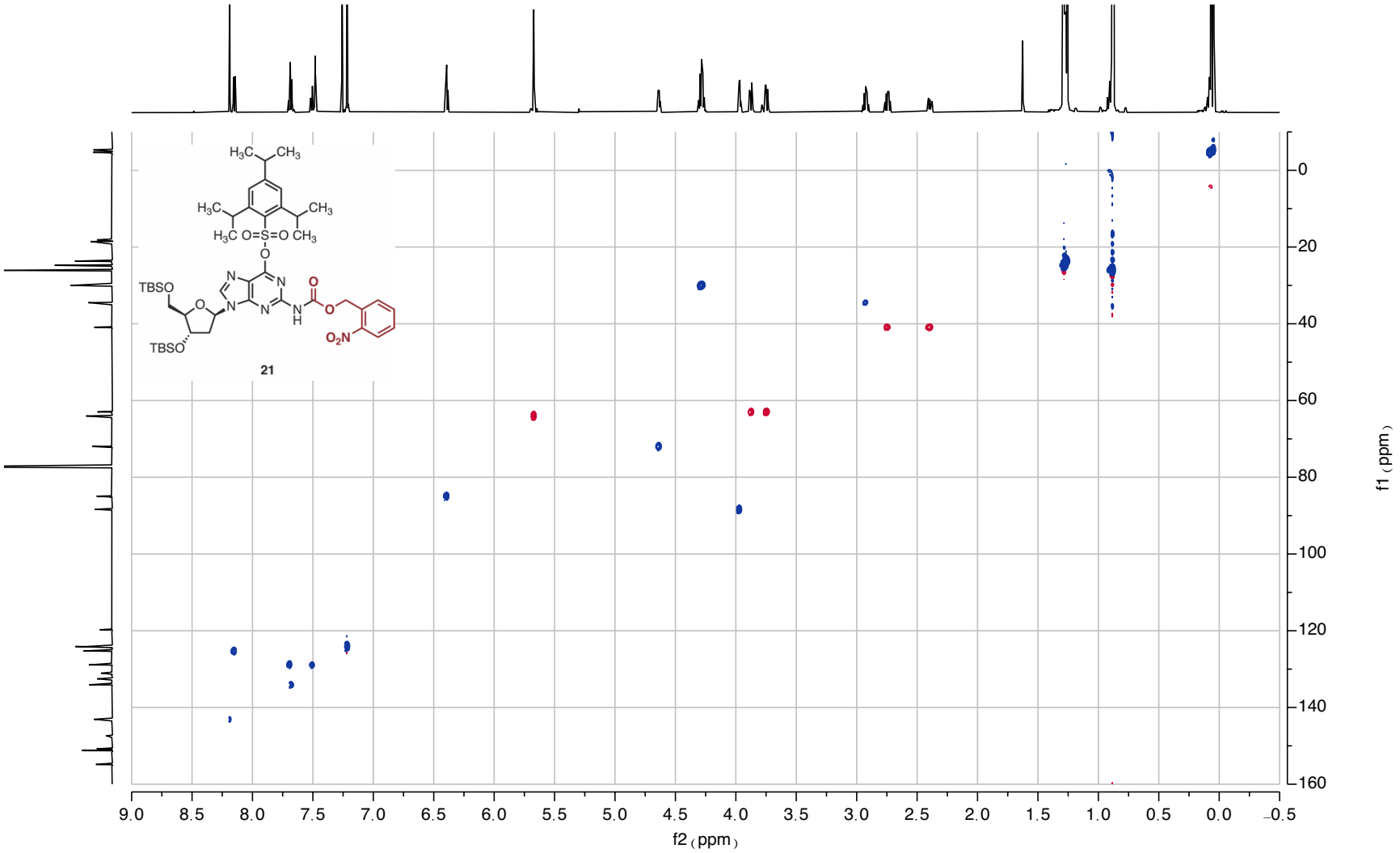
^{1}H NMR, 600 MHz, $CDCl_3$



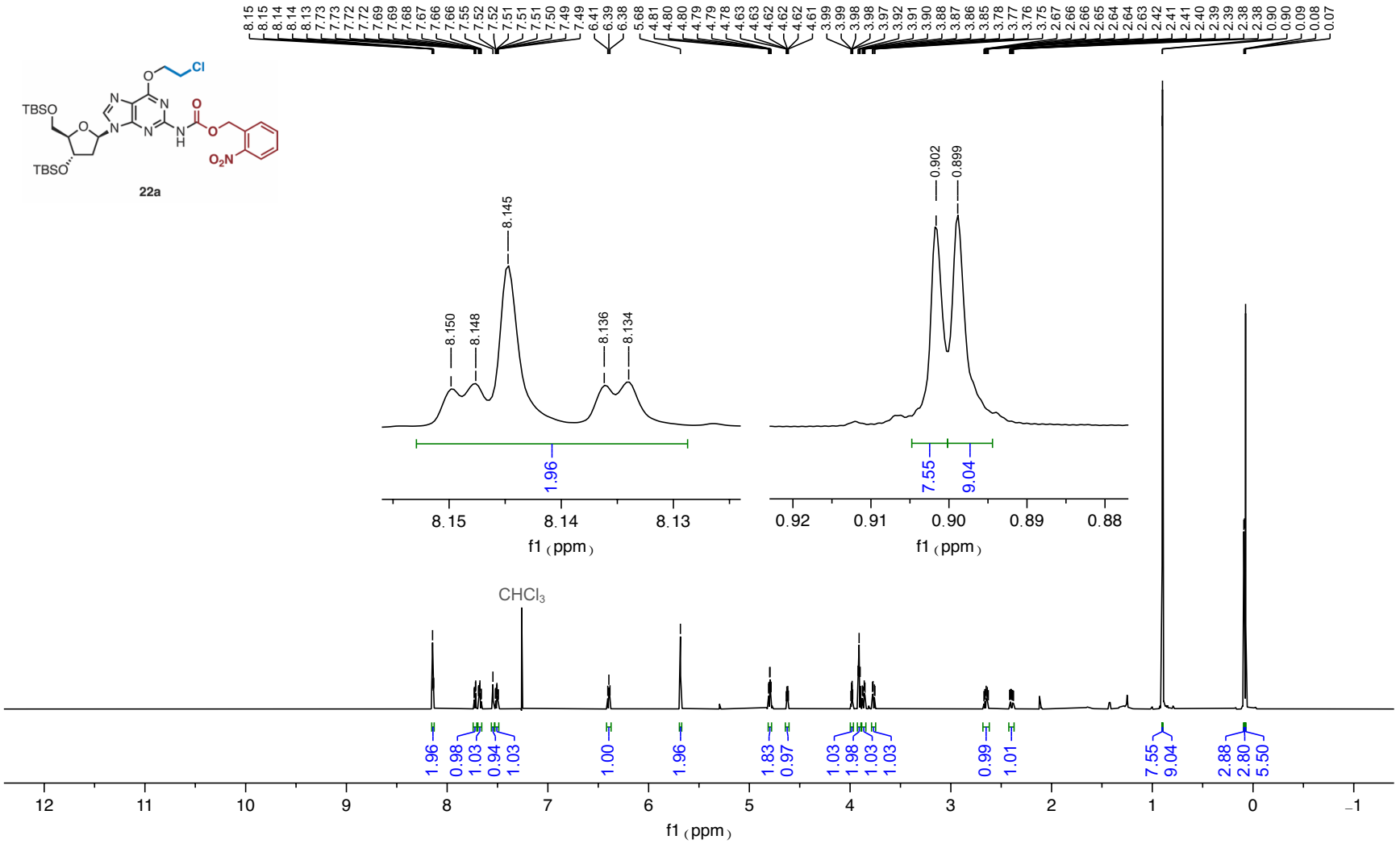
$^{13}C\{^{1}H\}$ NMR, 151 MHz, $CDCl_3$



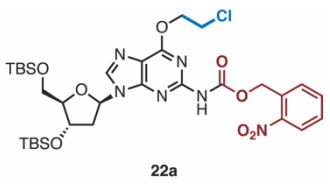
^{1}H–^{13}C HSQC{^{1}H}{^{13}C} NMR, 600/151 MHz, $CDCl_3$, $^{1}J_{CH}$ = 146 Hz, Blue = CH/CH_3, Red = CH_2



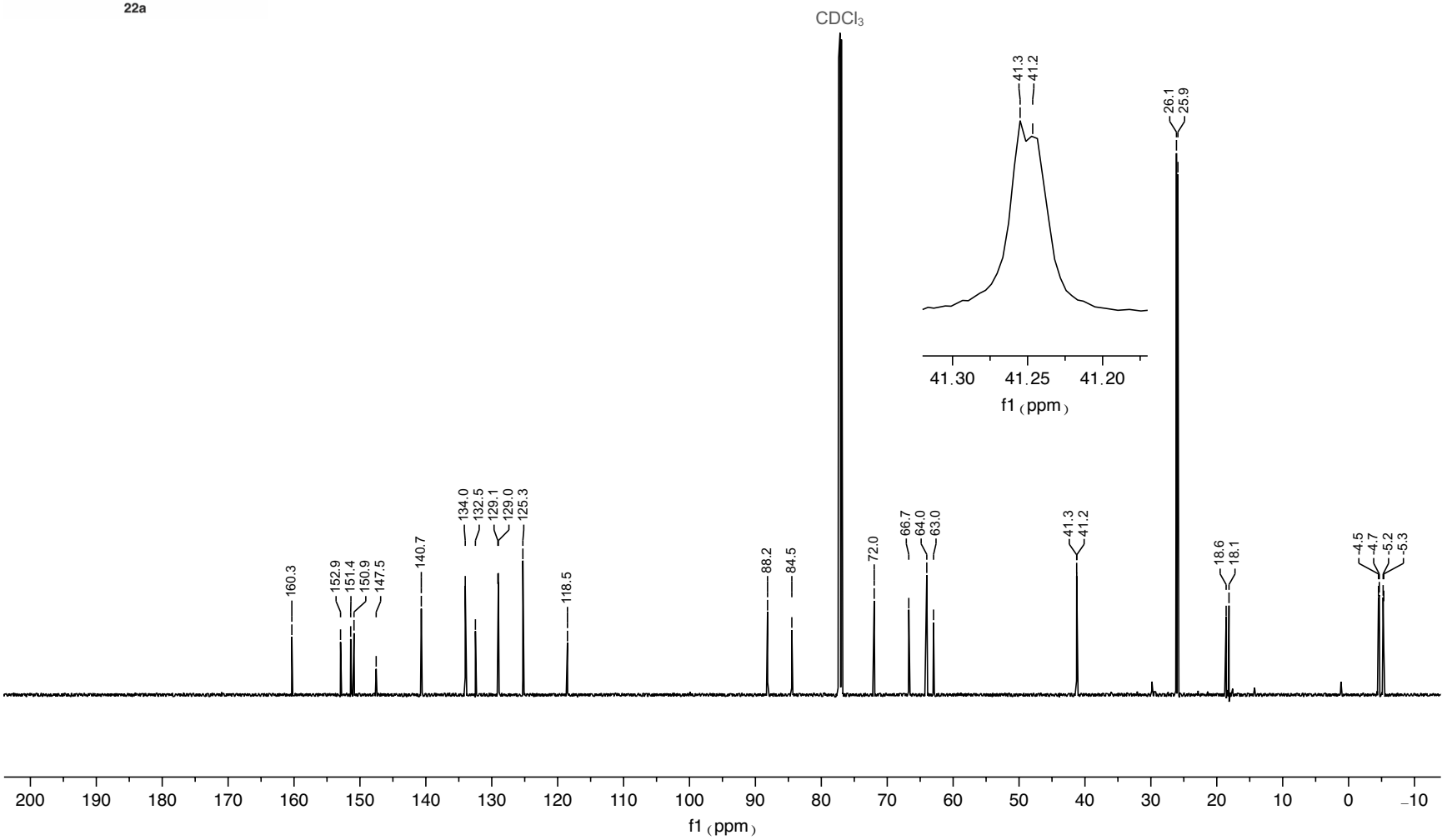
^{1}H NMR, 600 MHz, $CDCl_3$



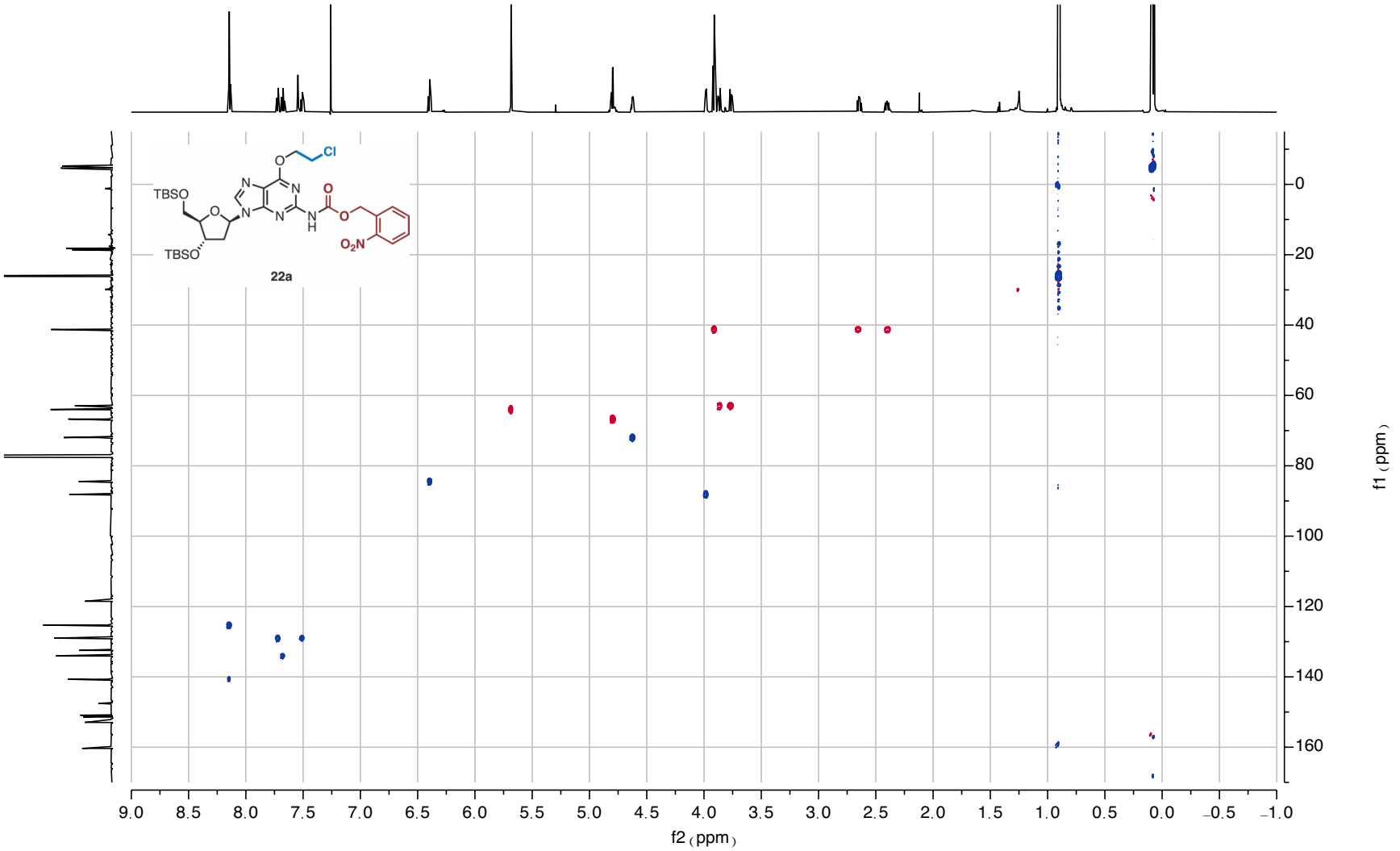
$^{13}C\{^{1}H\}$ NMR, 151 MHz, $CDCl_3$



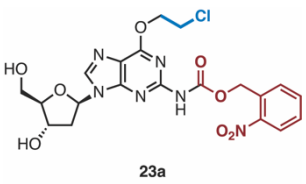
22a



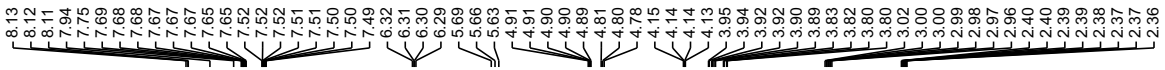
^{1}H–^{13}C HSQC{^{1}H}{^{13}C} NMR, 600/151 MHz, $CDCl_3$, $^{1}J_{CH}$ = 146 Hz, Blue = CH/CH_3, Red = CH_2

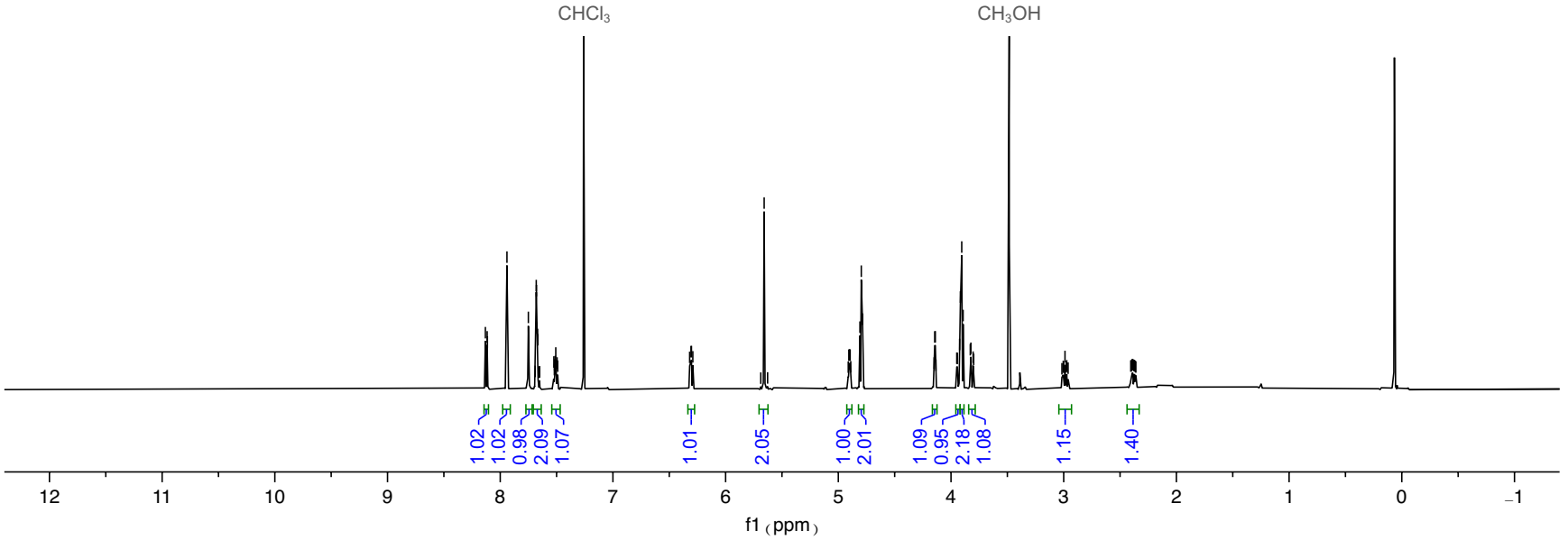


^{1}H NMR, 500 MHz, $CDCl_3$

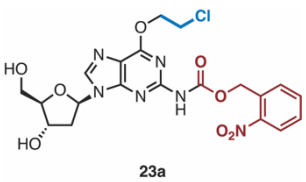


23a

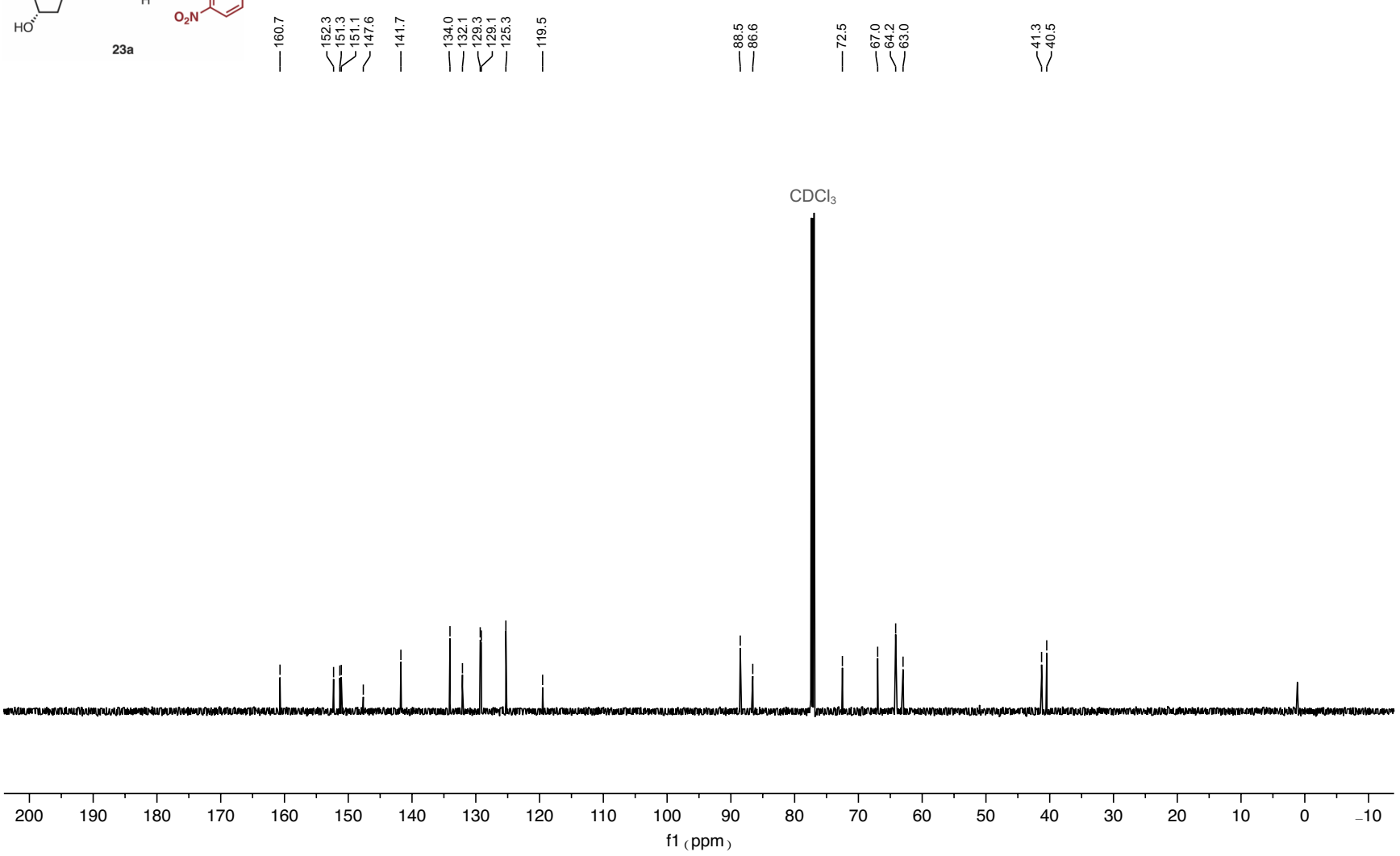




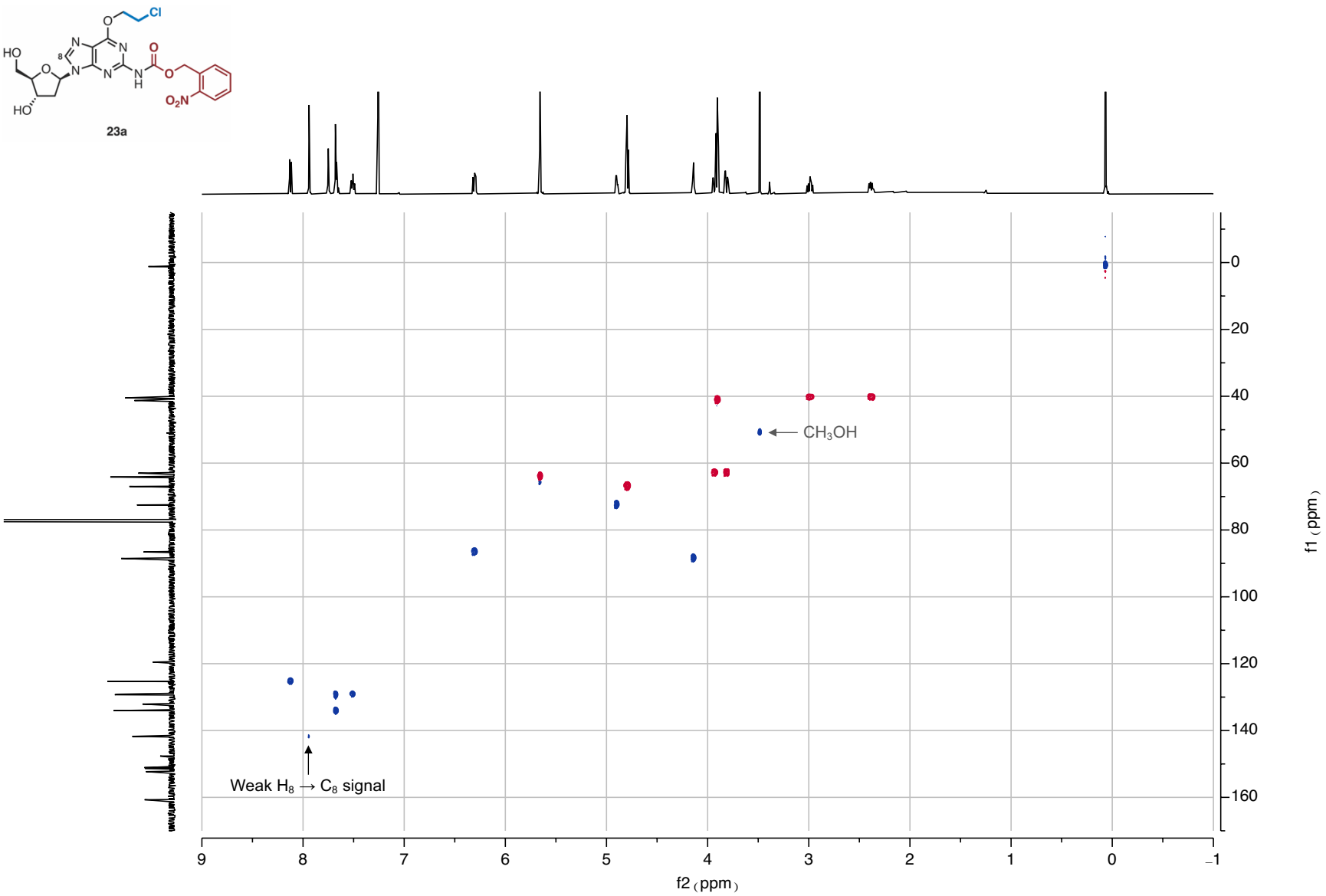
$^{13}C\{^{1}H\}$ NMR, 151 MHz, $CDCl_3$



