Supporting Information

A Convenient Route to Synthesize N^2-(Isobutyryl)-9-(carboxymethyl)guanine for *aeg*-PNA backbone

Dhrubajyoti Datta[a]*

[a]Department of Chemistry, Chemical Biology Unit, Indian Institute of Science Education and Research Pune, Dr. Homi Bhabha Road, Pune 411008, India

Table of Contents

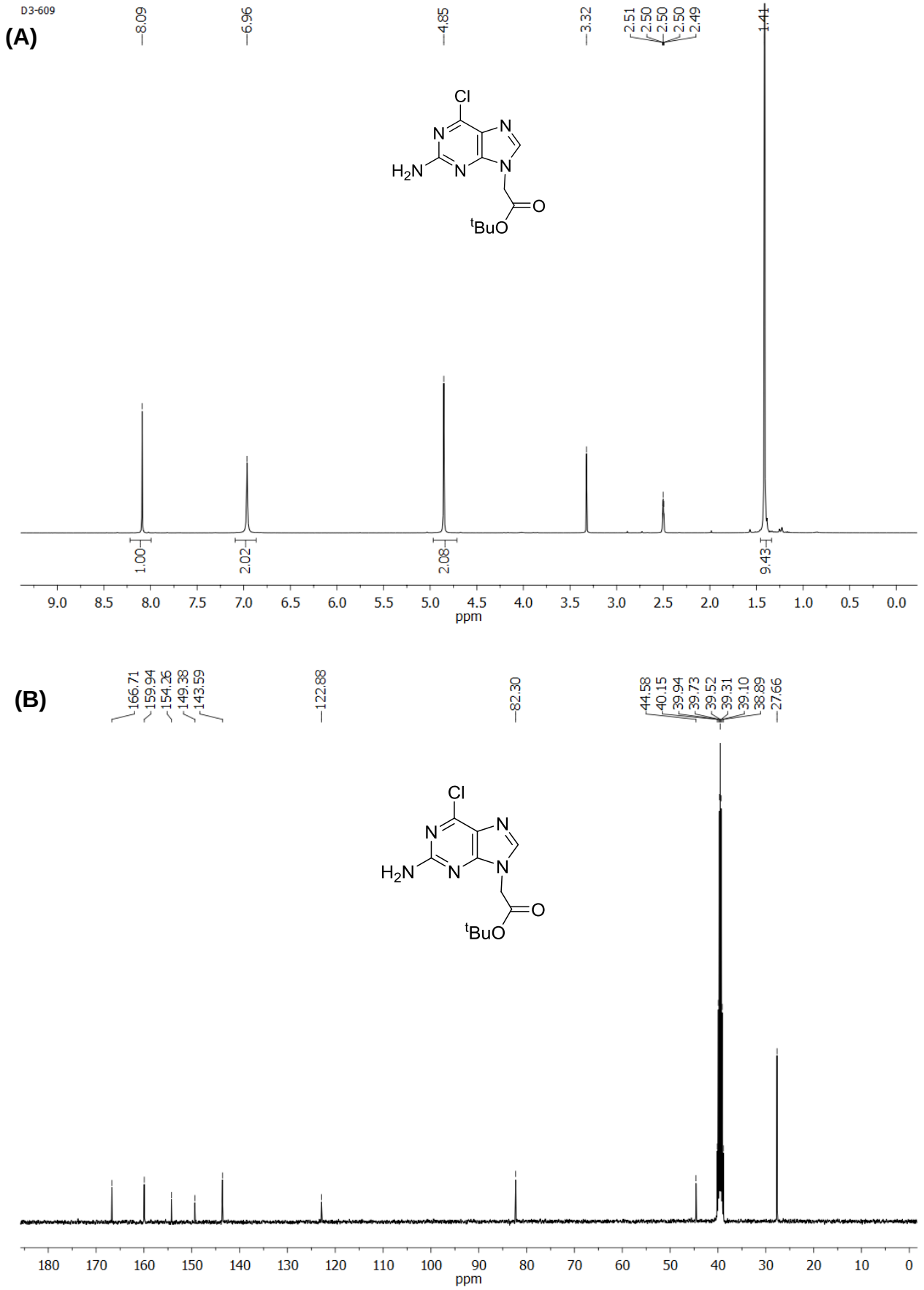


Figure S1: (A) ^{1}H and **(B)** ^{13}C NMR of **6**

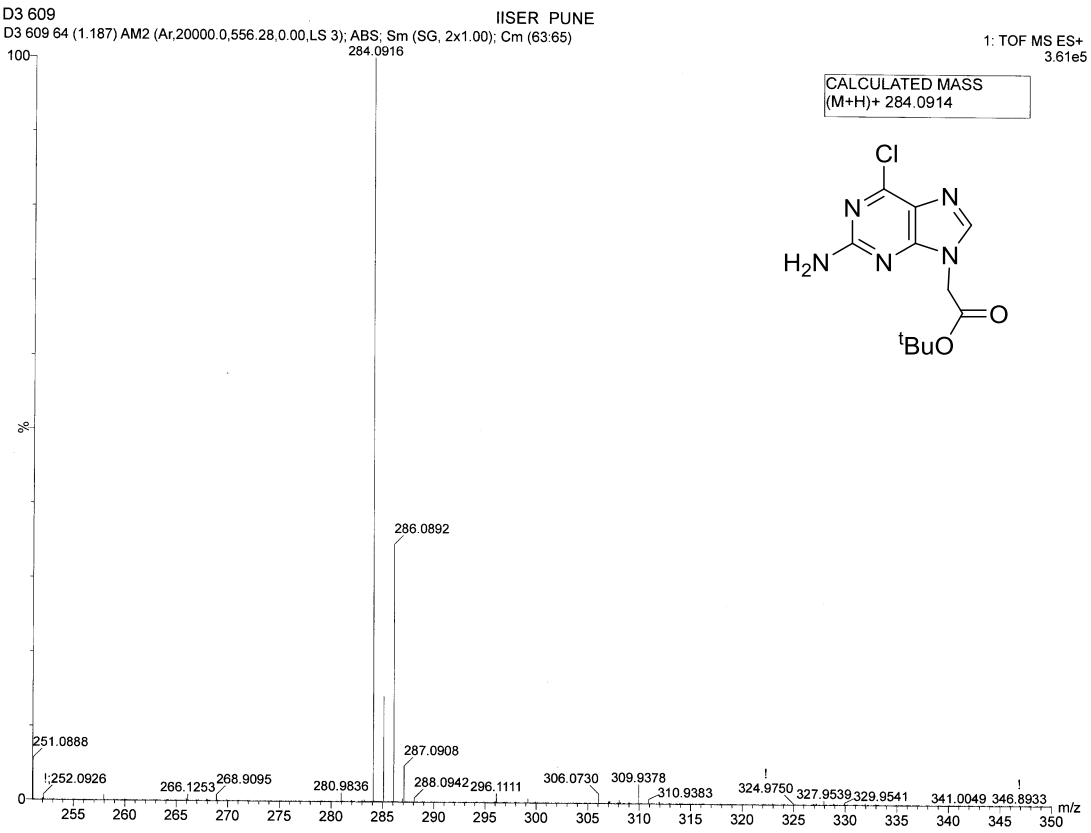