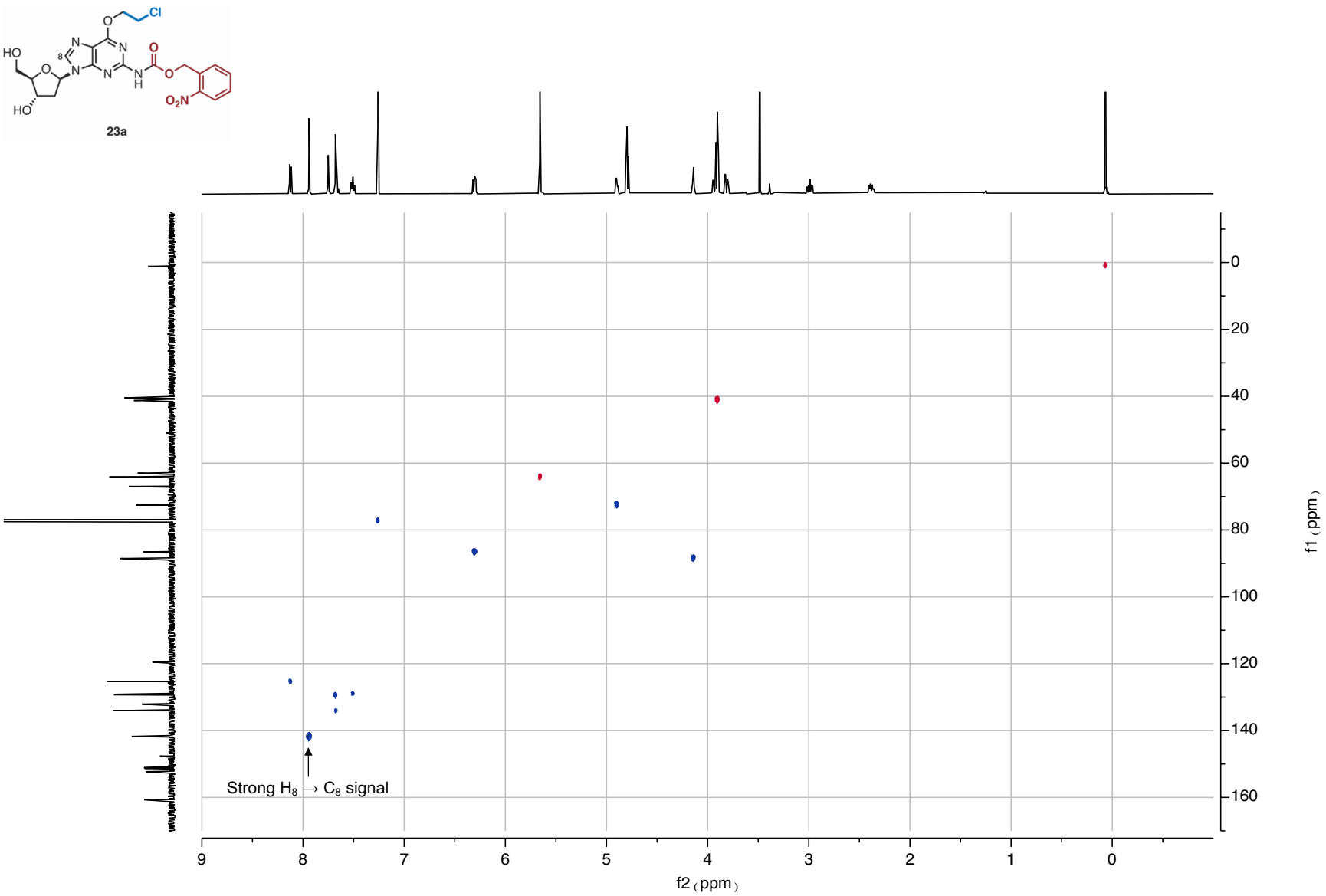
23a



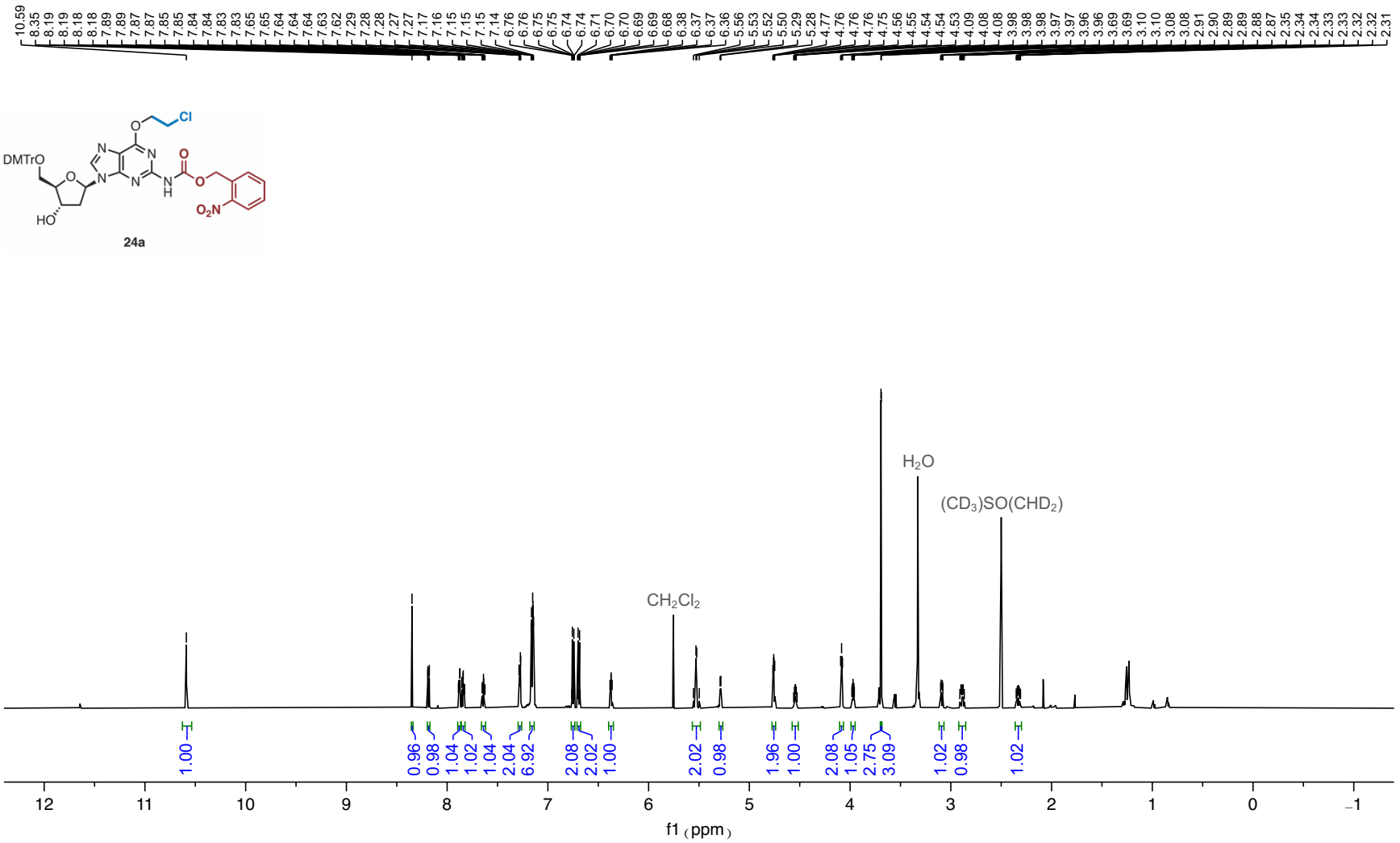
^{1}H–^{13}C HSQC{^{1}H}{^{13}C} NMR, 500/126 MHz, $CDCl_3$, $^{1}J_{CH}$ = 146 Hz, Blue = CH/CH_3, Red = CH_2



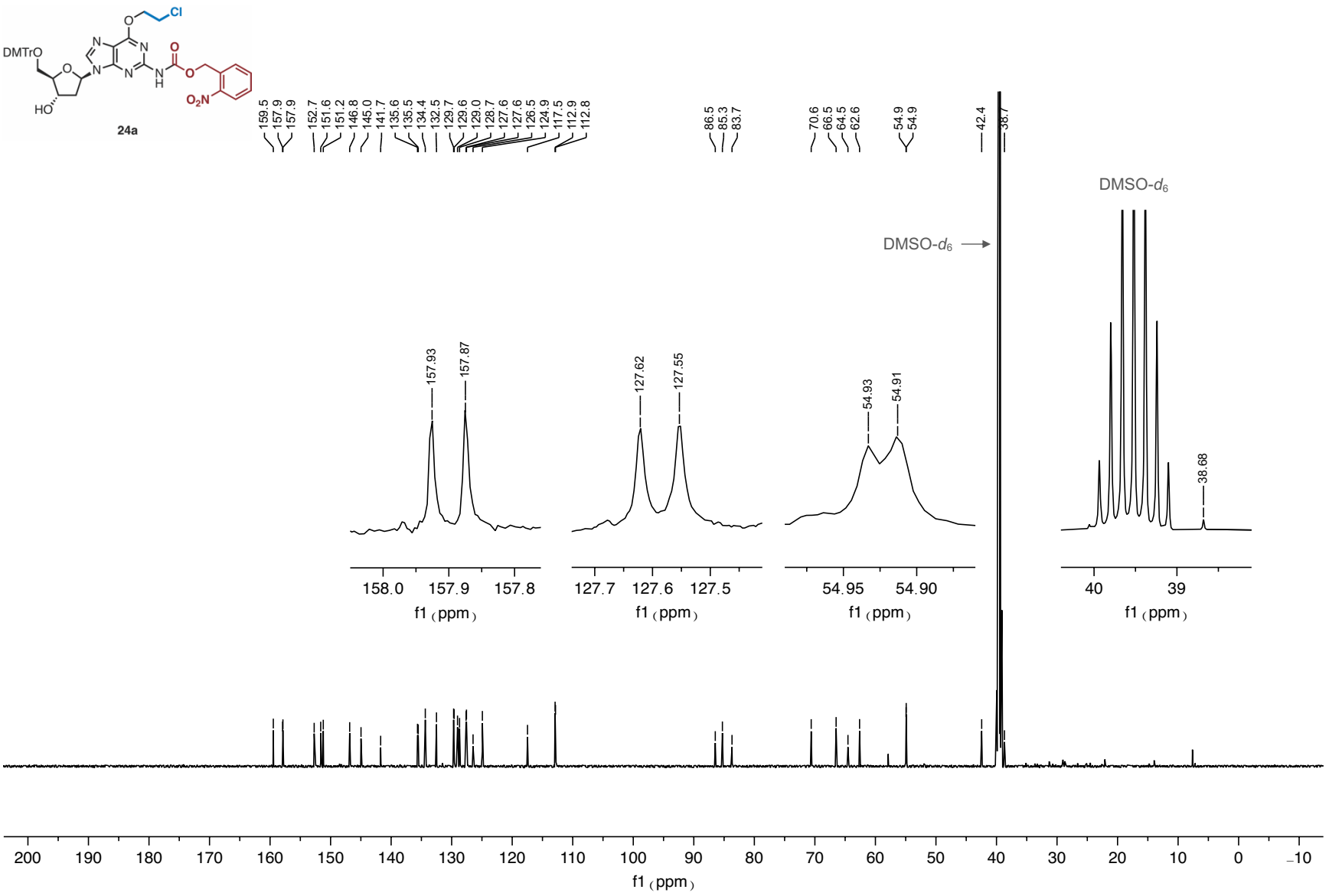
^{1}H–^{13}C HSQC{^{1}H}{^{13}C} NMR, 500/126 MHz, $CDCl_3$, $^{1}J_{CH}$ = 210 Hz, Blue = CH/CH_3, Red = CH_2



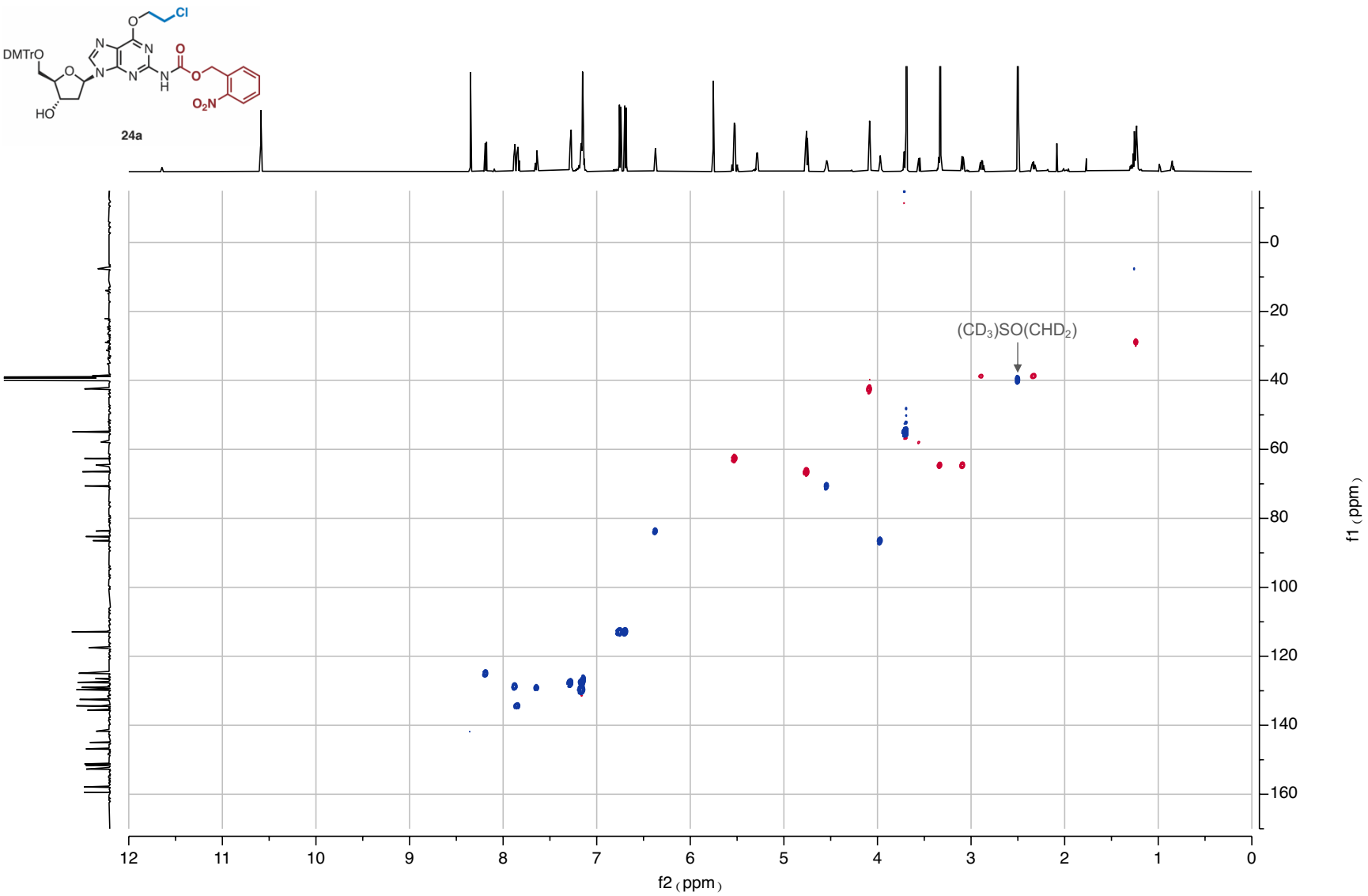
^{1}H NMR, 600 MHz, DMSO-d_6



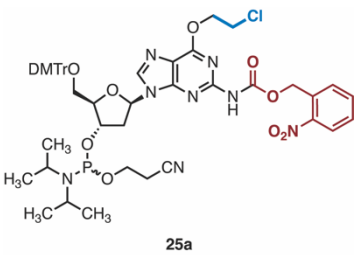
$^{13}C\{^{1}H\}$ NMR, 151 MHz, DMSO-d_6



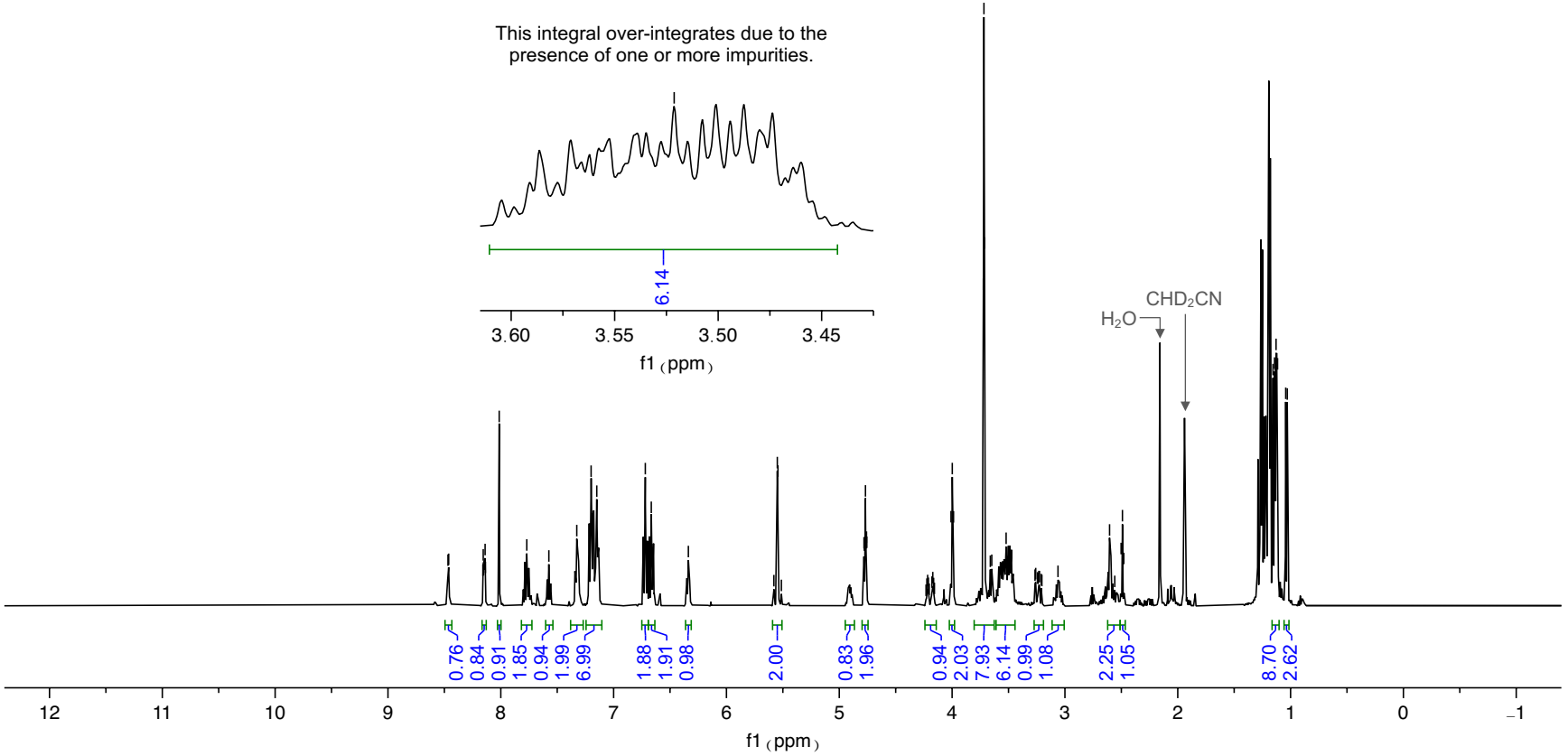
^{1}H–^{13}C HSQC{^{1}H}{^{13}C} NMR, 600/151 MHz, DMSO-d_6, $^{1}J_{CH}$ = 146 Hz, Blue = CH/CH_3, Red = CH_2



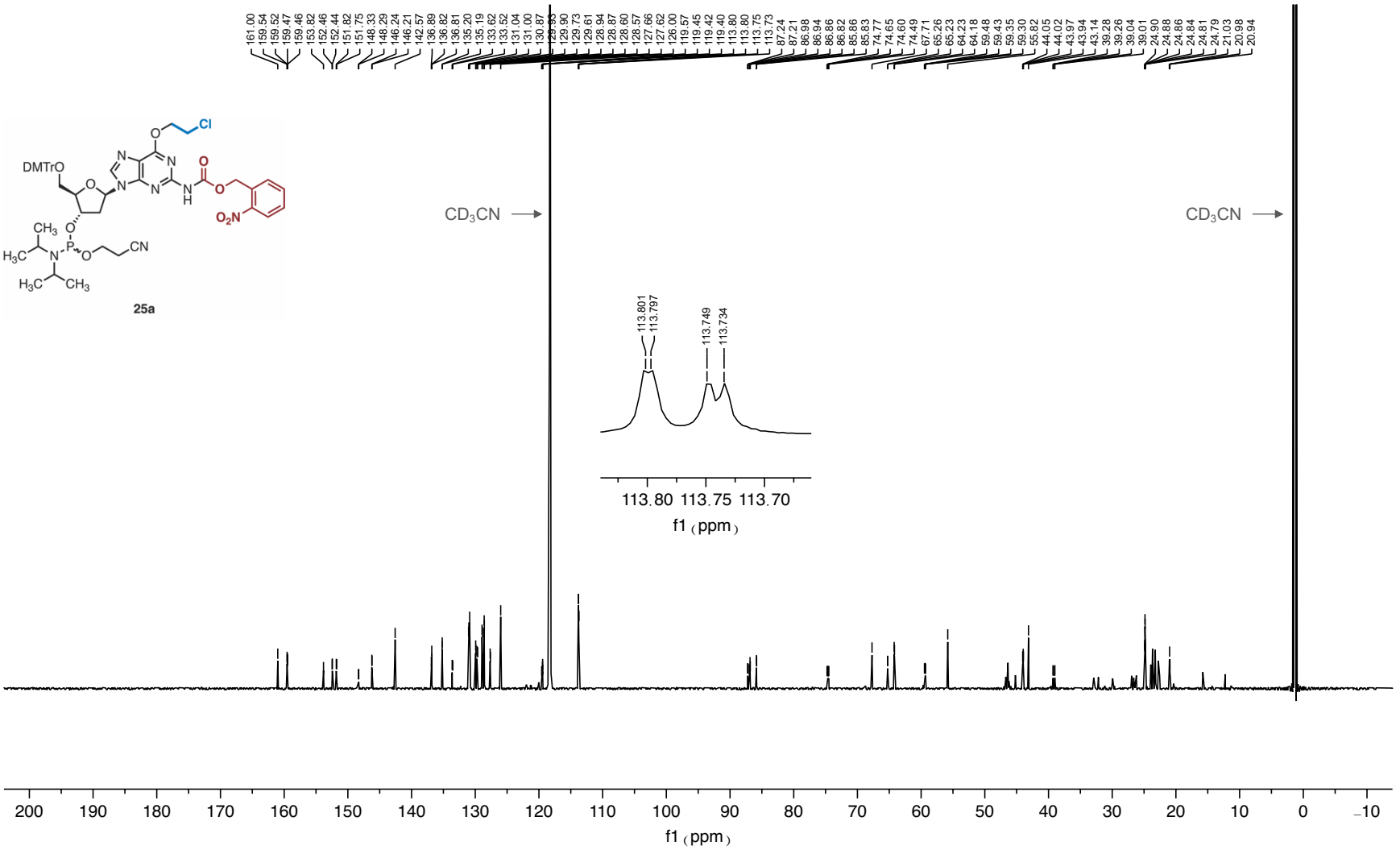
^{1}H NMR, 500 MHz, CD_3CN



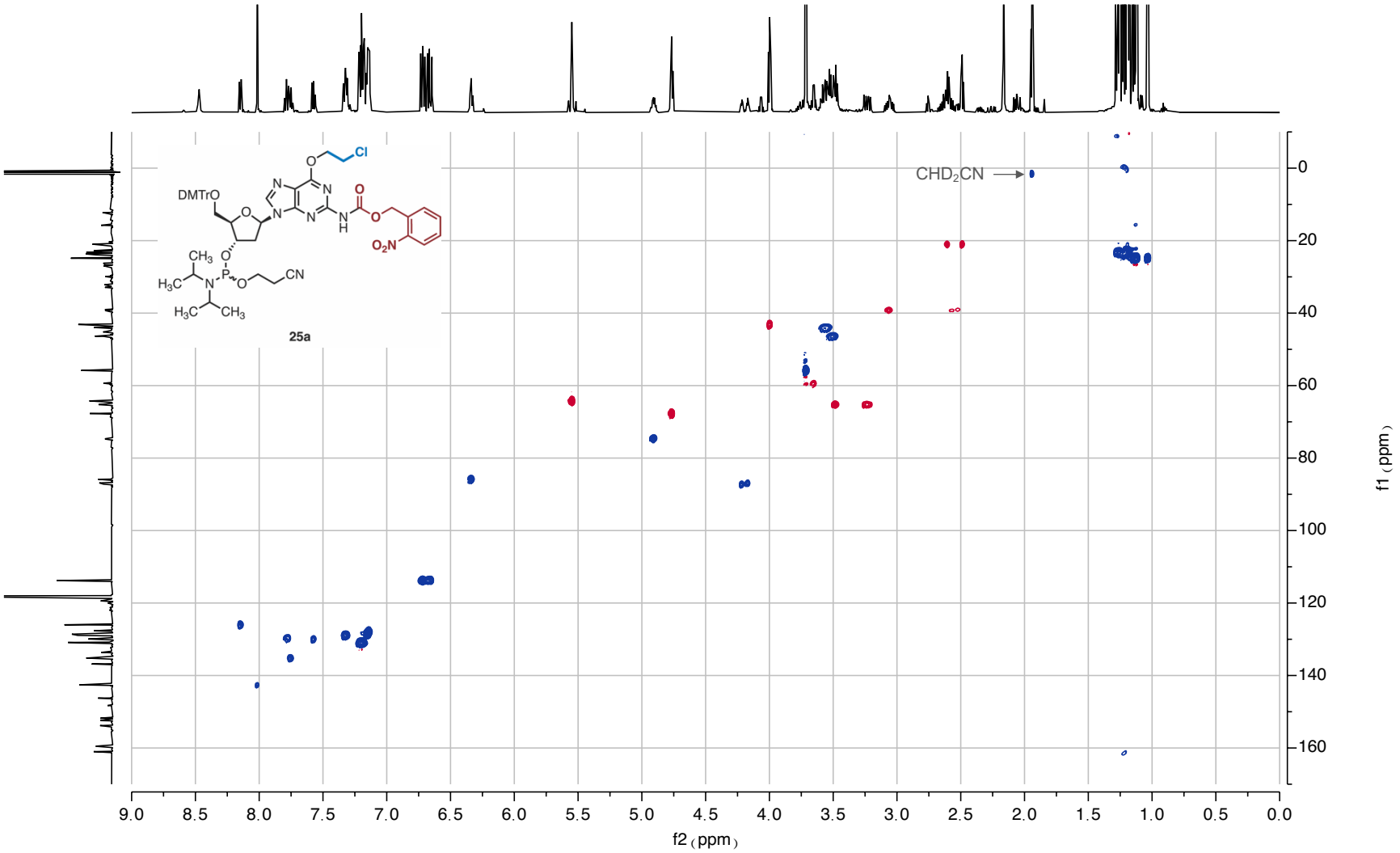
25a



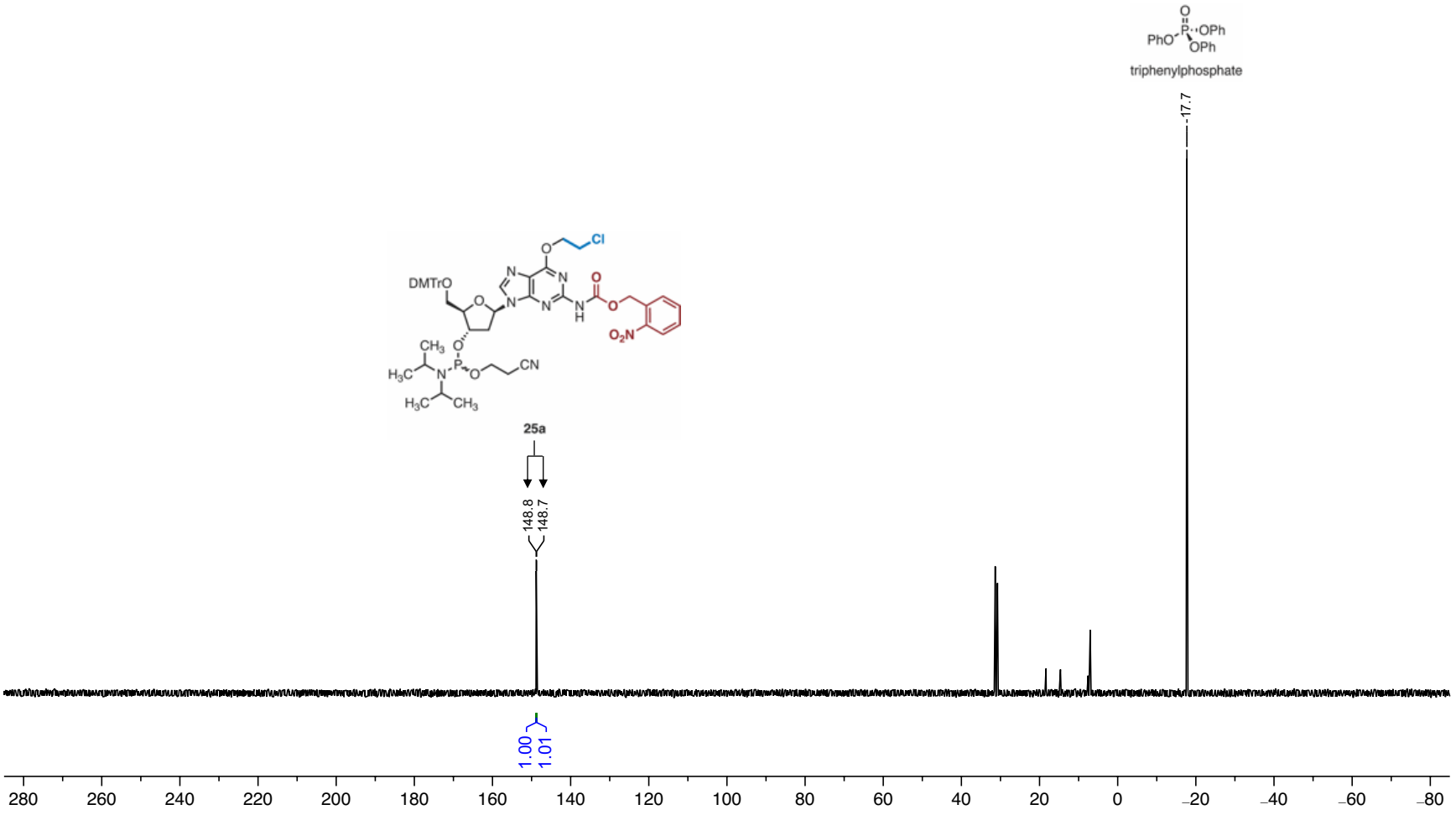
$^{13}C\{^{1}H\}$ NMR, 151 MHz, CD_3CN



^{1}H–^{13}C HSQC{^{1}H}{^{13}C} NMR, 600/151 MHz, CD_3CN, $^{1}J_{CH}$ = 146 Hz, Blue = CH/CH_3, Red = CH_2

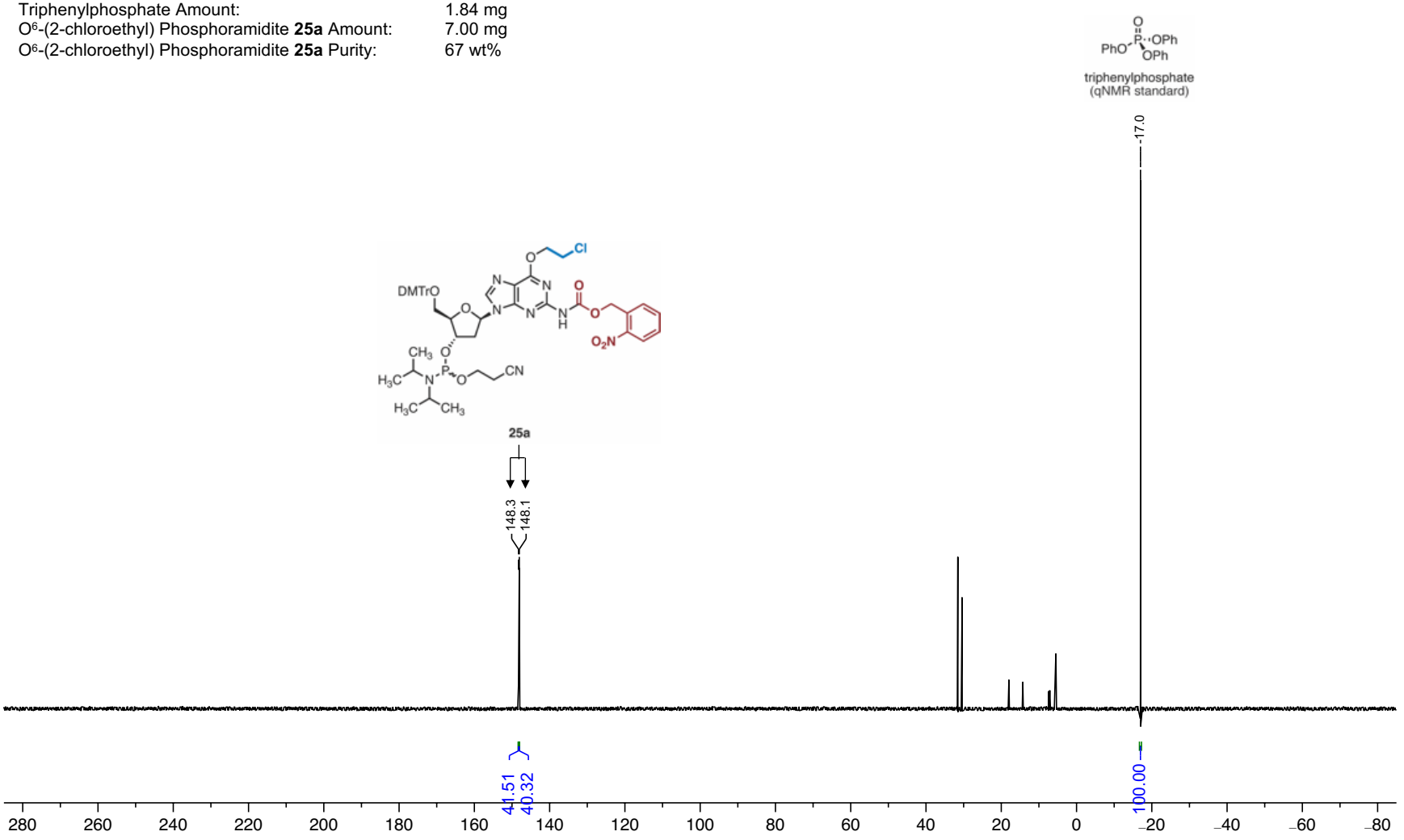


$^{31}P\{^{1}H\}$ NMR, 162 MHz, $CDCl_3$ (triphenyl phosphate added as internal reference)

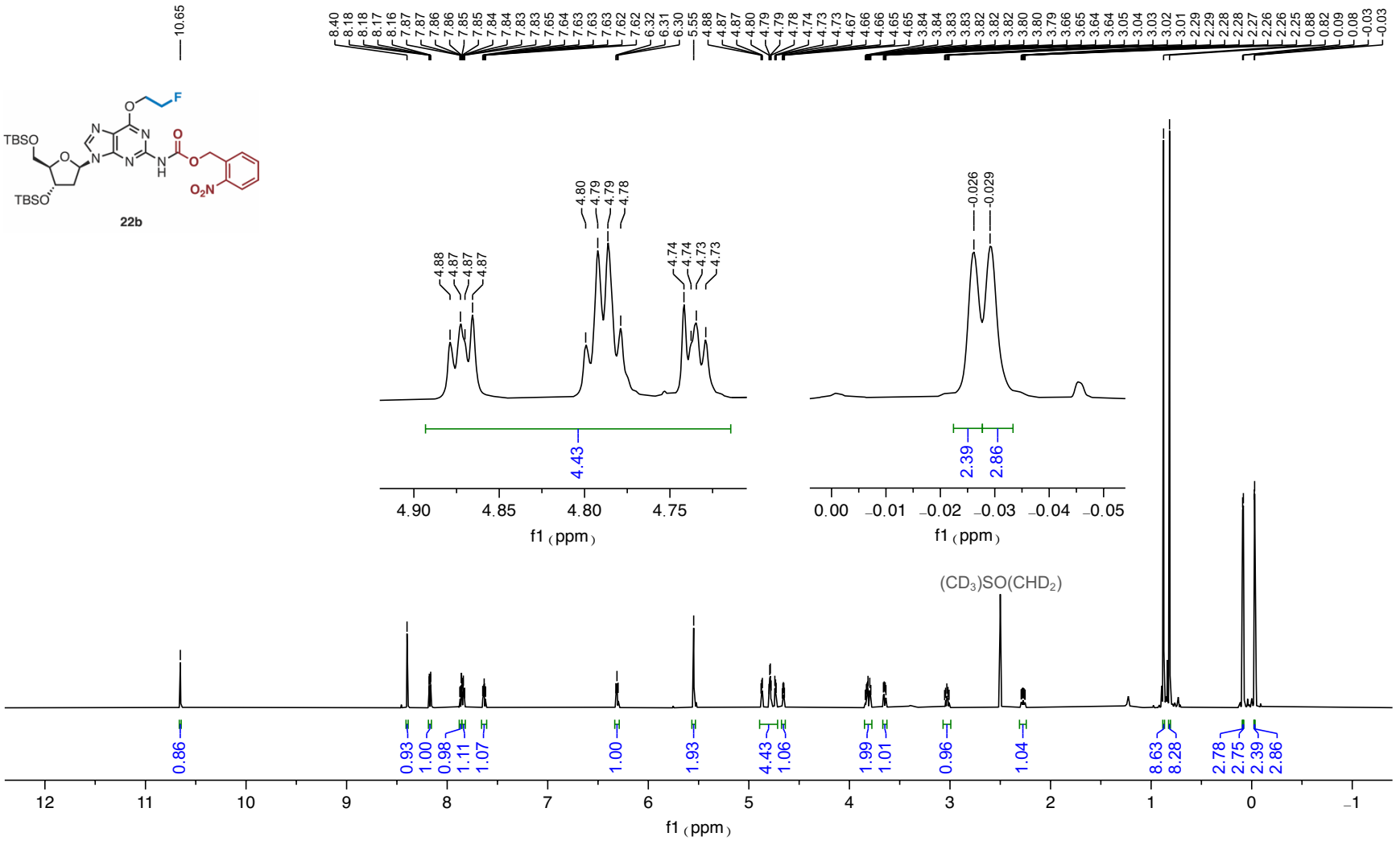


$^{31}P\{^{1}H\}$ qNMR, 202 MHz, CD_3CN (qNMR standard = triphenyl phosphate)

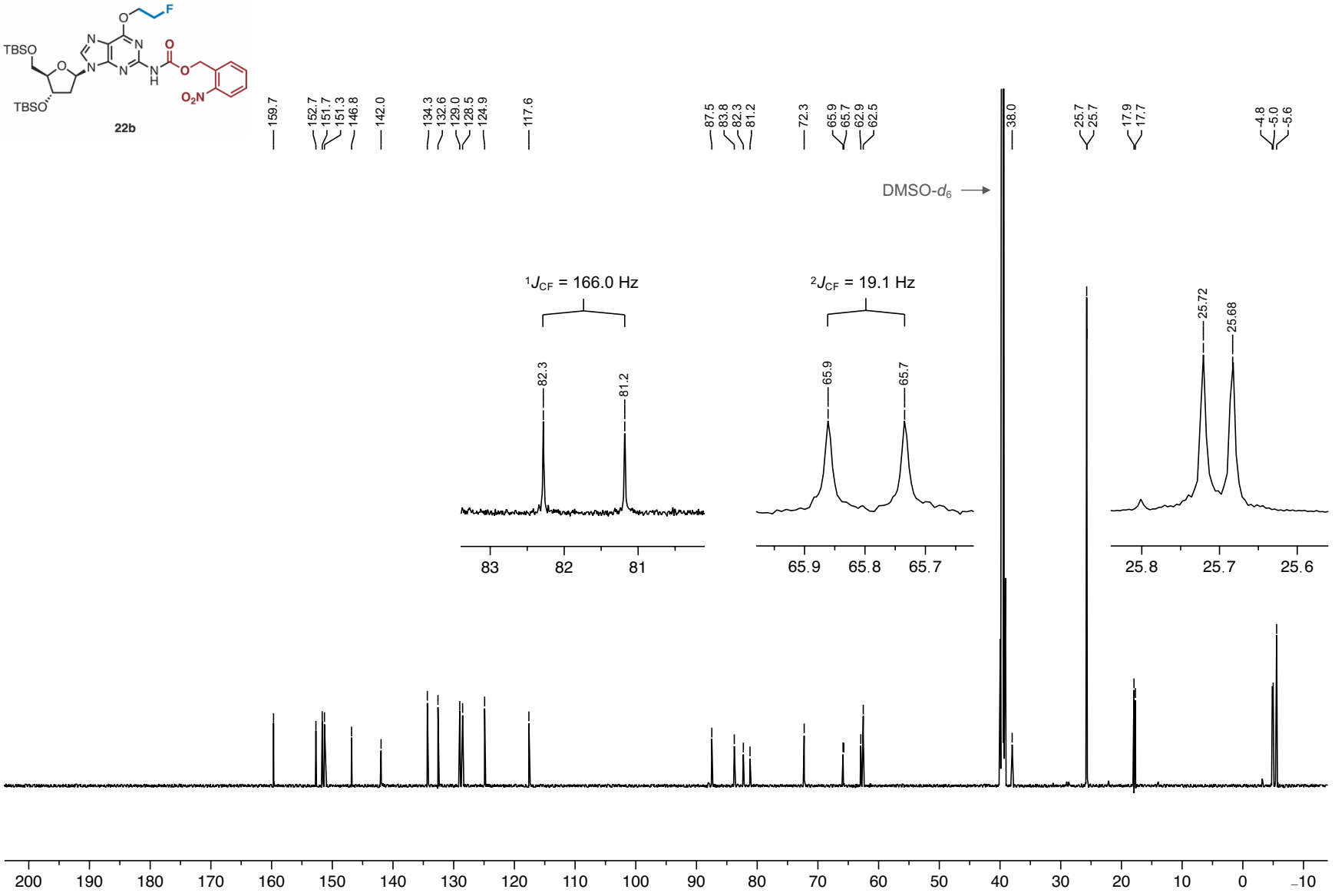
Triphenylphosphate Amount:	1.84 mg
O^6-(2-chloroethyl) Phosphoramidite **25a** Amount:	7.00 mg
O^6-(2-chloroethyl) Phosphoramidite **25a** Purity:	67 wt%



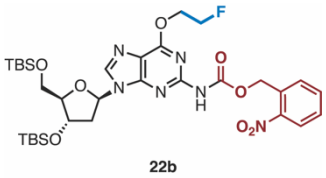
^{1}H NMR, 600 MHz, DMSO-d_6



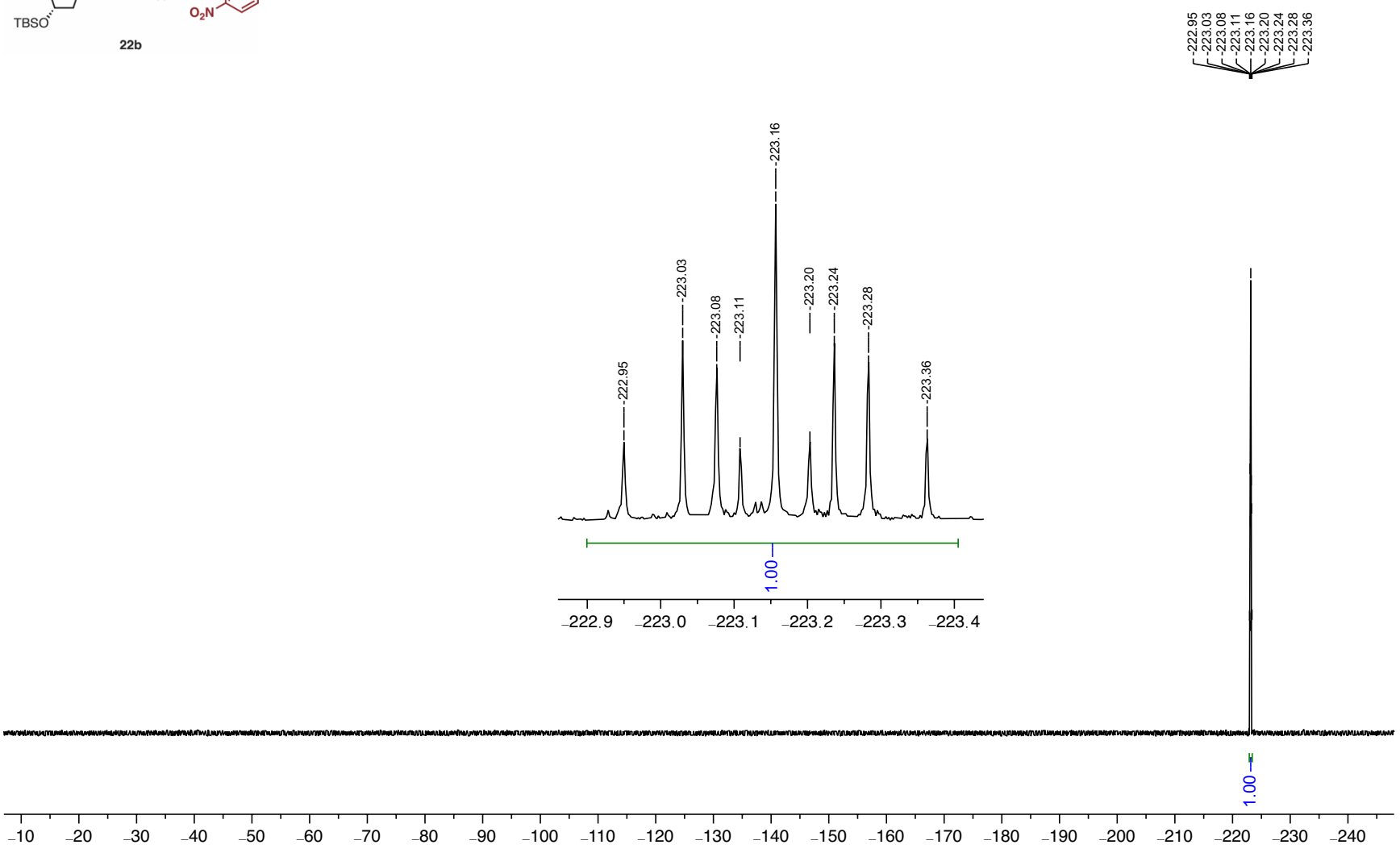
$^{13}C\{^{1}H\}$ NMR, 151 MHz, DMSO-d_6