Figure S2: HRMS data of **6**

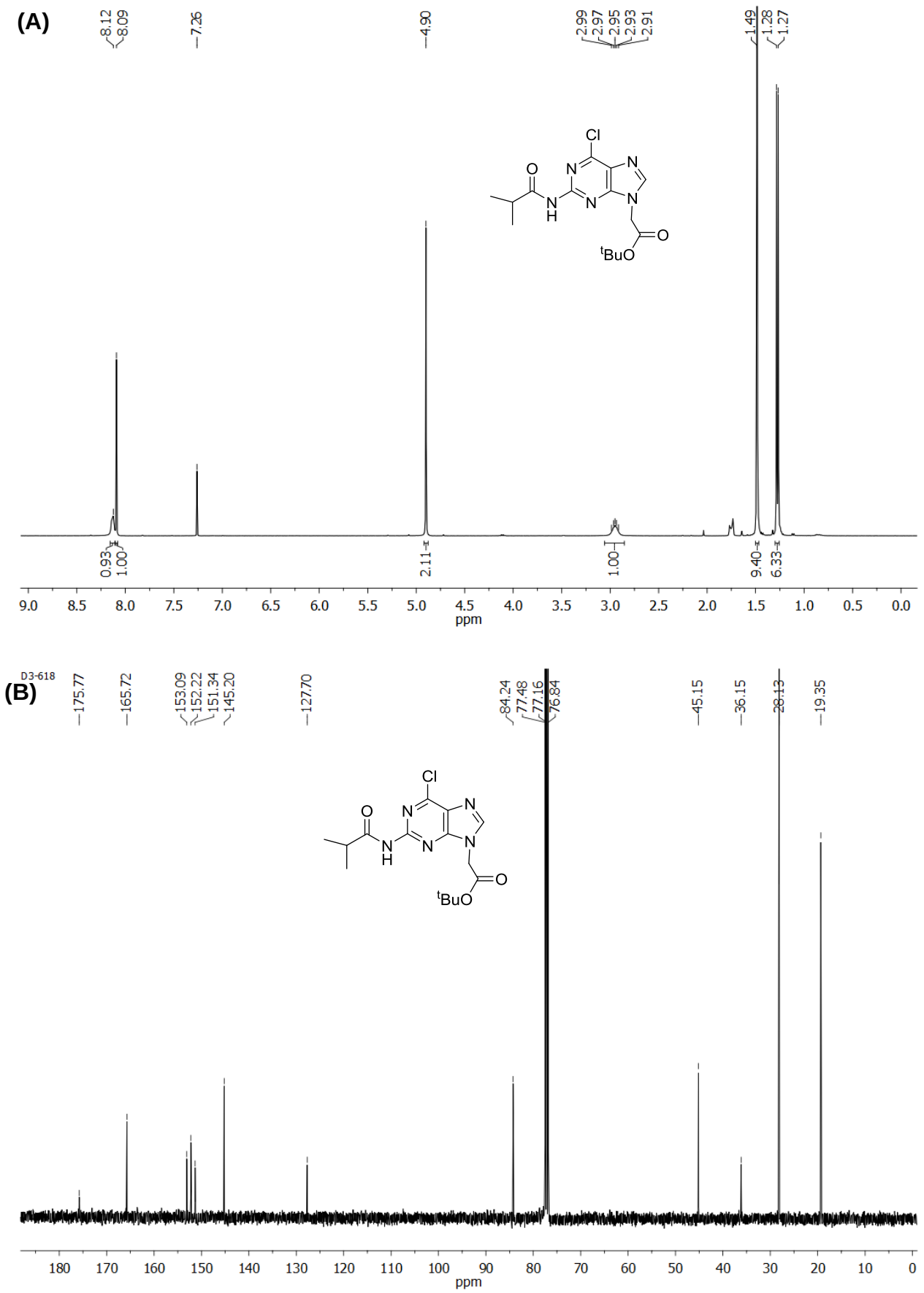


Figure S3: (A) ^{1}H and **(B)** ^{13}C NMR of **7**

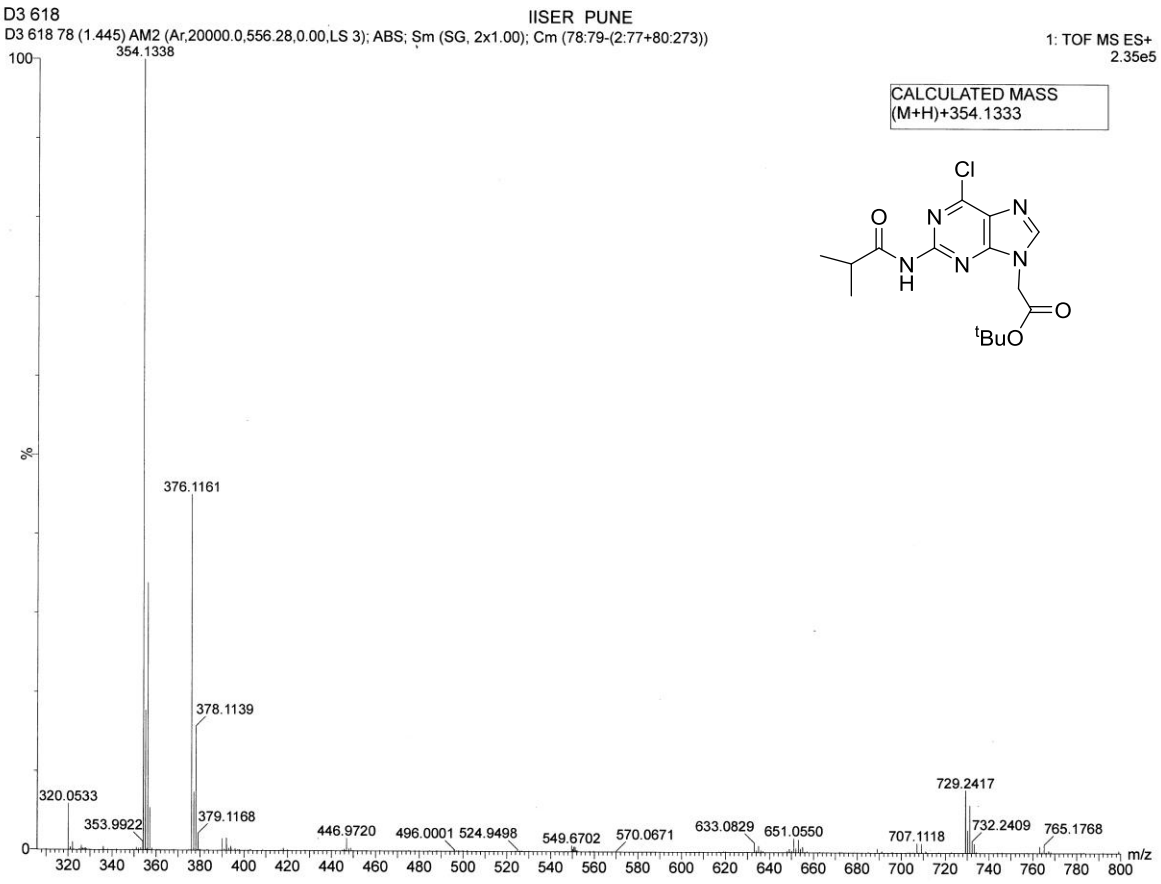