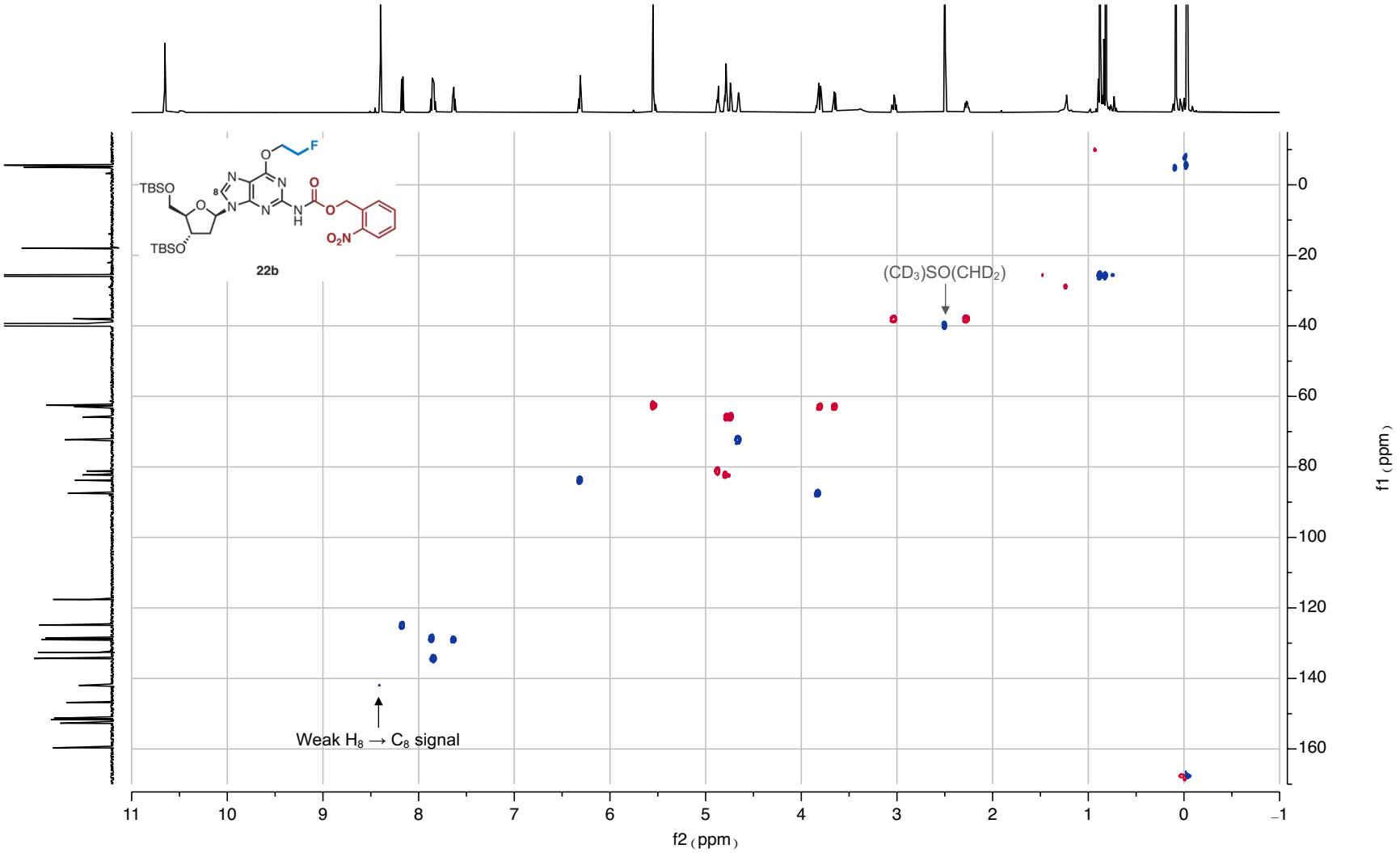
^{19}F NMR, 376 MHz, DMSO-d_6



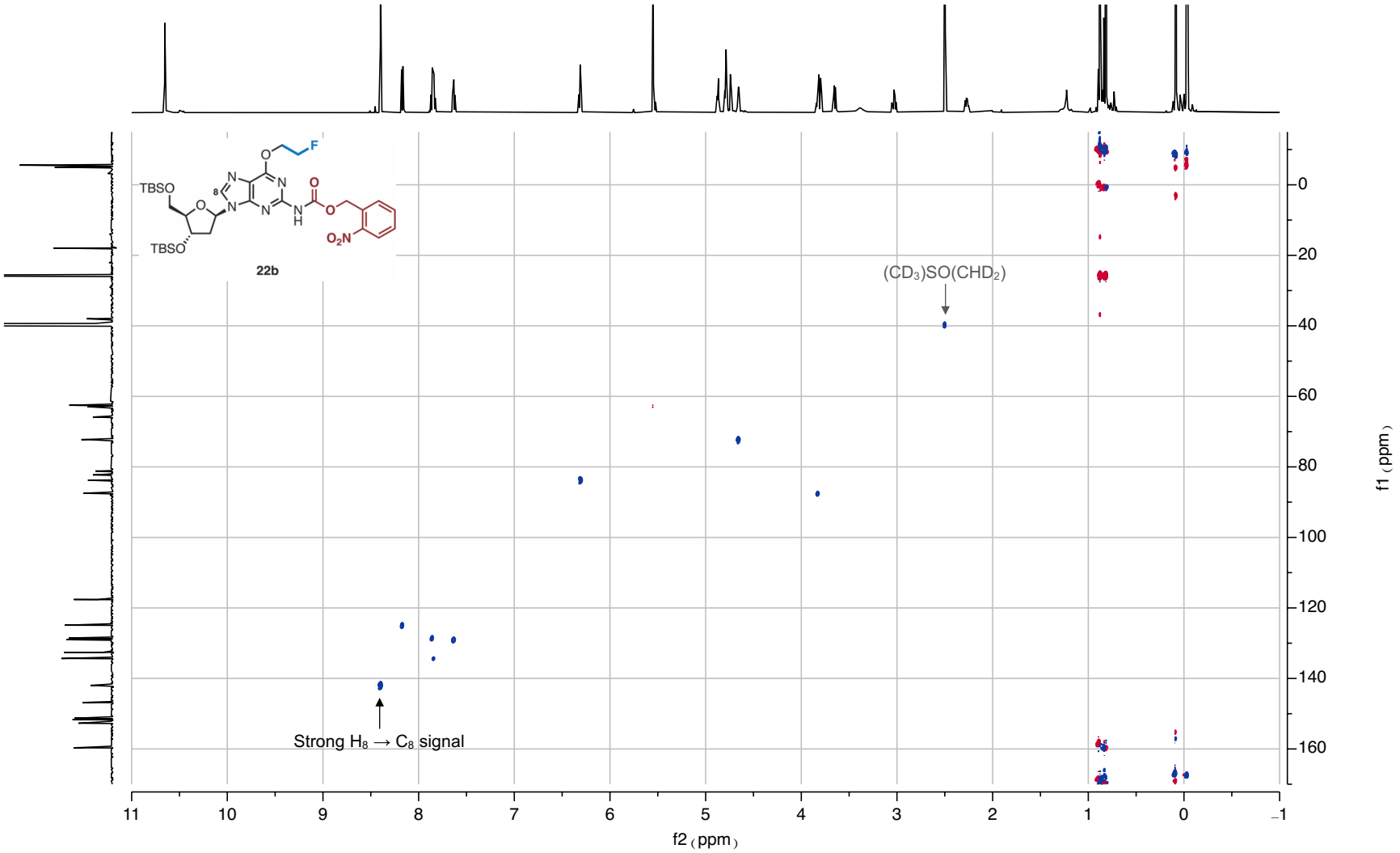
22b



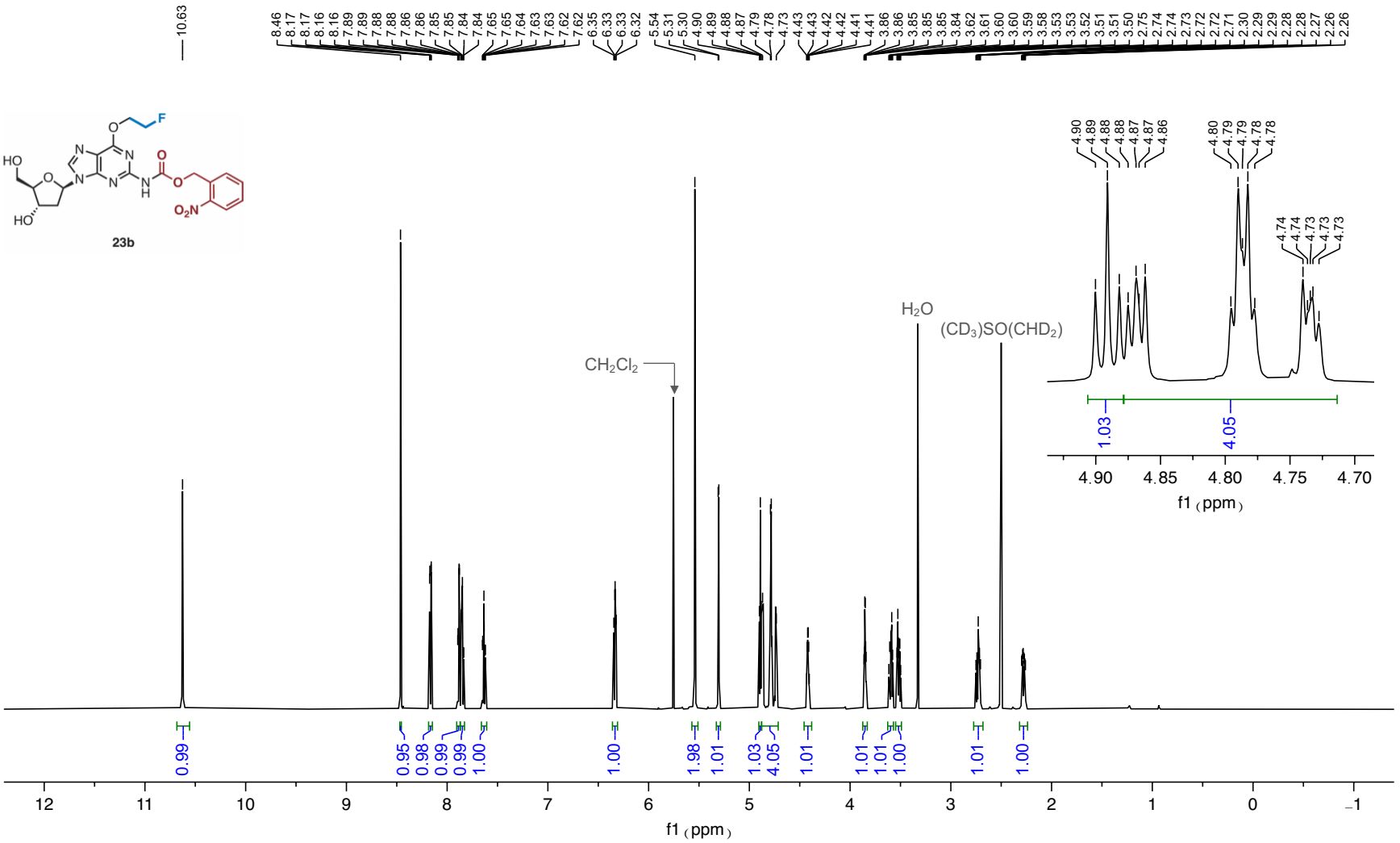
^{1}H–^{13}C HSQC{^{1}H}{^{13}C} NMR, 600/151 MHz, DMSO-d_6, $^{1}J_{CH}$ = 146 Hz, Blue = CH/CH_3, Red = CH_2



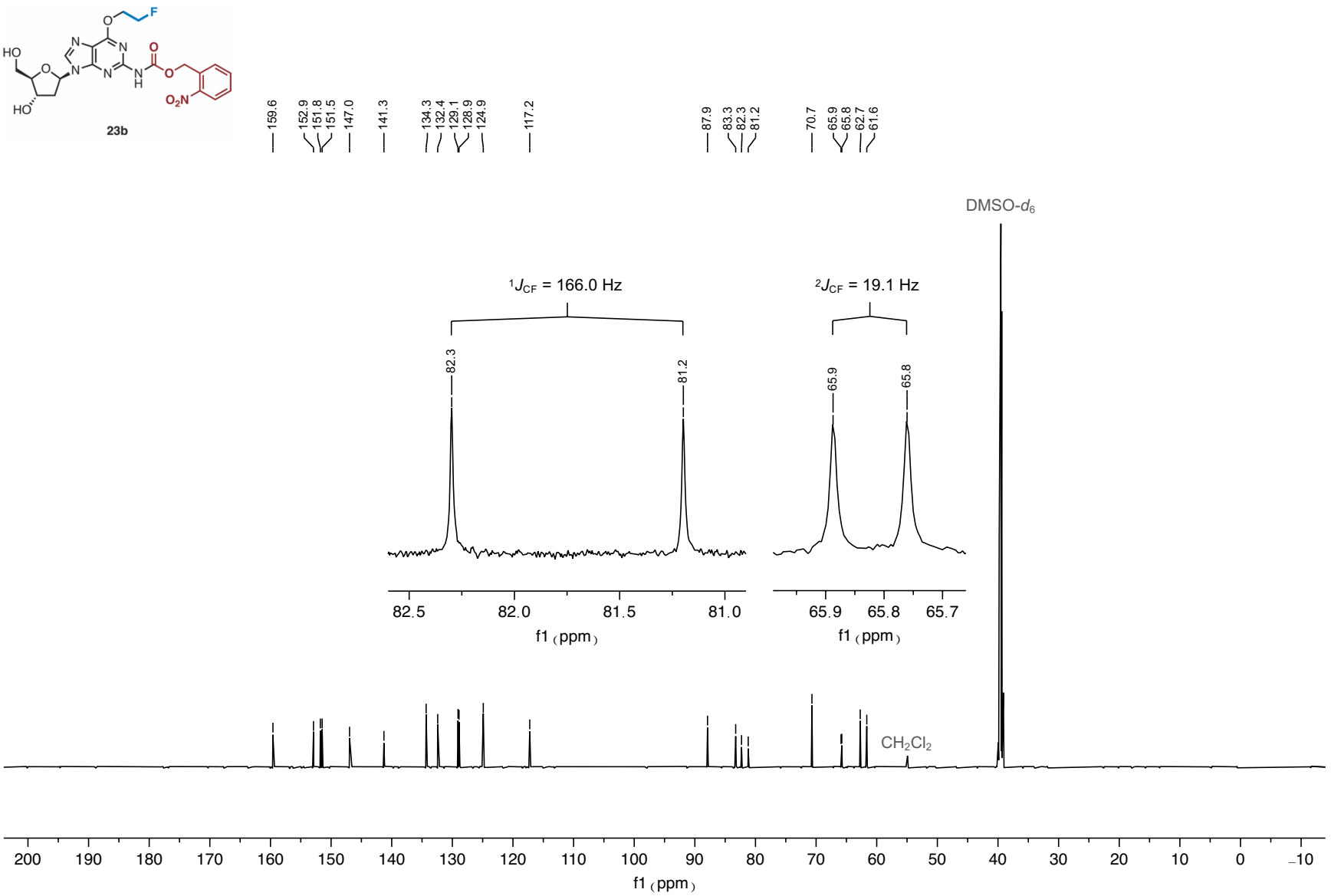
^{1}H–^{13}C HSQC{^{1}H}{^{13}C} NMR, 600 MHz, DMSO-d_6, $^{1}J_{CH}$ = 210 Hz, Blue = CH/CH_3, Red = CH_2



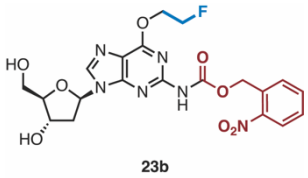
^{1}H NMR, 600 MHz, DMSO-d_6



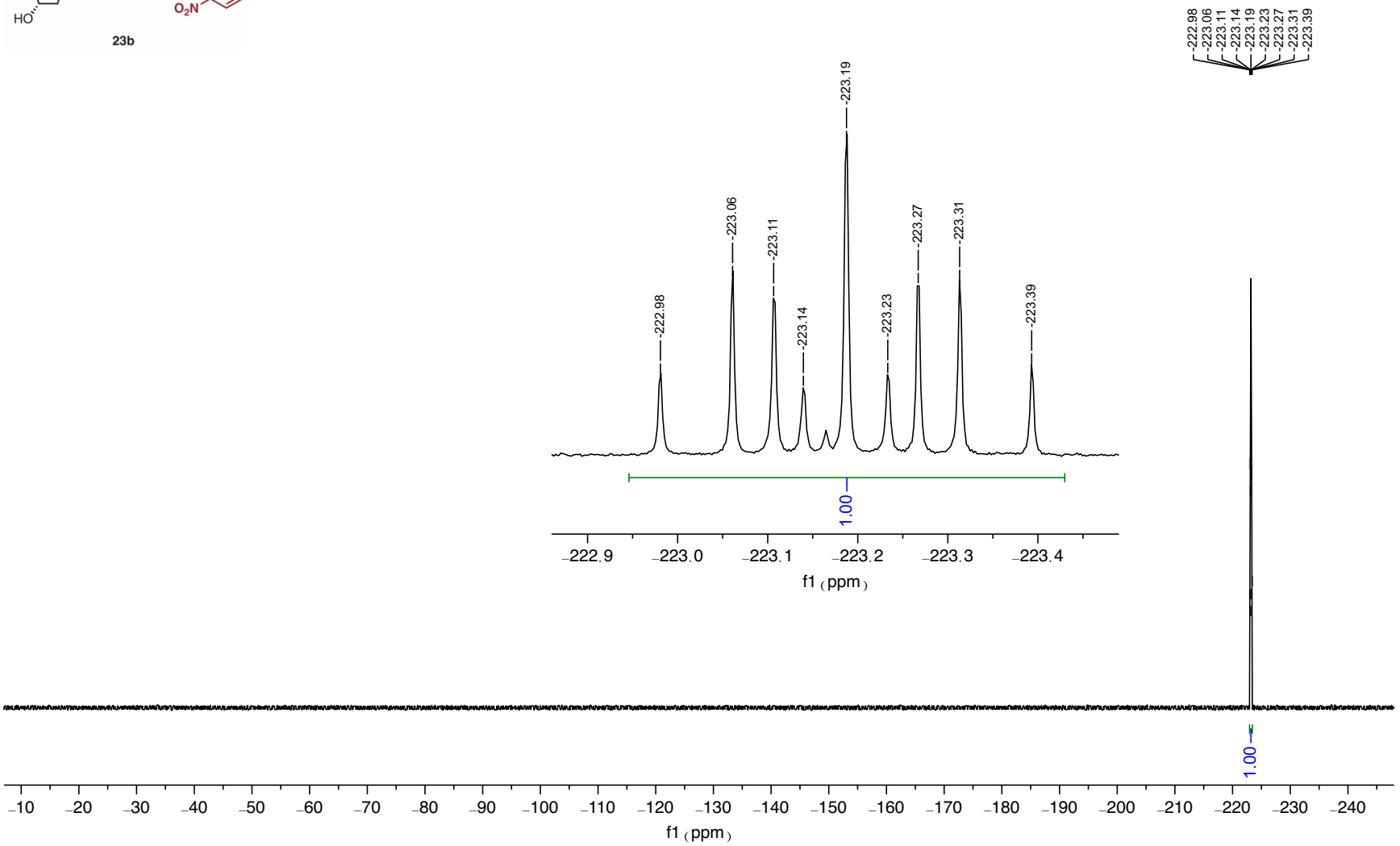
$^{13}C\{^{1}H\}$ NMR, 151 MHz, DMSO-d_6



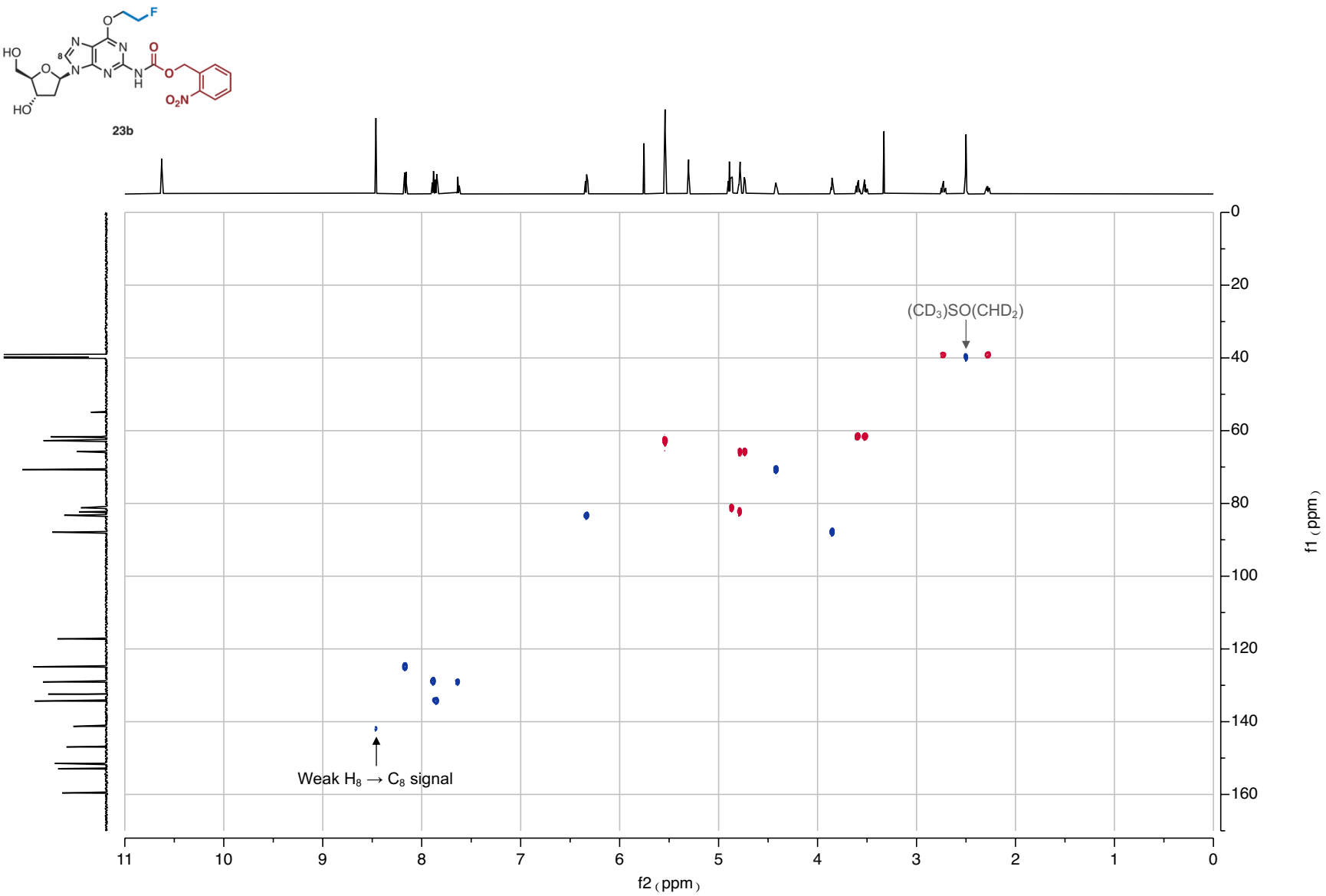
^{19}F NMR, 376 MHz, DMSO-d_6



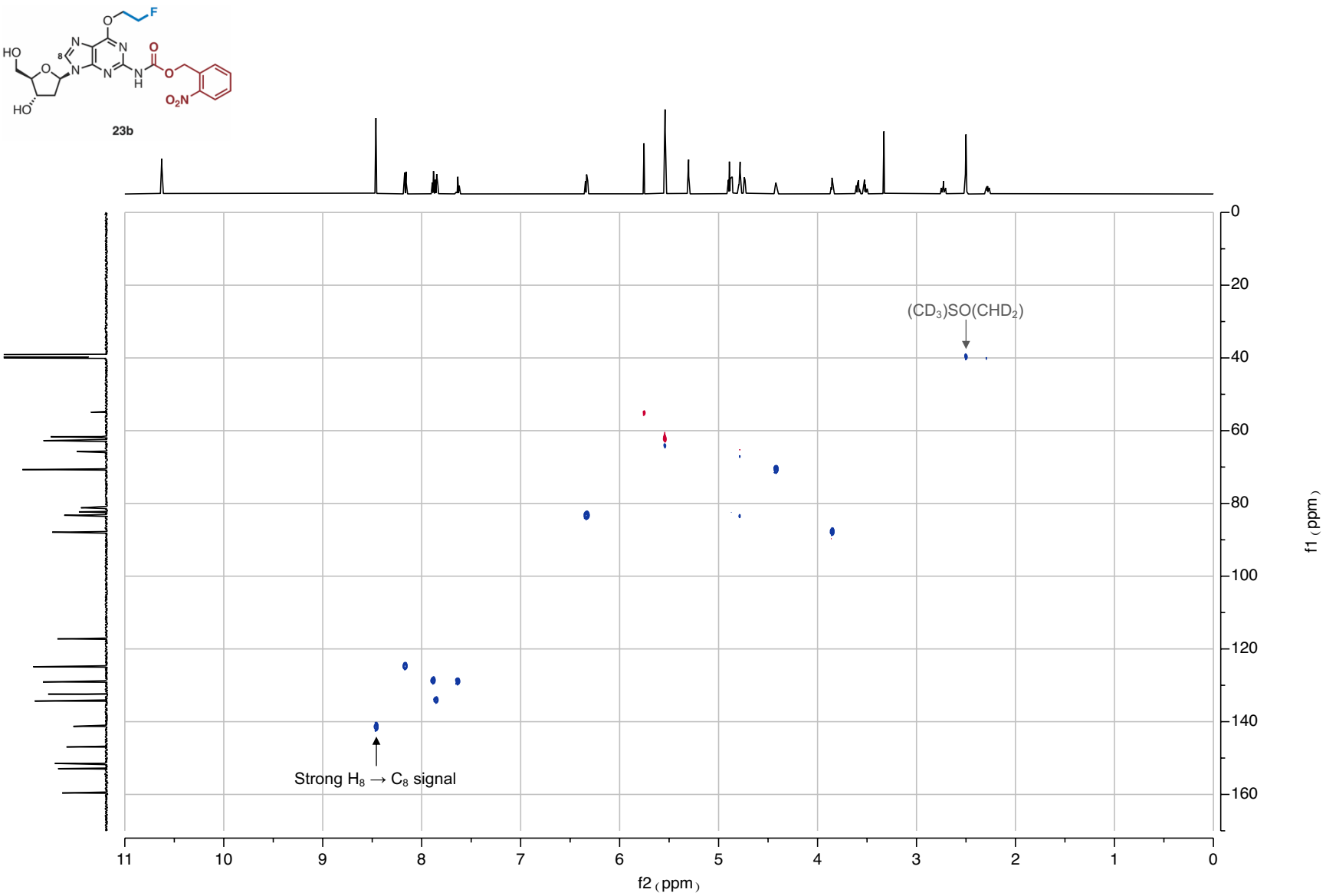
23b



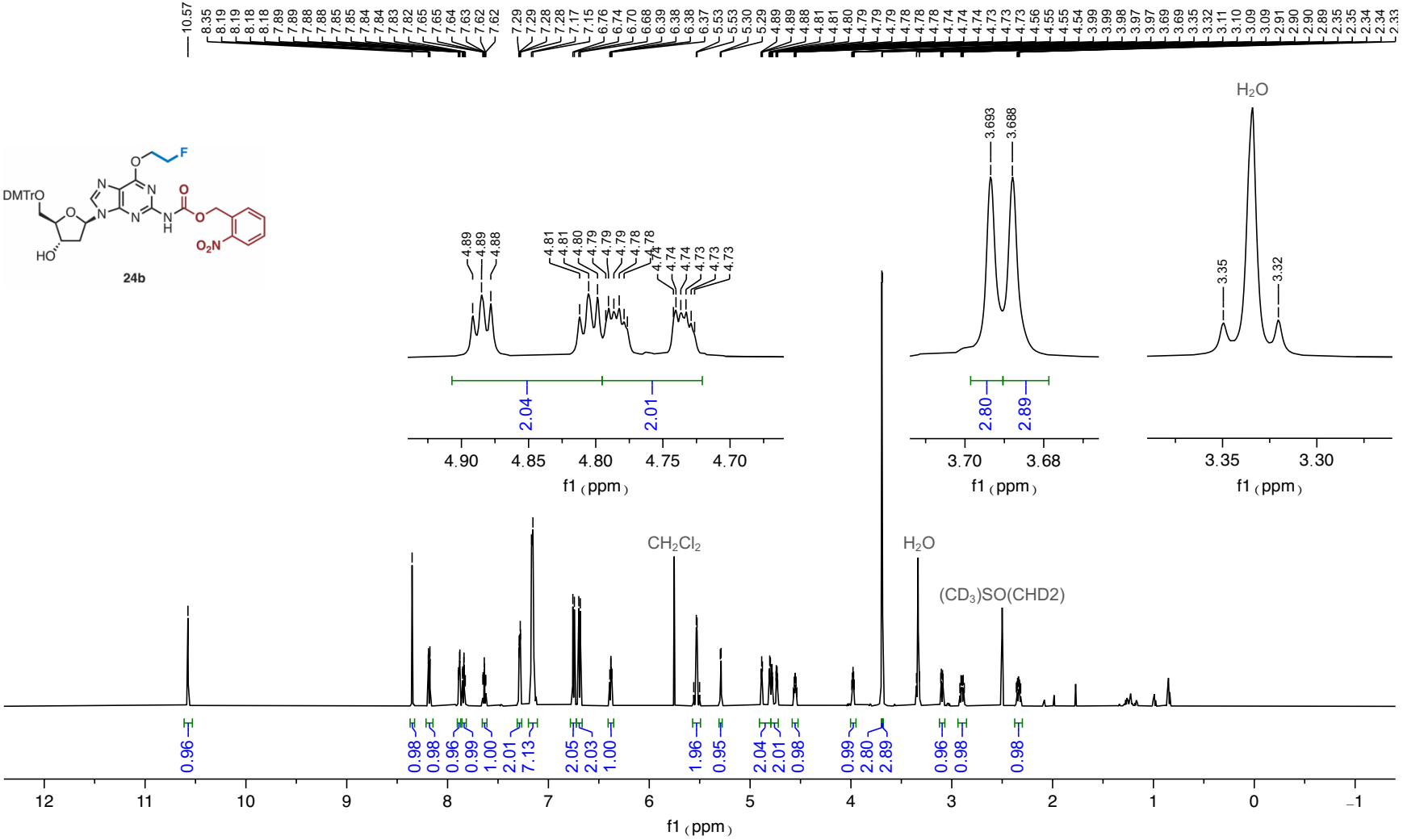
^{1}H–^{13}C HSQC{^{1}H}{^{13}C} NMR, 600/151 MHz, DMSO-d_6, $^{1}J_{CH}$ = 146 Hz, Blue = CH/CH_3, Red = CH_2



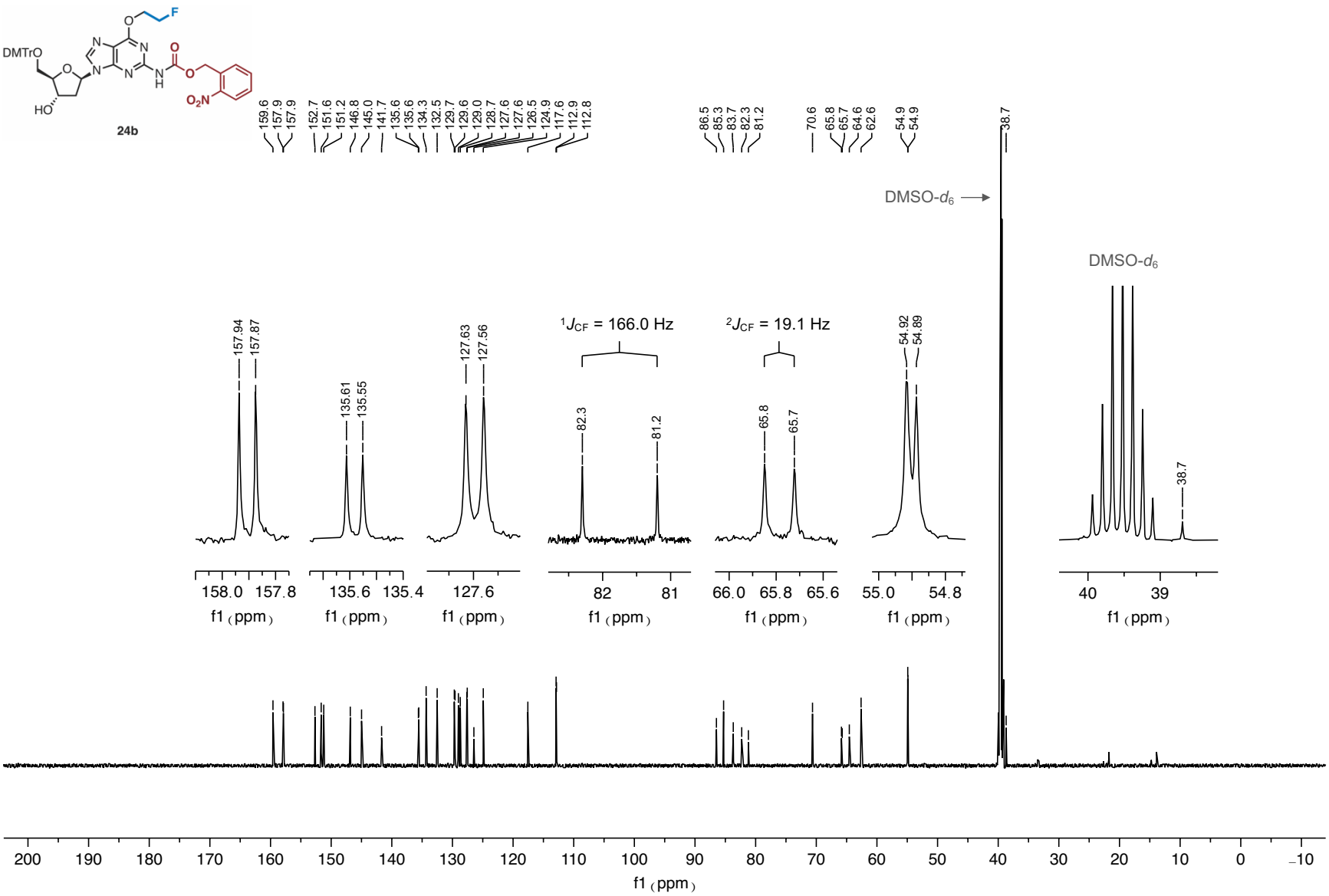
^{1}H–^{13}C HSQC{^{1}H}{^{13}C} NMR, 600/151 MHz, DMSO-d_6, $^{1}J_{CH}$ = 210 Hz, Blue = CH/CH_3, Red = CH_2



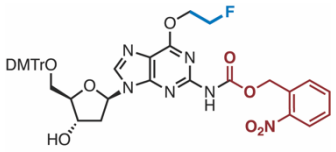
^{1}H NMR, 600 MHz, DMSO-d_6



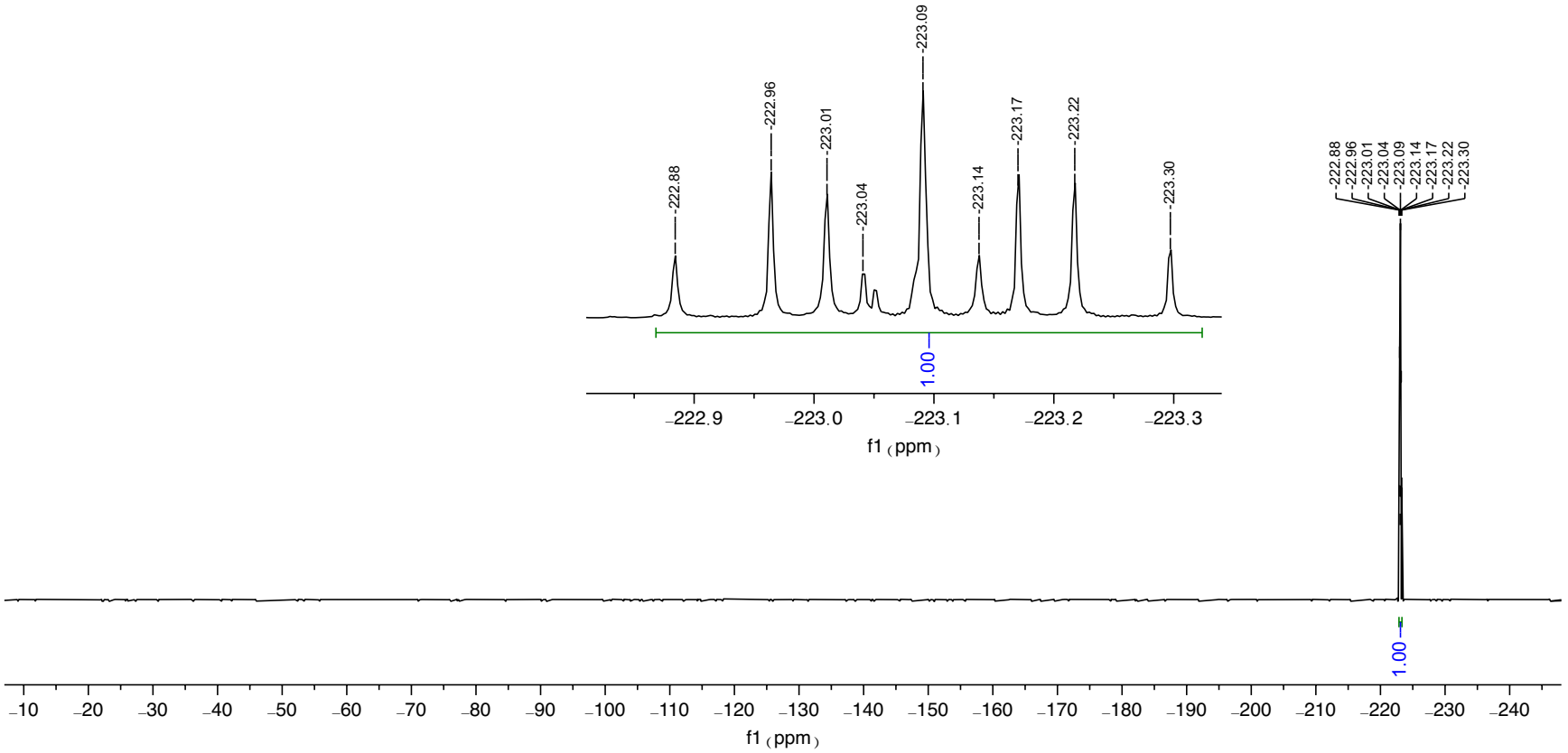
$^{13}C\{^{1}H\}$ NMR, 151 MHz, DMSO-d_6



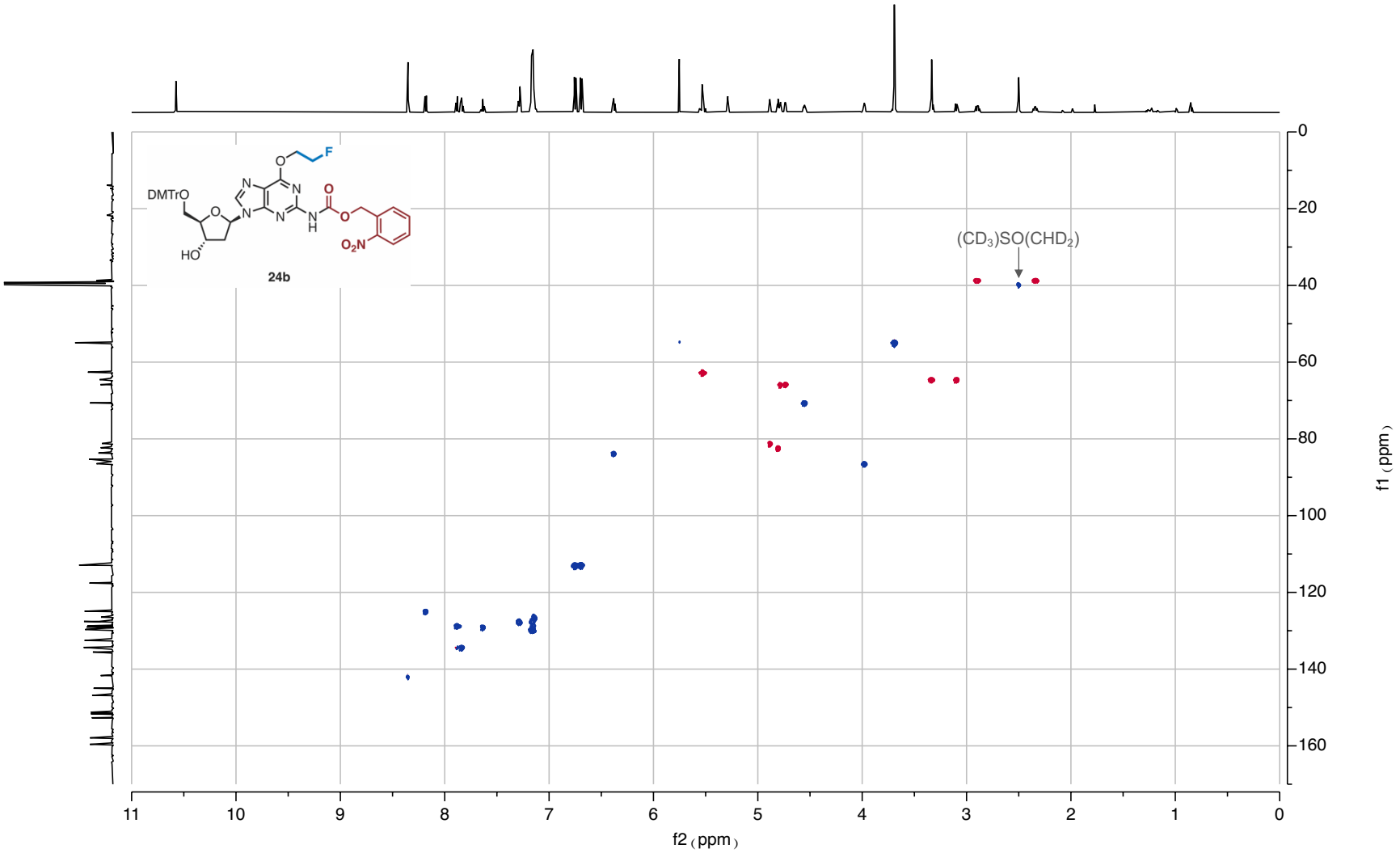
^{19}F NMR, 376 MHz, DMSO-d_6



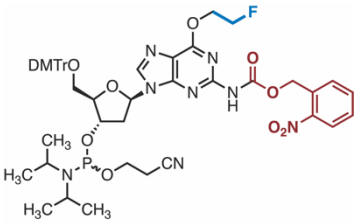
24b



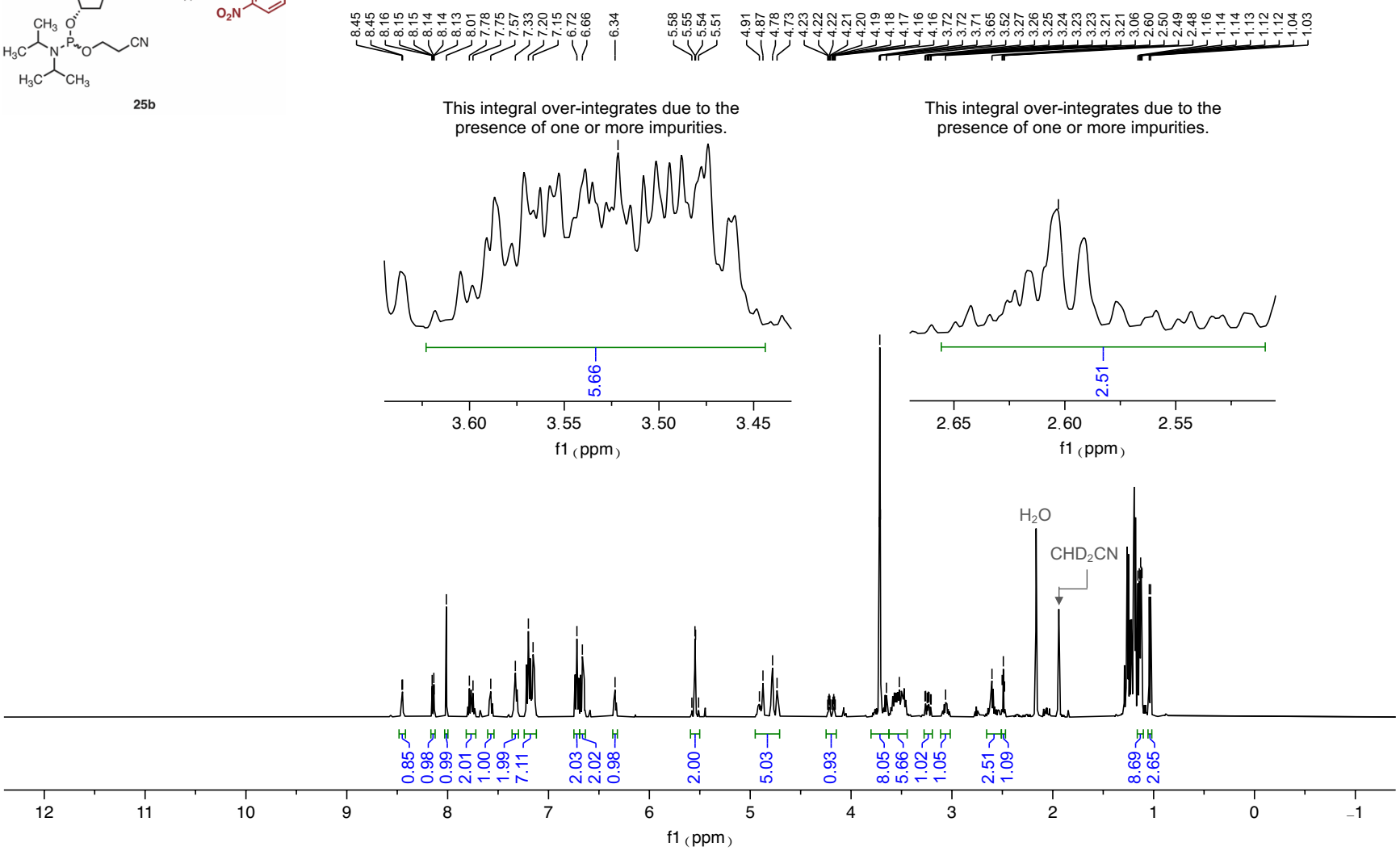
^{1}H–^{13}C HSQC{^{1}H}{^{13}C} NMR, 600/151 MHz, DMSO-d_6, $^{1}J_{CH}$ = 146 Hz, Blue = CH/CH_3, Red = CH_2