Figure S4: HRMS data of **7**

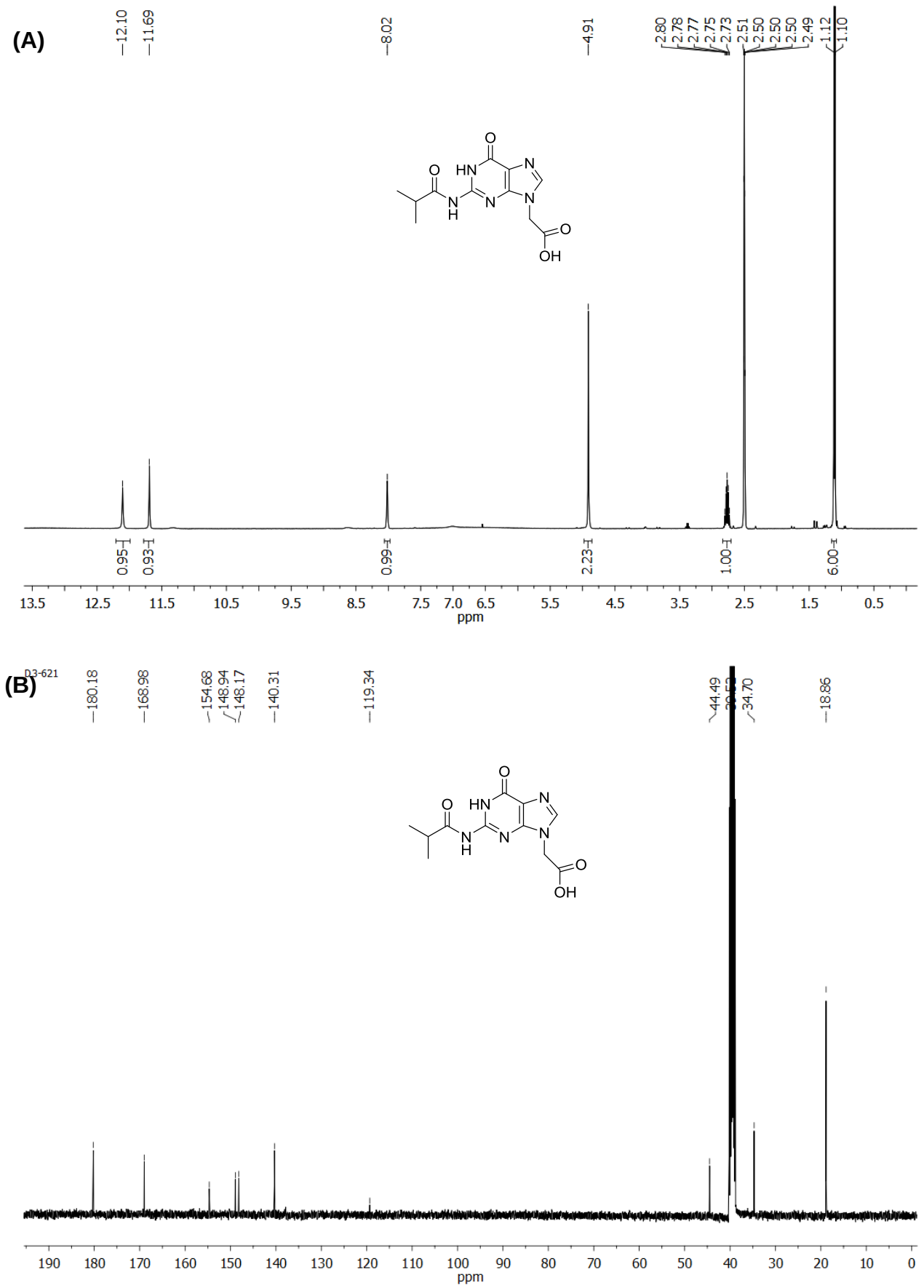


Figure S5: (A) ^{1}H and **(B)** ^{13}C NMR of **1**