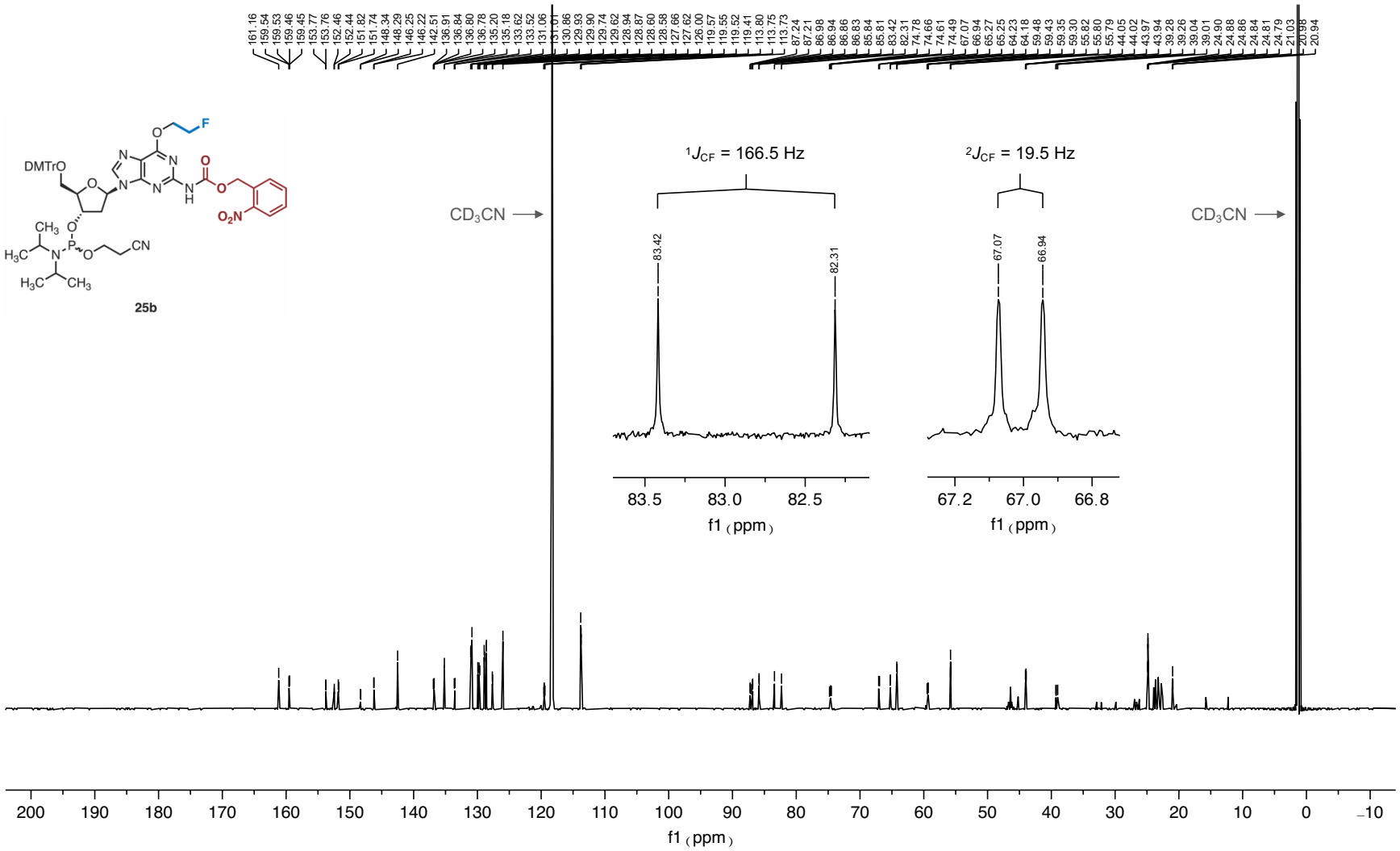
^{1}H NMR, 500 MHz, CD_3CN



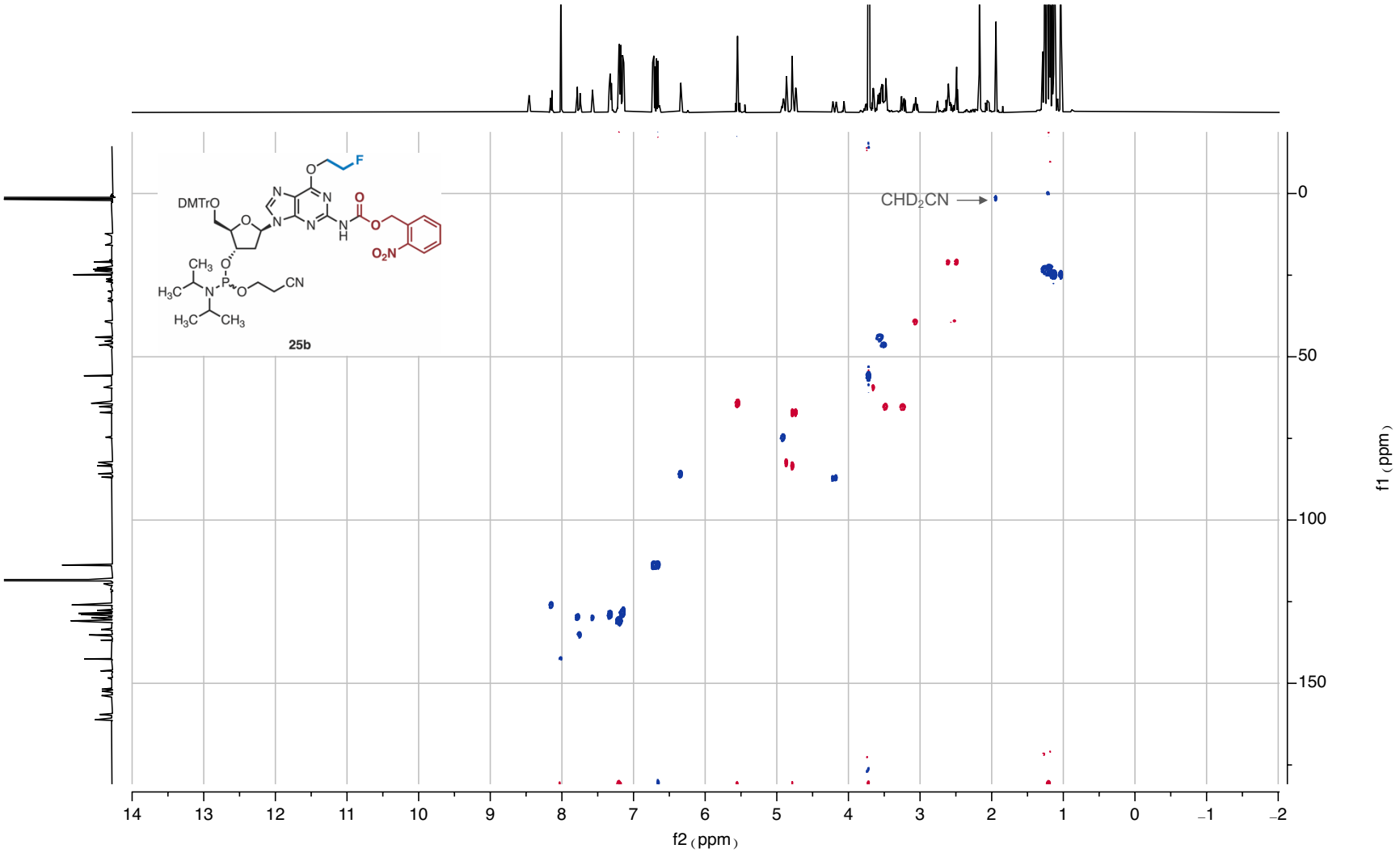
25b



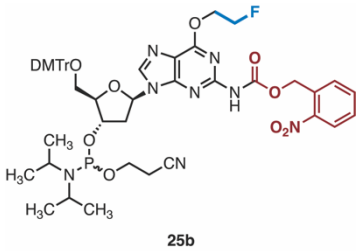
$^{13}C\{^{1}H\}$ NMR, 151 MHz, CD_3CN



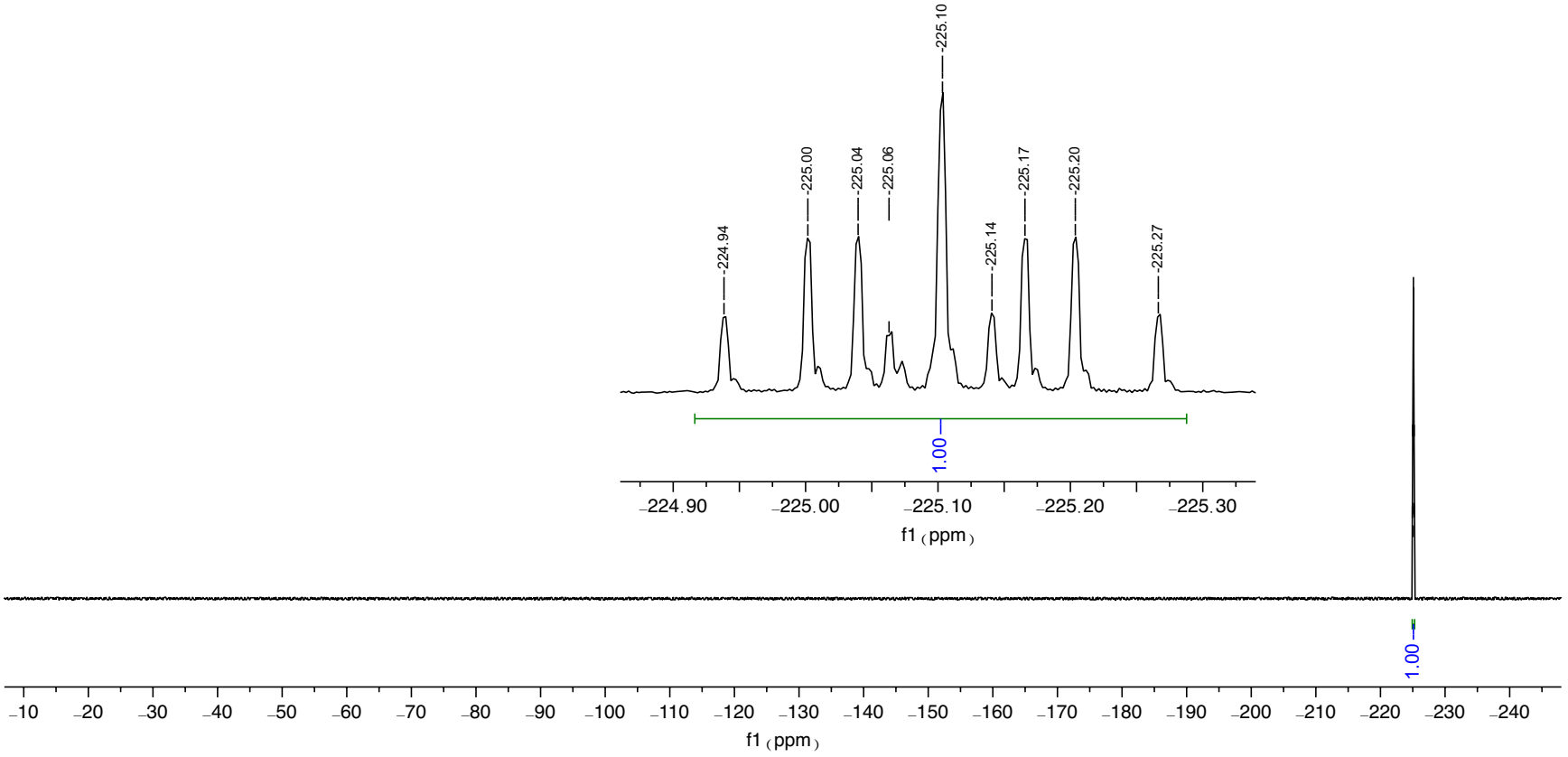
^{1}H–^{13}C HSQC{^{1}H}{^{13}C} NMR, 600/151 MHz, CD_3CN, $^{1}J_{CH}$ = 146 Hz, Blue = CH/CH_3, Red = CH_2



^{19}F NMR, 471 MHz, CD_3CN

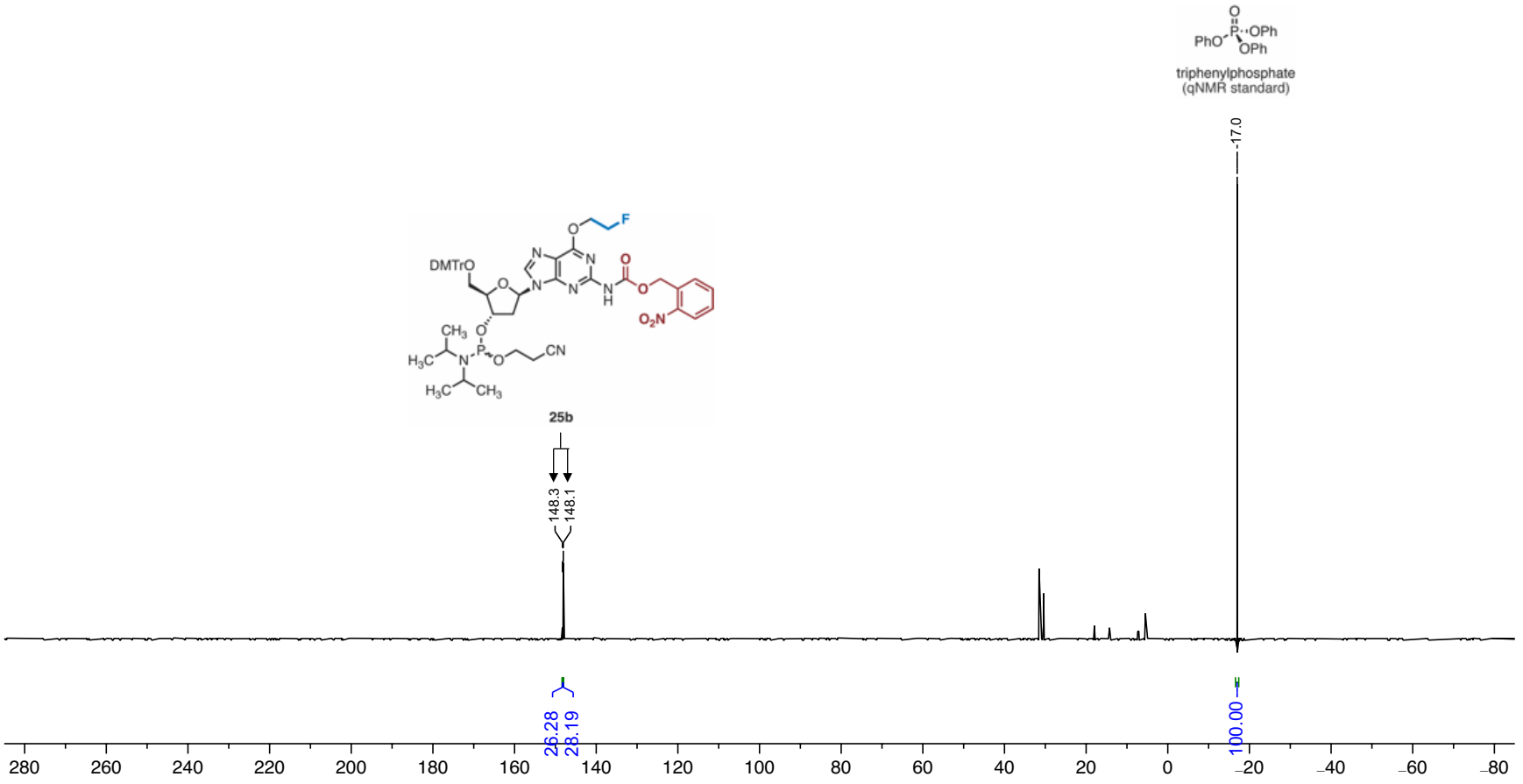


25b

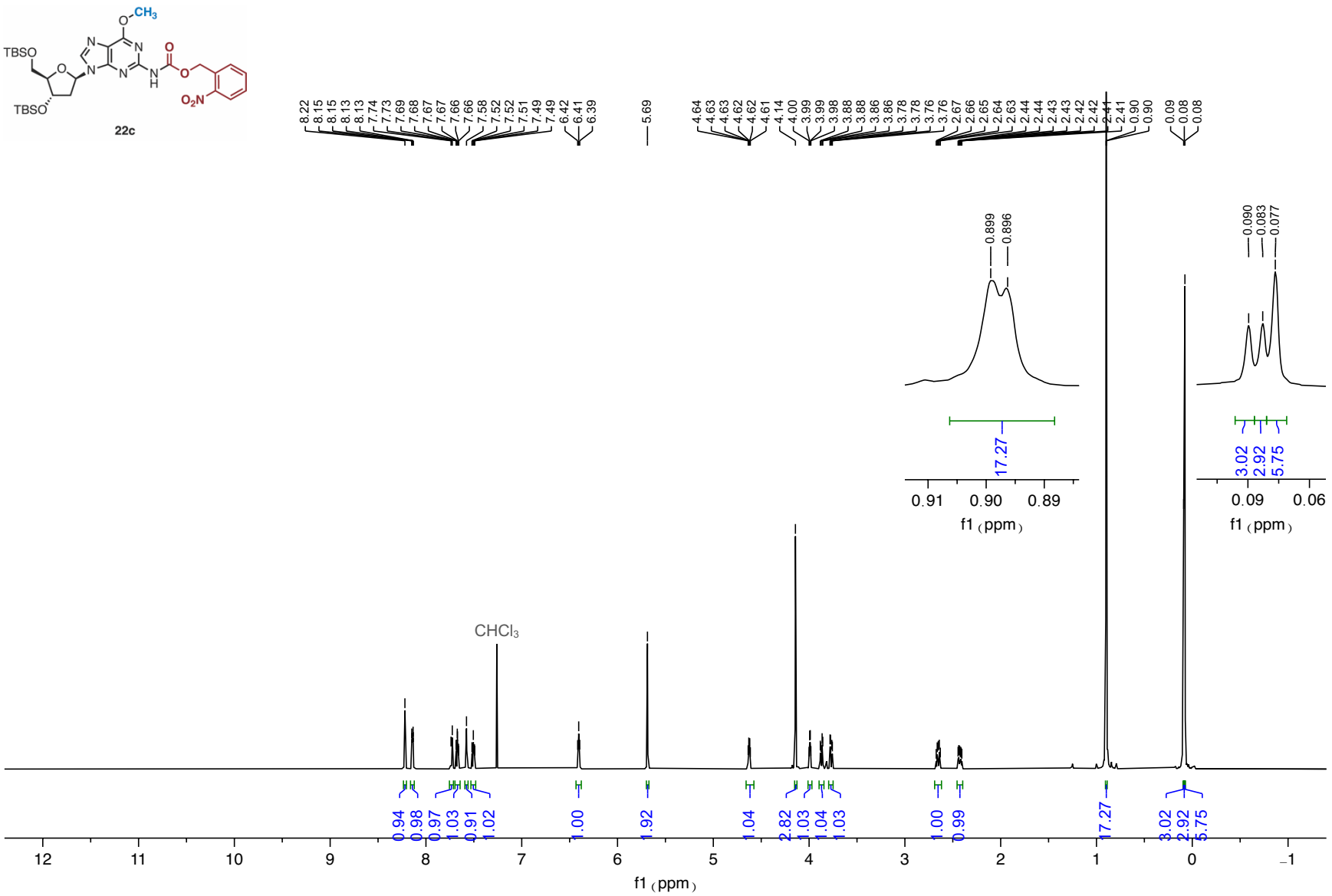


$^{31}P\{^{1}H\}$ qNMR, 202 MHz, CD_3CN (qNMR standard = triphenyl phosphate)

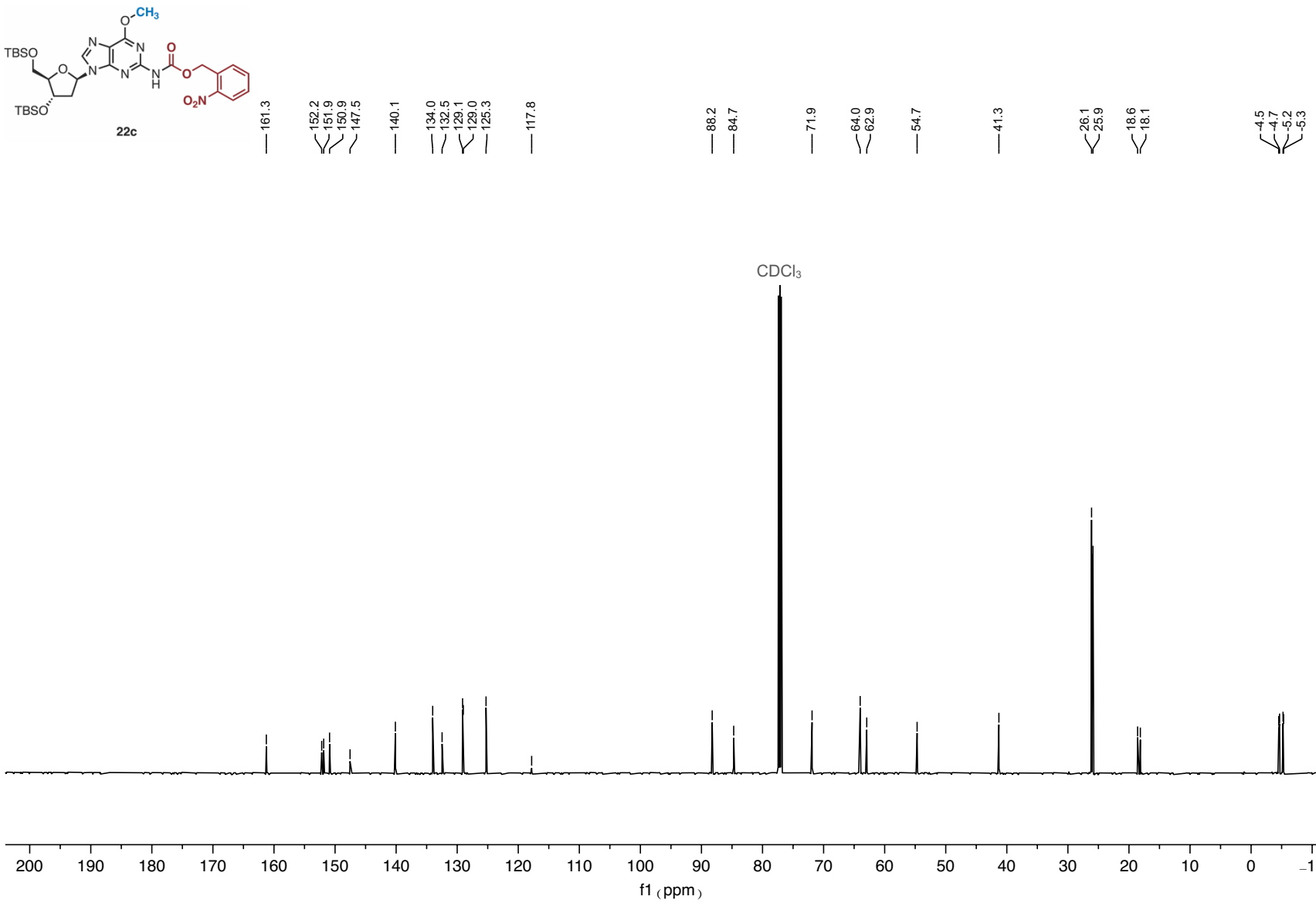
Triphenylphosphate Amount:	2.67 mg
O^6-(2-fluoroethyl) Phosphoramidite **25b** Amount:	6.42 mg
O^6-(2-fluoroethyl) Phosphoramidite **25b** Purity:	69 wt%



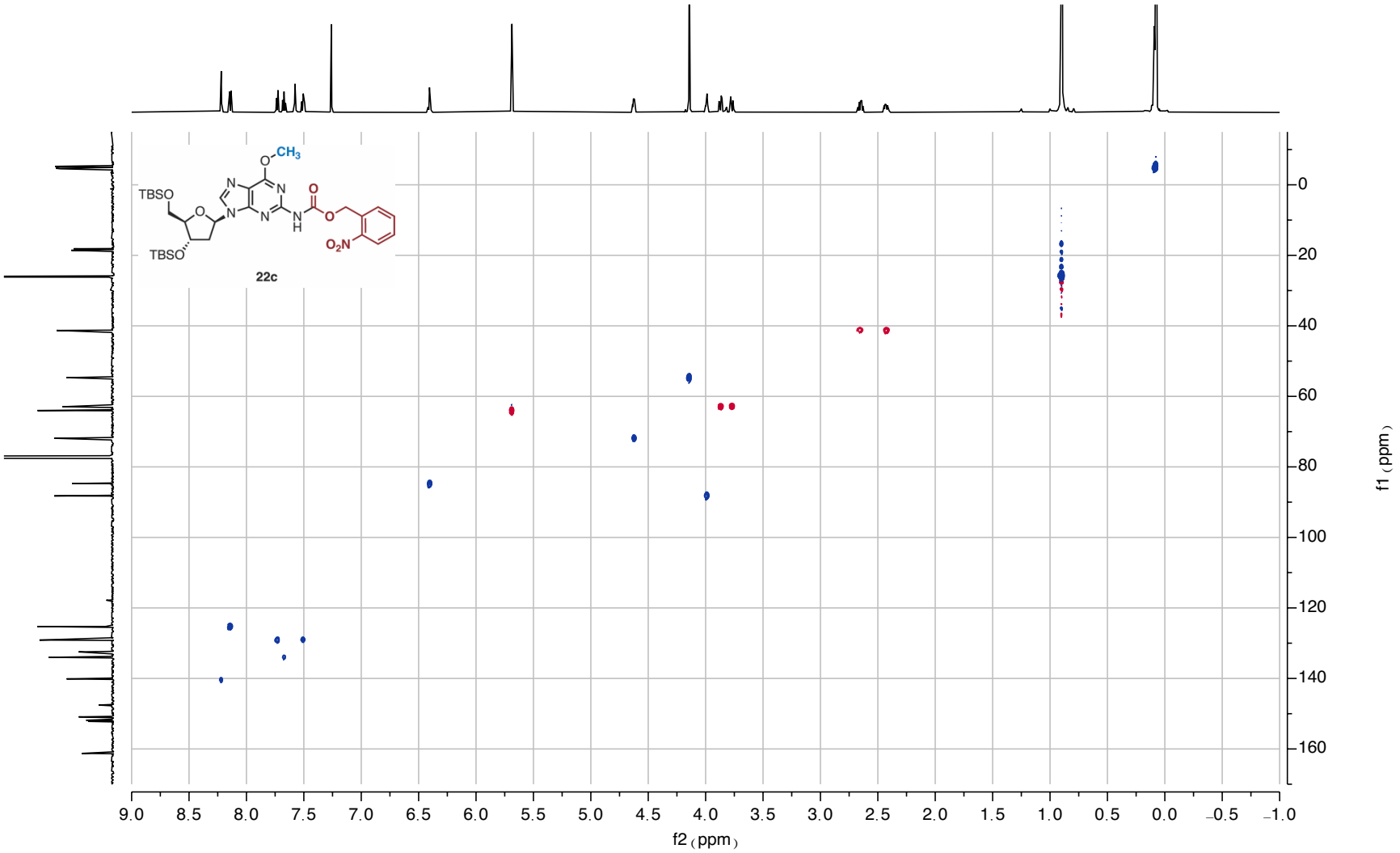
^{1}H NMR, 600 MHz, $CDCl_3$



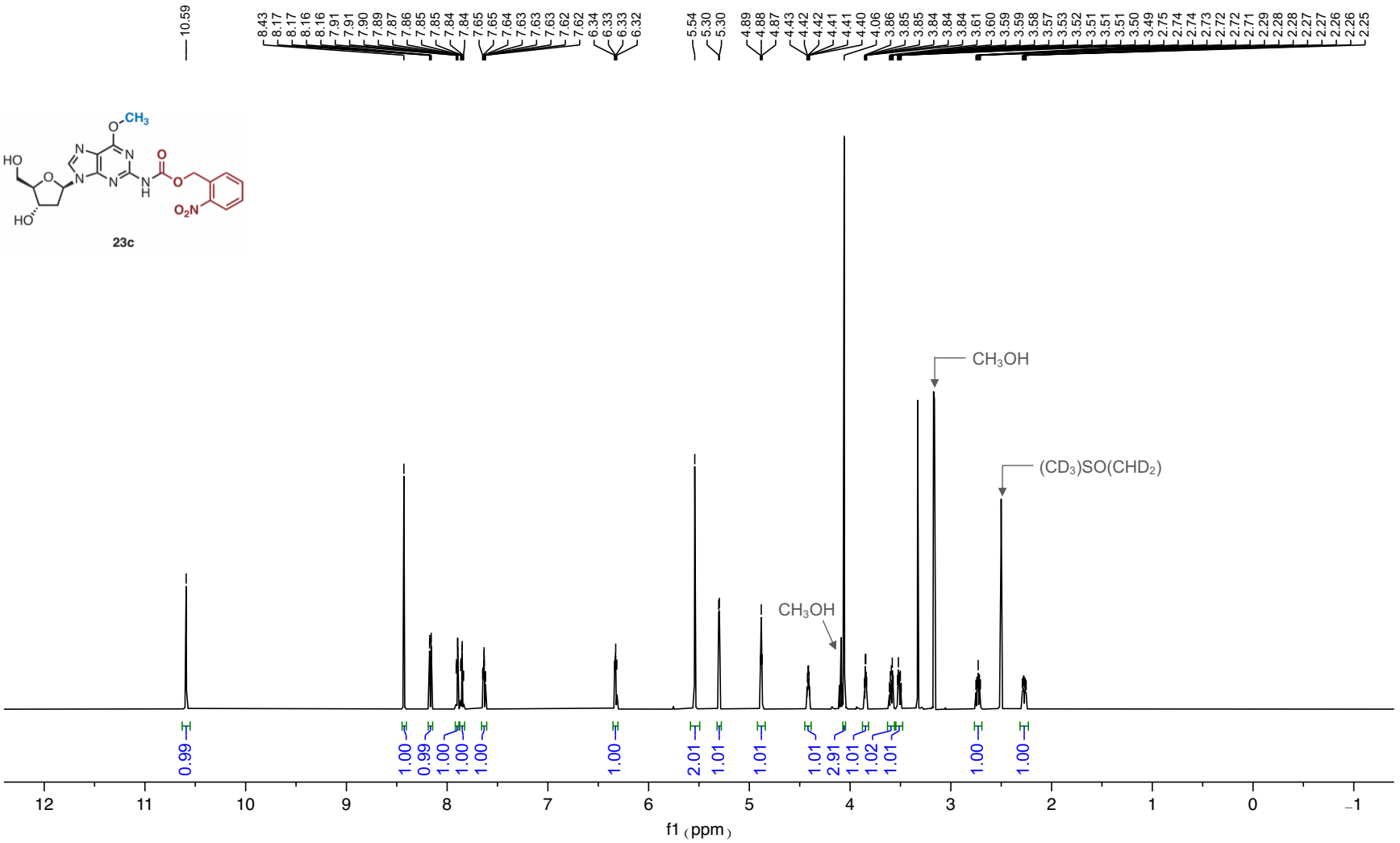
$^{13}C\{^{1}H\}$ NMR, 151 MHz, $CDCl_3$



^{1}H–^{13}C HSQC{^{1}H}{^{13}C} NMR, 600/151 MHz, $CDCl_3$, $^{1}J_{CH}$ = 146 Hz, Blue = CH/CH_3, Red = CH_2



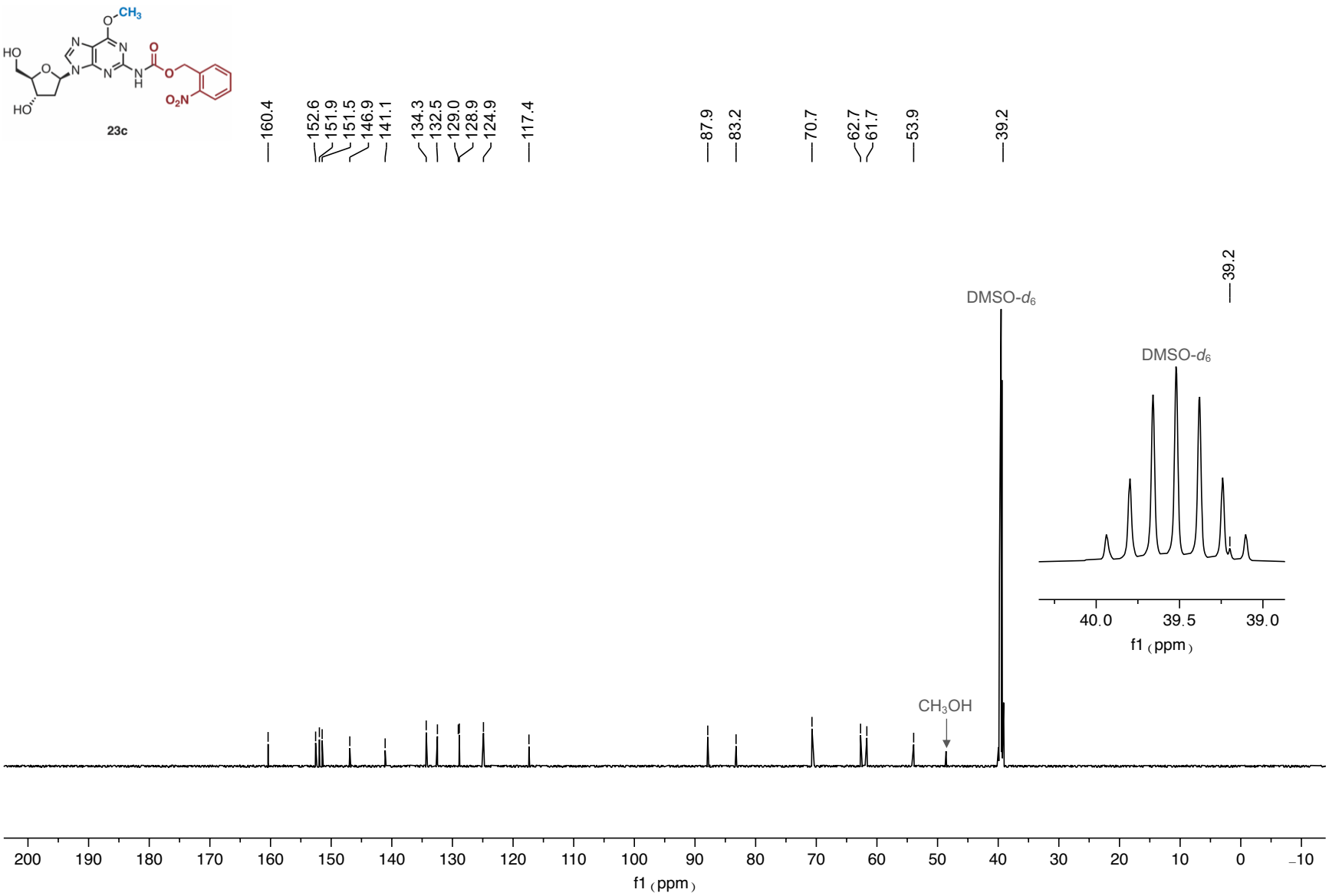
^{1}H NMR, 600 MHz, DMSO-d_6



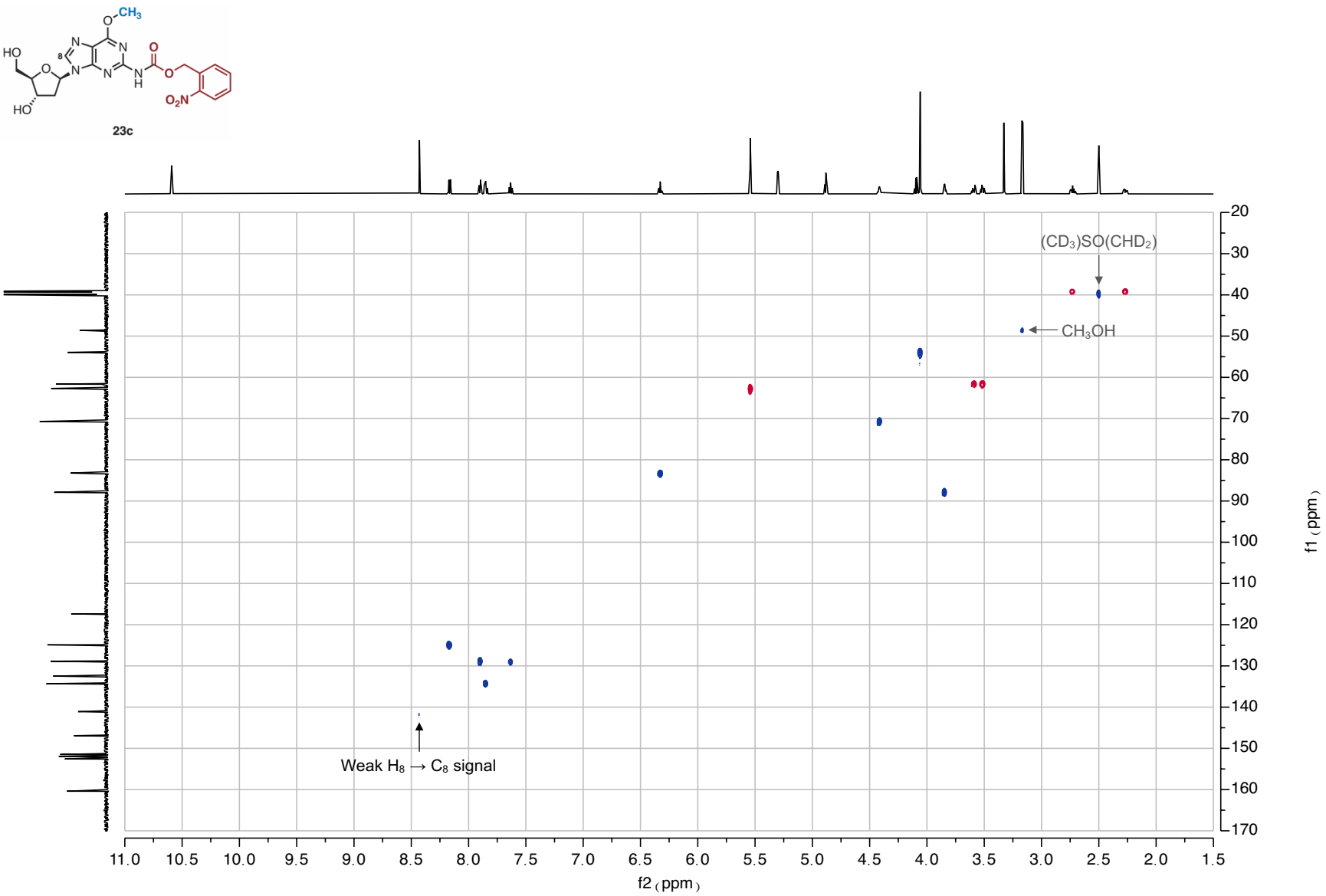
23c

$^{13}C\{^{1}H\}$ NMR, 151 MHz, DMSO-d_6

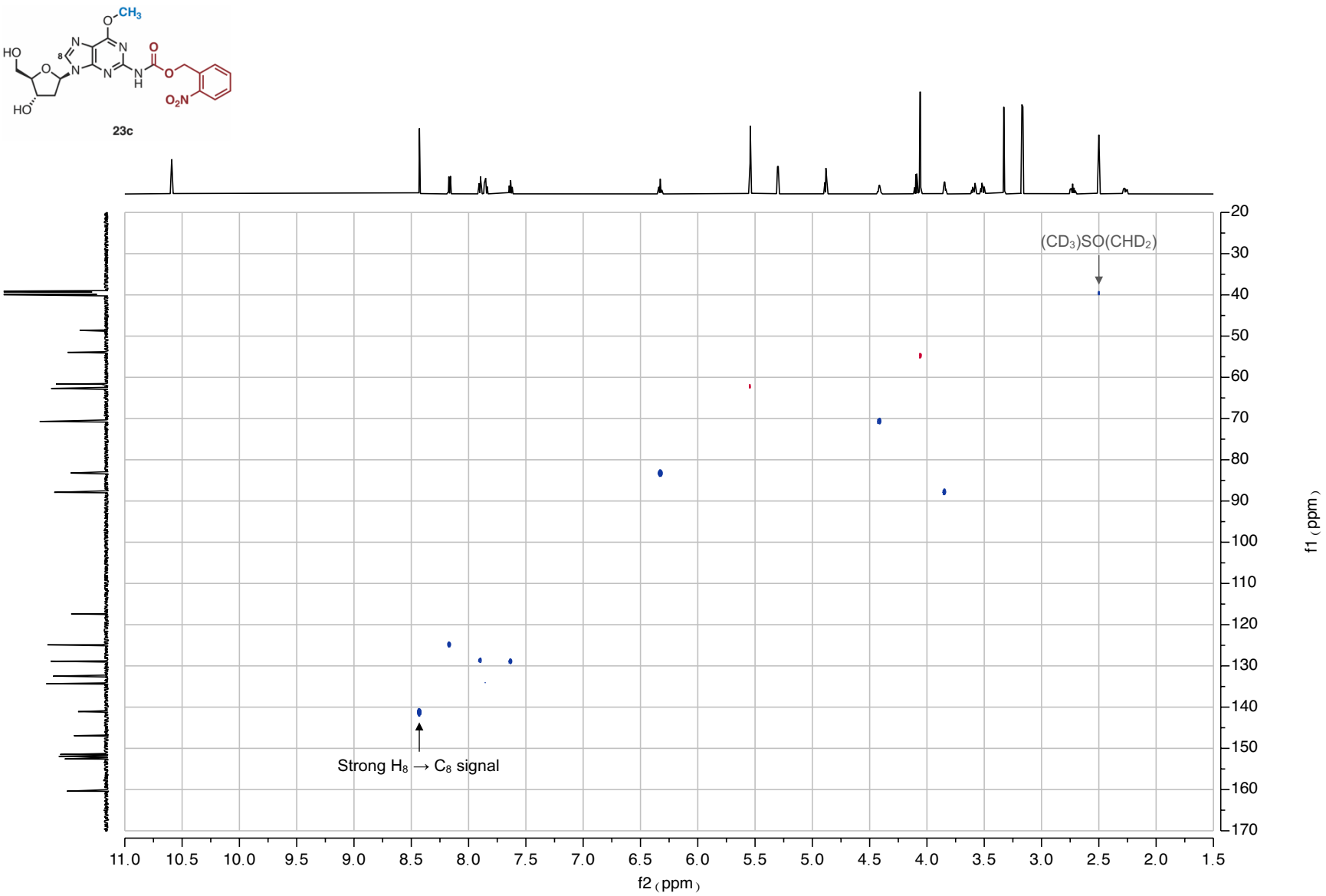
23c



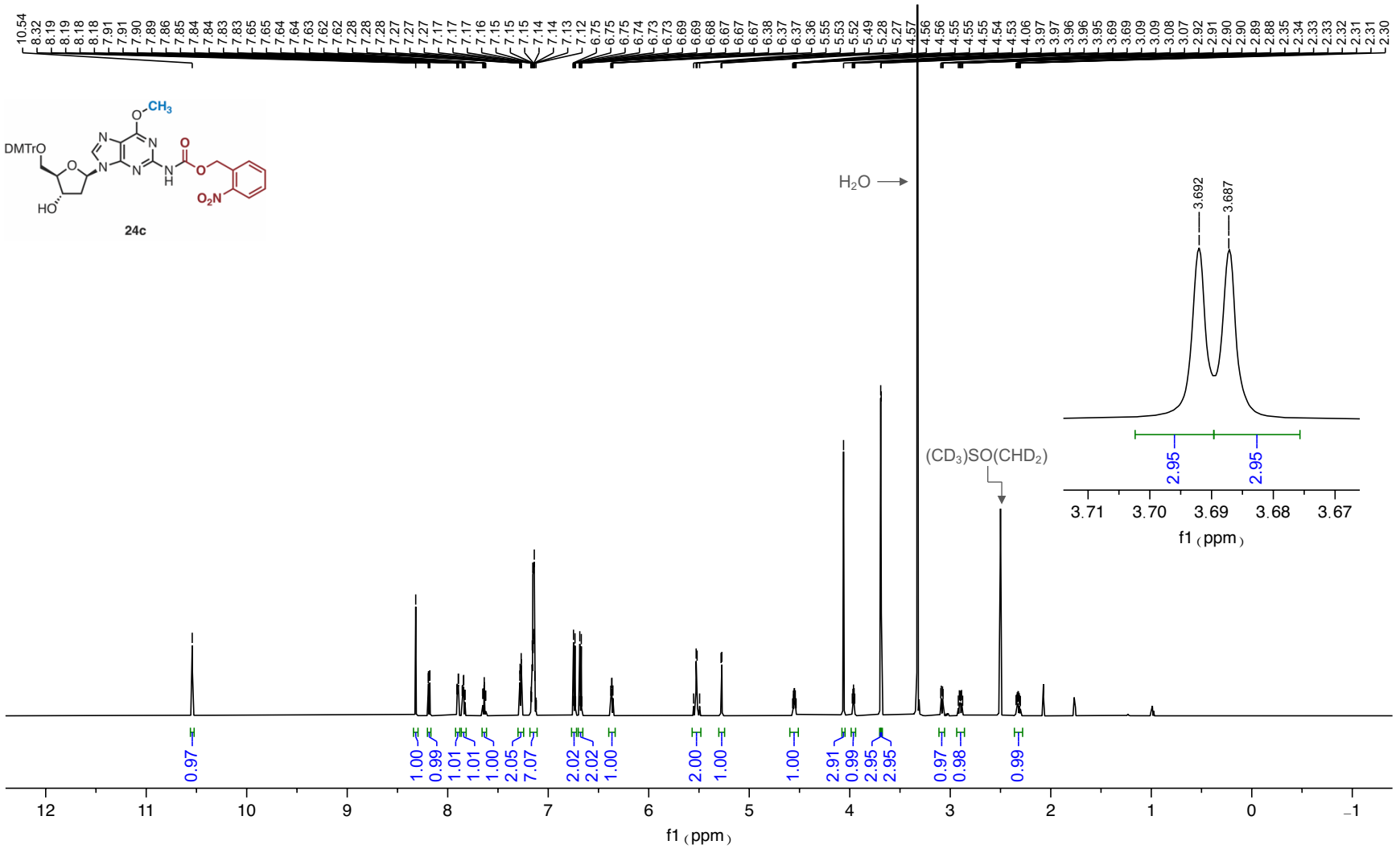
^{1}H–^{13}C HSQC{^{1}H}{^{13}C} NMR, 600/151 MHz, DMSO-d_6, $^{1}J_{CH}$ = 146 Hz, Blue = CH/CH_3, Red = CH_2



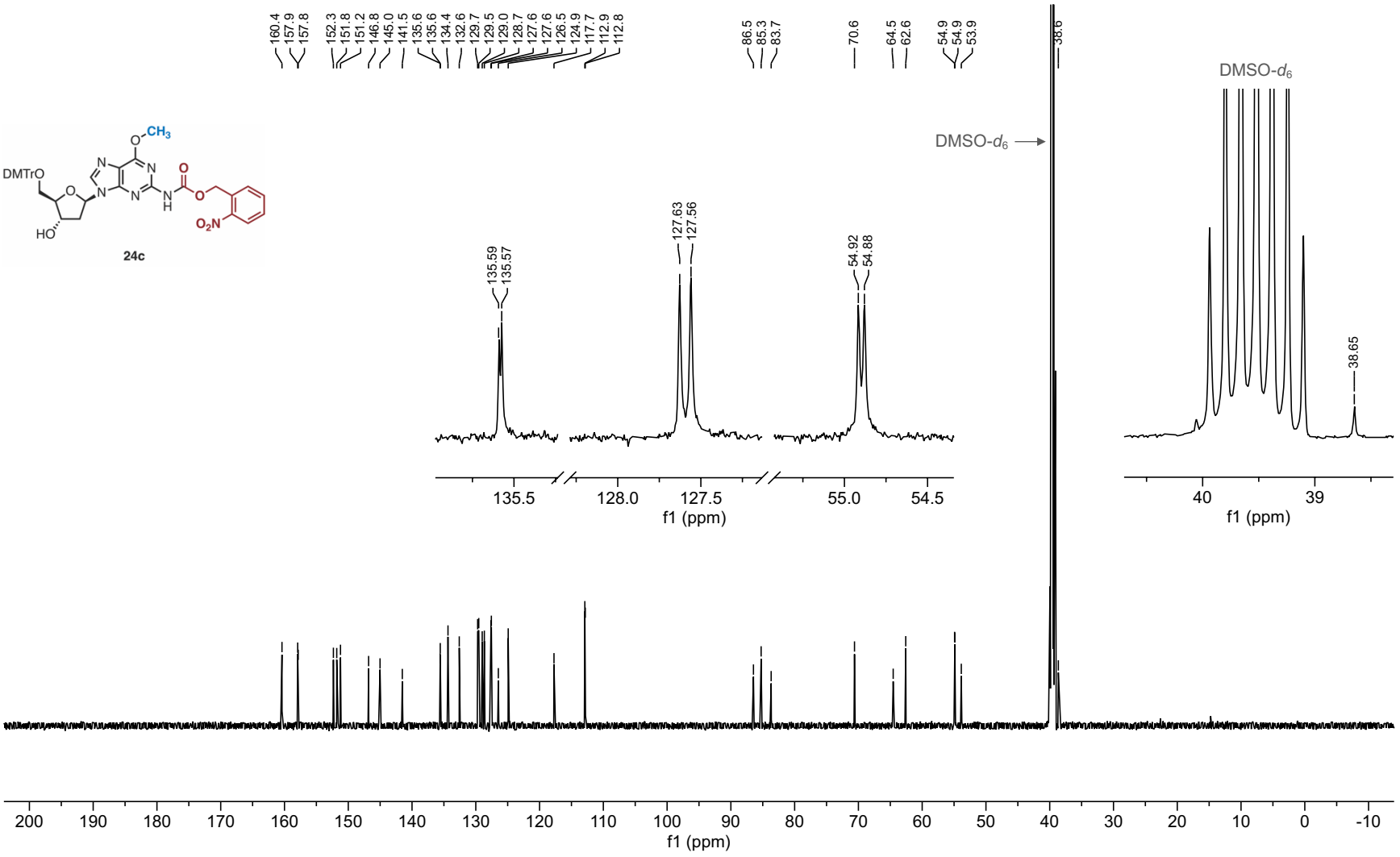
^{1}H–^{13}C HSQC{^{1}H}{^{13}C} NMR, 600/151 MHz, DMSO-d_6, $^{1}J_{CH}$ = 210 Hz, Blue = CH/CH_3, Red = CH_2



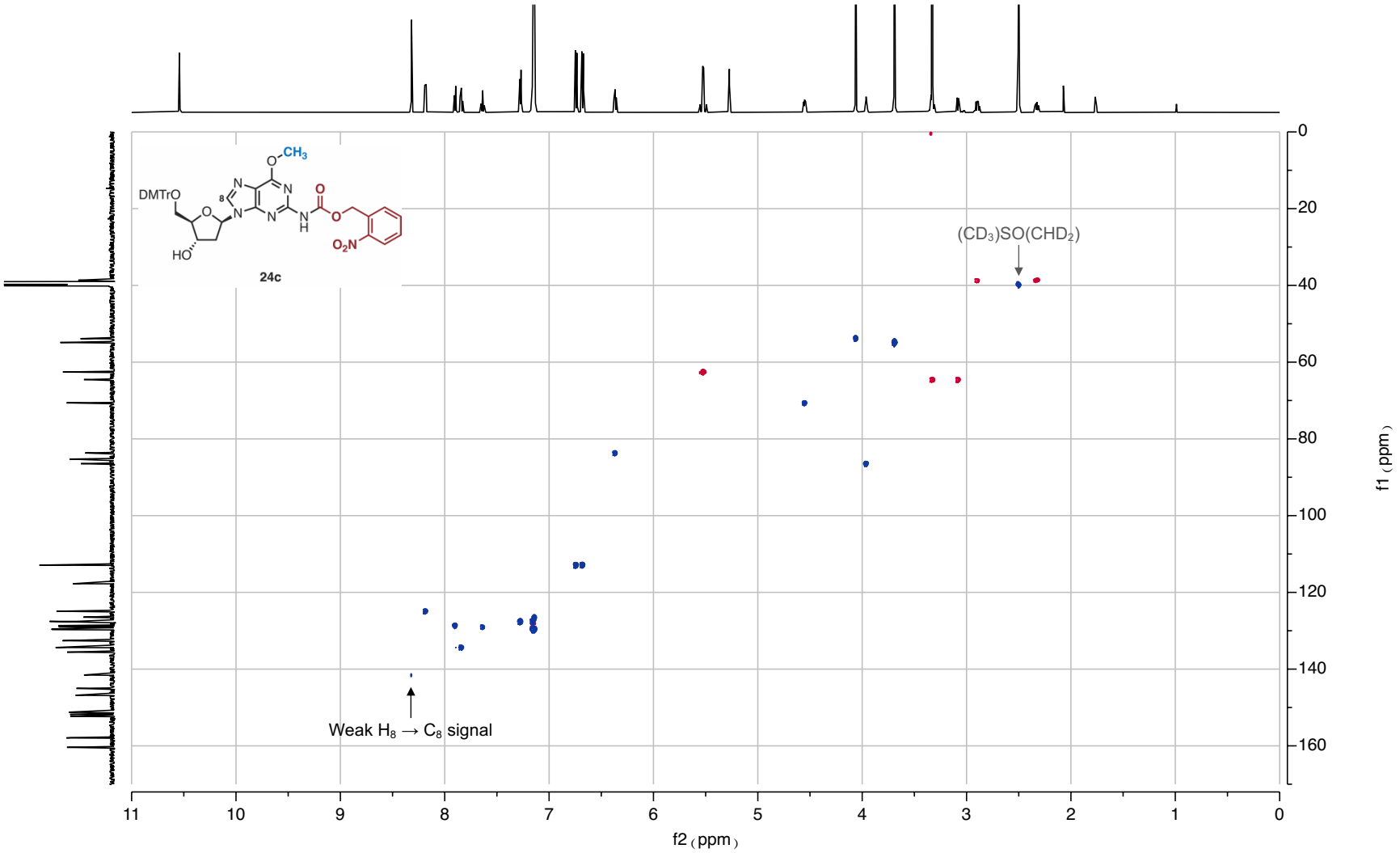
^{1}H NMR, 600 MHz, DMSO-d_6



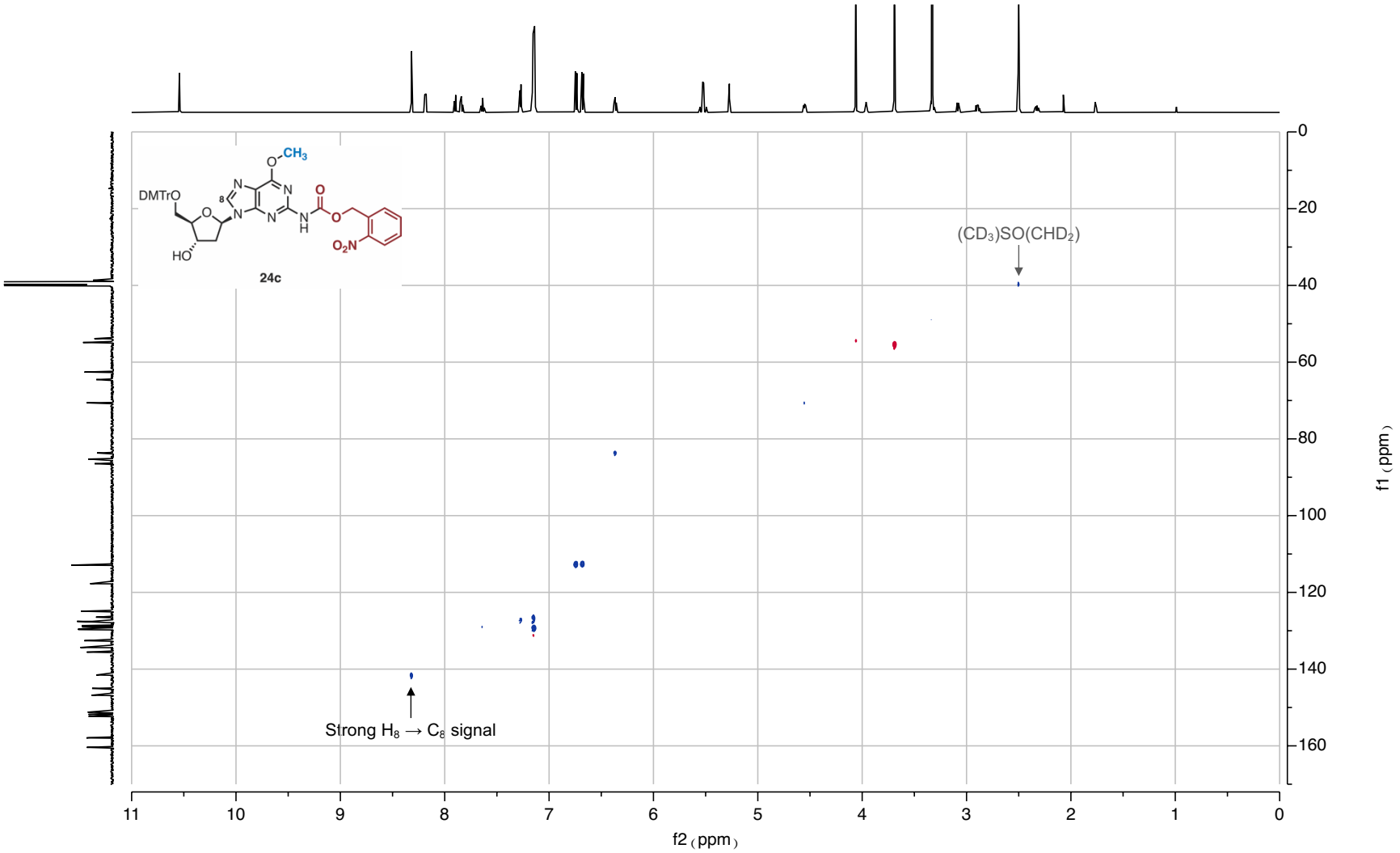
$^{13}C\{^{1}H\}$ NMR, 151 MHz, DMSO-d_6



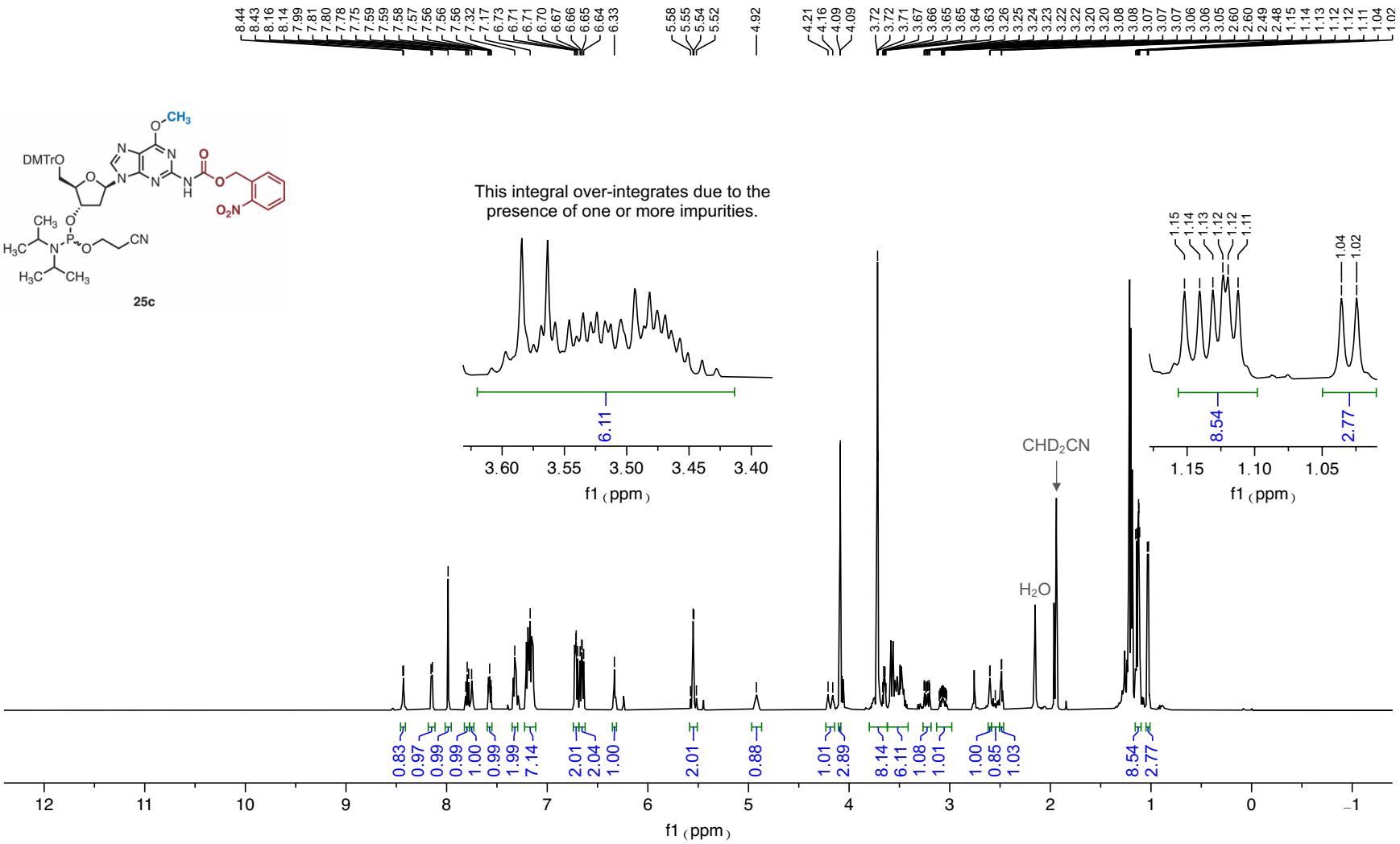
^{1}H–^{13}C HSQC{^{1}H}{^{13}C} NMR, 600/151 MHz, DMSO-d_6, $^{1}J_{CH}$ = 146 Hz, Blue = CH/CH_3, Red = CH_2



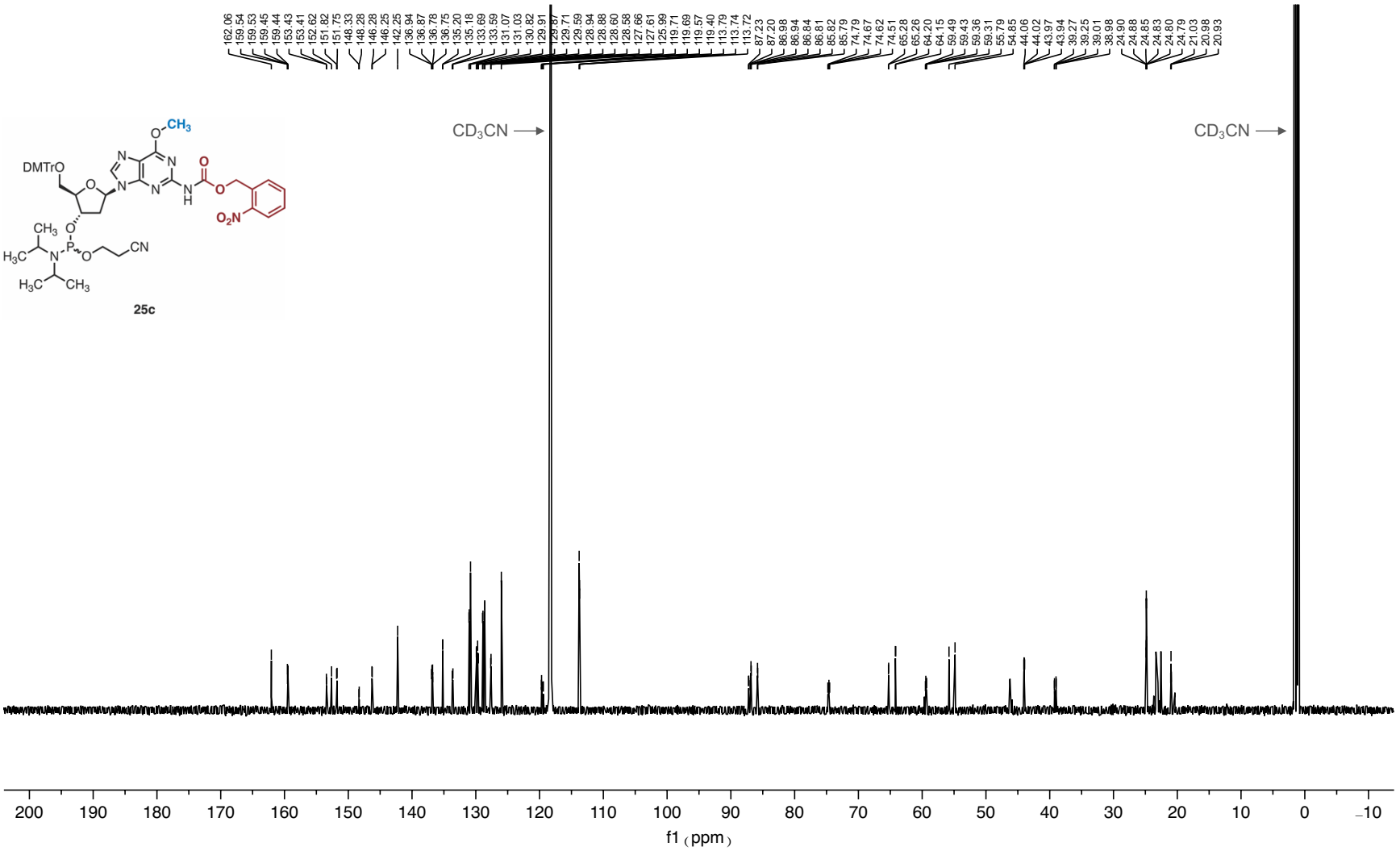
^{1}H–^{13}C HSQC{^{1}H}{^{13}C} NMR, 600/151 MHz, DMSO-d_6, $^{1}J_{CH}$ = 210 Hz, Blue = CH/CH_3, Red = CH_2



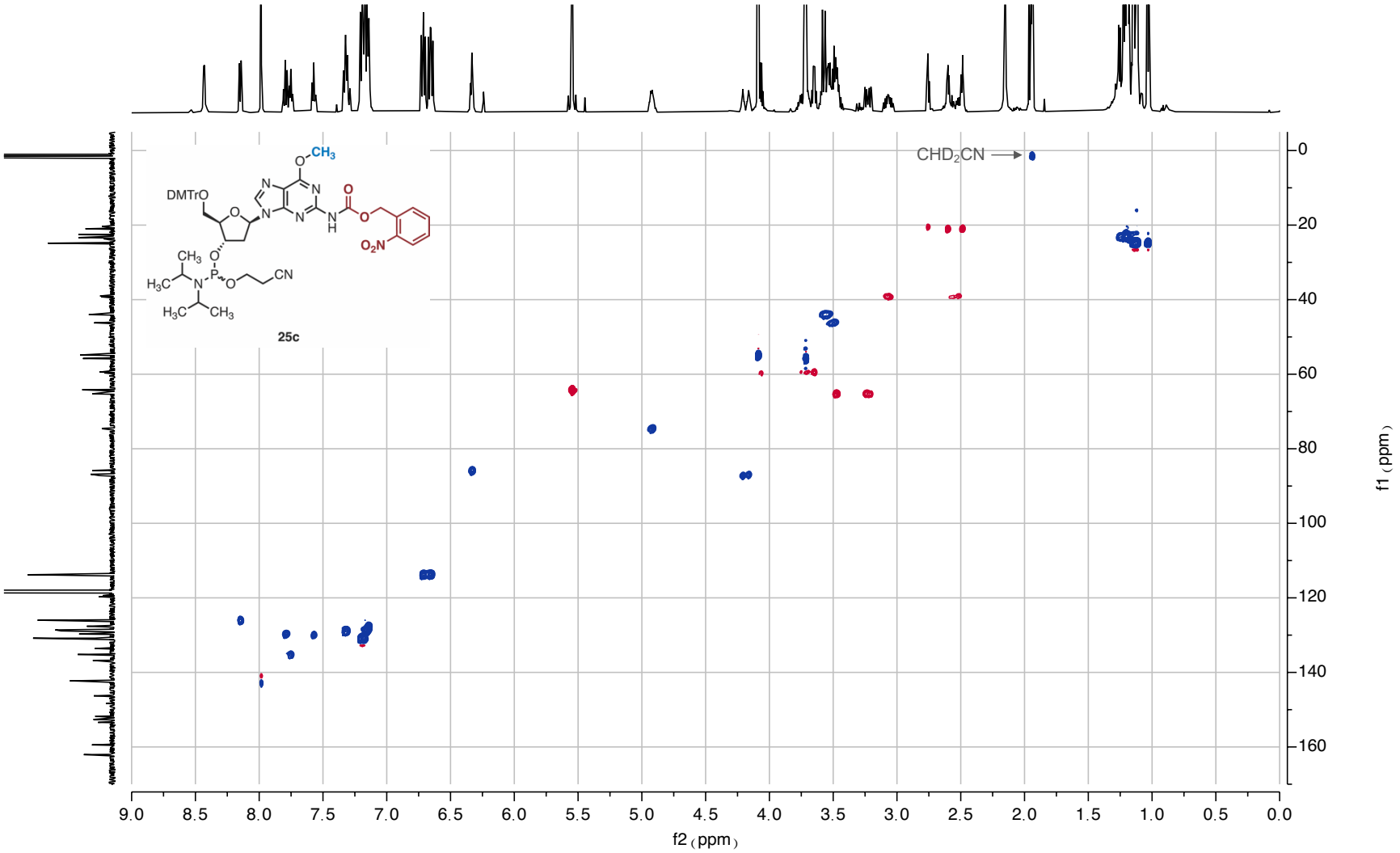
^{1}H NMR, 600 MHz, CD_3CN



$^{13}C\{^{1}H\}$ NMR, 151 MHz, CD_3CN

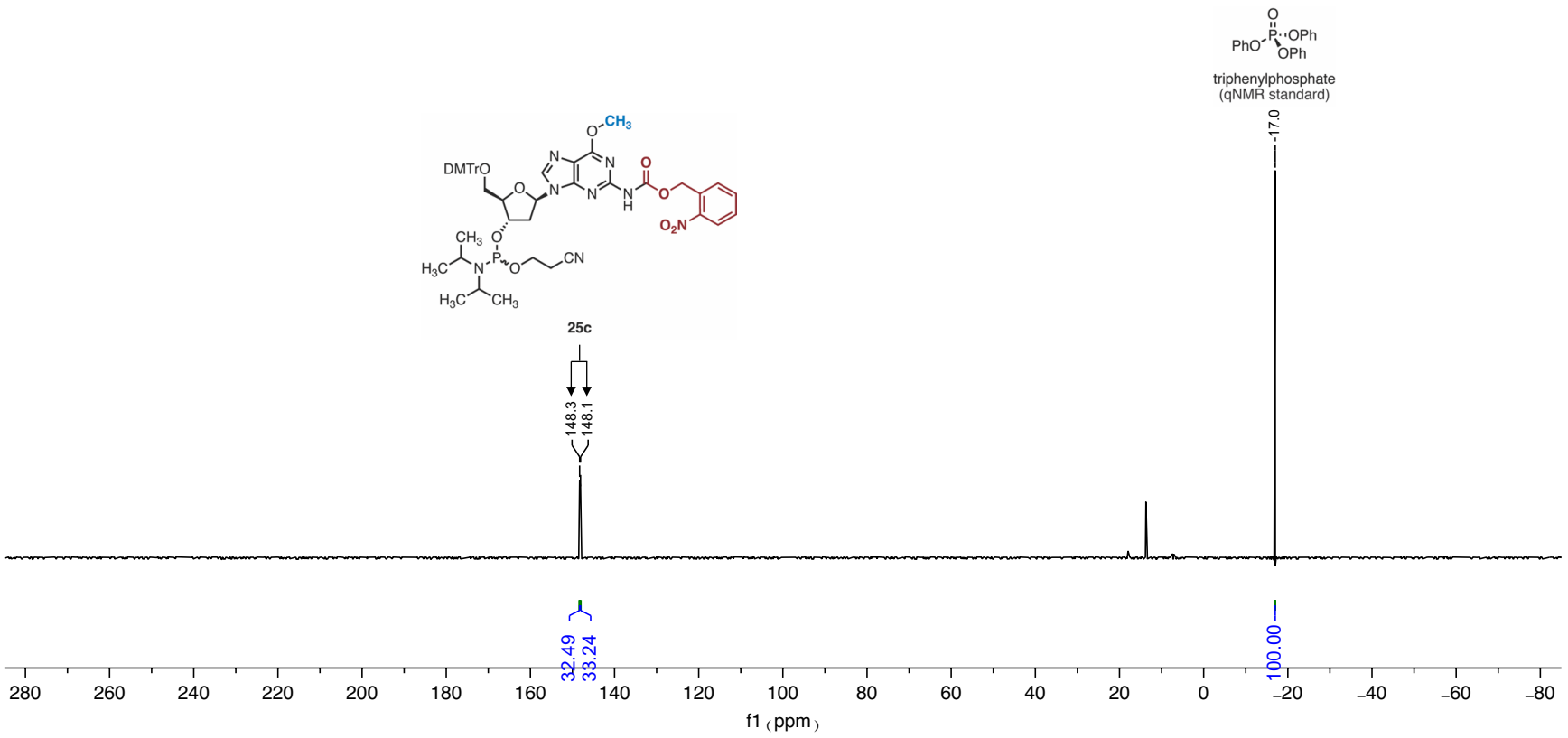


$^{1}H–^{13}C$ HSQC{^{1}H}{^{13}C} NMR, 600/151 MHz, CD_3CN, $^{1}J_{CH}$ = 146 Hz, Blue = CH/CH_3, Red = CH_2



$^{31}P\{^{1}H\}$ qNMR, 202 MHz, CD_3CN (qNMR standard = triphenyl phosphate)

Triphenylphosphate Amount:	2.88 mg
O^6-methyl Phosphoramidite **25c** Amount:	6.46 mg
O^6-methyl Phosphoramidite **25c** Purity:	87 wt%



Bibliography.

(1) Picco, G.; Chen, E. D.; Alonso, L. G.; Behan, F. M.; Goncalves, E.; Bignell, G.; Matchan, A.; Fu, B.; Banerjee, R.; Anderson, E.; Butler, A.; Benes, C. H.; McDermott, U.; Dow, D.; Iorio, F.; Stronach, E.; Yang, F.; Yusa, K.; Saez-Rodriguez, J.; Garnett, M. J. Functional Linkage of Gene Fusions to Cancer Cell Fitness Assessed by Pharmacological and CRISPR-Cas9 Screening. *Nat. Commun.* **2019,** *10*, 2198.
(2) Iorio, F.; Knijnenburg, T. A.; Vis, D. J.; Bignell, G. R.; Menden, M. P.; Schubert, M.; Aben, N.; Goncalves, E.; Barthorpe, S.; Lightfoot, H.; Cokelaer, T.; Greninger, P.; van Dyk, E.; Chang, H.; de Silva, H.; Heyn, H.; Deng, X.; Egan, R. K.; Liu, Q.; Mironenko, T.; Mitropoulos, X.; Richardson, L.; Wang, J.; Zhang, T.; Moran, S.; Sayols, S.; Soleimani, M.; Tamborero, D.; Lopez-Bigas, N.; Ross-Macdonald, P.; Esteller, M.; Gray, N. S.; Haber, D. A.; Stratton, M. R.; Benes, C. H.; Wessels, L. F. A.; Saez-Rodriguez, J.; McDermott, U.; Garnett, M. J. A Landscape of Pharmacogenomic Interactions in Cancer. *Cell* **2016,** *166*, 740–754.
(3) Churchwell, M. I.; Beland, F. A.; Doerge, D. R. Quantification of O^6-Methyl and O^6-Ethyl Deoxyguanosine Adducts in C57BL/6N/$Tk^{+/-}$ Mice Using LC/MS/MS. *J. Chromatogr. B* **2006,** *844*, 60–66.
(4) Gregson, J. M.; McCloskey, J. A. Collision-Induced Dissociation of Protonated Guanine. *Int. J. Mass Spectrom.* **1997,** *165-166*, 475–485.
(5) Spartan 20; Wavefunction, Inc., Irvine, CA.
(6) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A. *Gaussian 16*, Revision C.01; Gaussian, Inc.: Wallingford CT, 2019.
(7) (a) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Self-consistent Molecular Orbital Methods. XX. A Basis Set for Correlated Wave Functions. *J. Chem. Phys.* **1980,** *72*, 650–654. (b) McLean, A. D.; Chandler, G. S. Contracted Gaussian Basis Sets for Molecular Calculations. I. Second Row Atoms, *Z*=11–18. *J. Chem. Phys.* **1980,** *72*, 5639–5648. (c) Hariharan, P. C.; Pople, J. A. The Influence of Polarization Functions on Molecular Orbital Hydrogenation Energies. *Theor. Chim. Acta* **1973,** *28*, 213–222. (d) Spitznagel, G. W.; Clark, T.; von Ragué Schleyer, P.; Hehre, W. J. An Evaluation of the Performance of Diffuse Function-Augmented Basis Sets for Second Row Elements, Na-Cl. *J. Comput. Chem.* **1987,** *8*, 1109–1116.
(8) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B Condens. Matter* **1988,** *37*, 785–789.
(9) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate *Ab Initio* Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010,** *132*, 154104.
(10) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009,** *113*, 6378–6396.
(11) Pangborn, A. B.; Giardello, M. A.; Grubbs, R. H.; Rosen, R. K.; Timmers, F. J. Safe and Convenient Procedure for Solvent Purification. *Organometallics* **1996,** *15*, 1518–1520.
(12) Hentschel, S.; Alzeer, J.; Angelov, T.; Scharer, O. D.; Luedtke, N. W. Synthesis of DNA Interstrand Cross-Links Using a Photocaged Nucleobase. *Angew. Chem., Int. Ed. Engl.* **2012,** *51*, 3466–3469.
(13) Williamson, R. T.; Buevich, A. V.; Martin, G. E.; Parella, T. LR-HSQMBC: A Sensitive NMR Technique to Probe Very Long-Range Heteronuclear Coupling Pathways. *J. Org. Chem.* **2014,** *79*, 3887–3894.
(14) Lin, K.; Gueble, S. E.; Sundaram, R. K.; Huseman, E. D.; Bindra, R. S.; Herzon, S. B. Mechanism-Based Design of Agents That Selectively Target Drug-Resistant Glioma. *Science* **2022,** *377*, 502–511.
(15) (a) Yu, C.; Mannan, A. M.; Yvone, G. M.; Ross, K. N.; Zhang, Y. L.; Marton, M. A.; Taylor, B. R.; Crenshaw, A.; Gould, J. Z.; Tamayo, P.; Weir, B. A.; Tsherniak, A.; Wong, B.; Garraway, L. A.; Shamji, A. F.; Palmer, M. A.; Foley, M. A.; Winckler, W.; Schreiber, S. L.; Kung, A. L.; Golub, T. R. High-Throughput Identification of Genotype-Specific Cancer Vulnerabilities in Mixtures of Barcoded Tumor

Cell Lines. *Nat Biotechnol* **2016,** *34*, 419–423. (b) Corsello, S. M.; Nagari, R. T.; Spangler, R. D.; Rossen, J.; Kocak, M.; Bryan, J. G.; Humeidi, R.; Peck, D.; Wu, X.; Tang, A. A.; Wang, V. M.; Bender, S. A.; Lemire, E.; Narayan, R.; Montgomery, P.; Ben-David, U.; Garvie, C. W.; Chen, Y.; Rees, M. G.; Lyons, N. J.; McFarland, J. M.; Wong, B. T.; Wang, L.; Dumont, N.; O'Hearn, P. J.; Stefan, E.; Doench, J. G.; Harrington, C. N.; Greulich, H.; Meyerson, M.; Vazquez, F.; Subramanian, A.; Roth, J. A.; Bittker, J. A.; Boehm, J. S.; Mader, C. C.; Tsherniak, A.; Golub, T. R. Discovering the Anticancer Potential of Non-Oncology Drugs by Systematic Viability Profiling. *Nat. Cancer* **2020,** *1*, 235–248.
(16) Bossuet-Greif, N.; Vignard, J.; Taieb, F.; Mirey, G.; Dubois, D.; Petit, C.; Oswald, E.; Nougayrede, J. P. The Colibactin Genotoxin Generates DNA Interstrand Cross-Links in Infected Cells. *mBio* **2018,** *9*.
(17) Carlson, B. L.; Pokorny, J. L.; Schroeder, M. A.; Sarkaria, J. N. Establishment, Maintenance and in Vitro and in Vivo Applications of Primary Human Glioblastoma Multiforme (GBM) Xenograft Models for Translational Biology Studies and Drug Discovery. *Curr. Clin. Pharmacol.* **2011,** *52*, 14.16.11–14.16.23.
(18) Engels, J. W.; Hofmann, T.; Zweig, K. A New Synthetic Approach for the Synthesis of N^2-Modified Guanosines. *Synthesis* **2005,** *2005*, 1797–1800.
(19) Kaburagi, Y.; Kishi, Y. Operationally Simple and Efficient Workup Procedure for TBAF-Mediated Desilylation: Application to Halichondrin Synthesis. *Org. Lett.* **2007,** *9*, 723–726.
(20) Bodell, W. J.; Pongracz, K. Chemical Synthesis and Detection of the Cross-Link 1-[N^3-(2′-Deoxycytidyl)]-2-[N^1-(2′-deoxyguanosinyl)]ethane in DNA Reacted with 1-(2-Chloroethyl)-1-nitrosourea. *Chem. Res. Toxicol.* **1993,** *6*, 434-438.
(21) Hoang, K. M.; Lees, N. R.; Herzon, S. B. Programmable Synthesis of 2-Deoxyglycosides. *J. Am. Chem. Soc.* **2019,** *141*, 8098–8103.