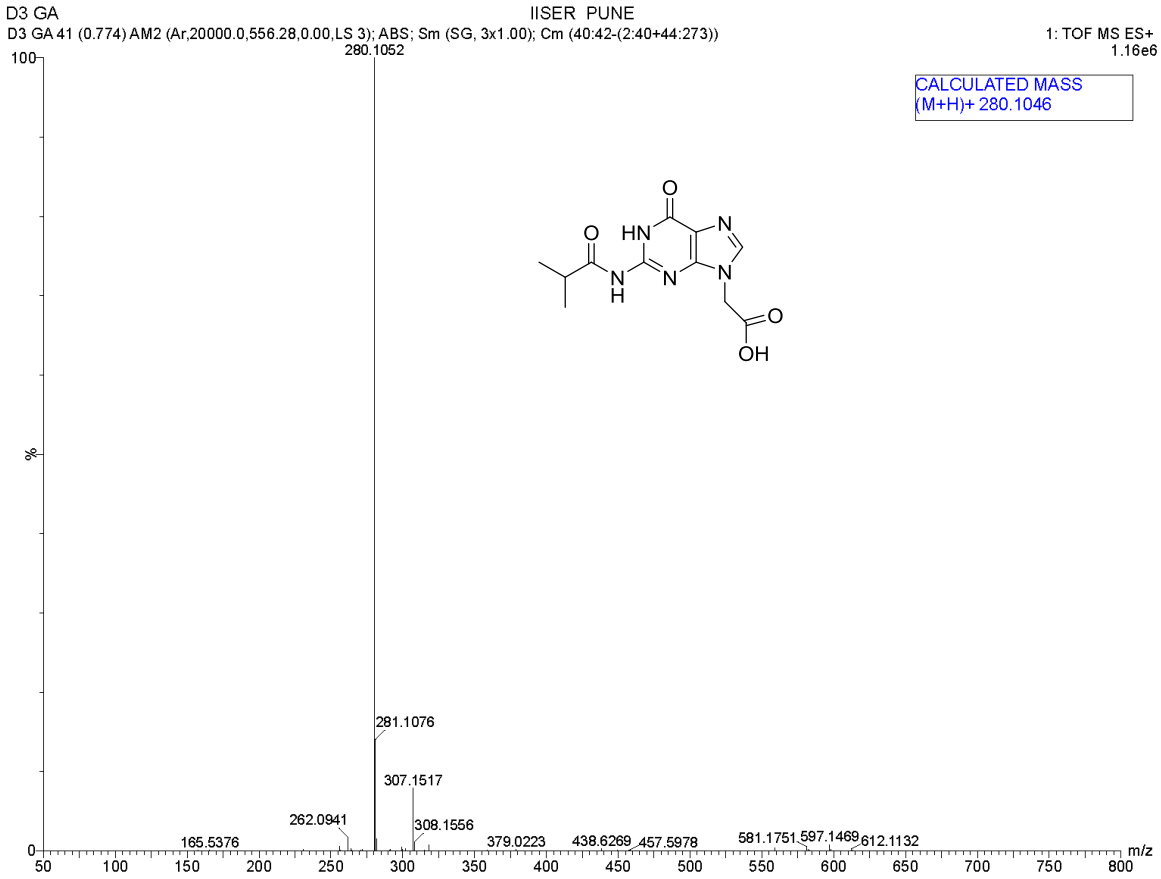


Figure S6: HRMS data of **1**

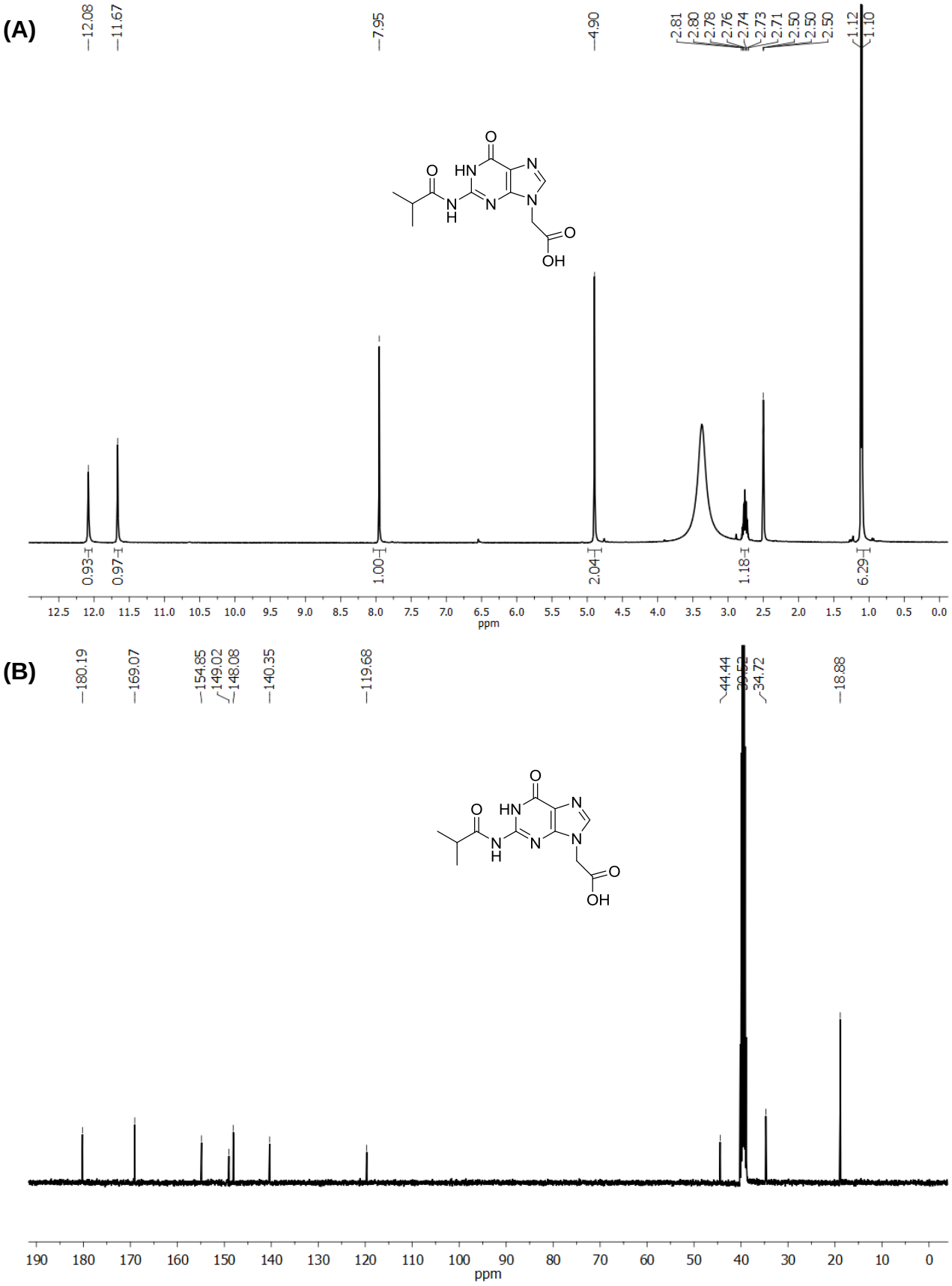


Figure S7: (A) ^{1}H and **(B)** ^{13}C NMR of **1**

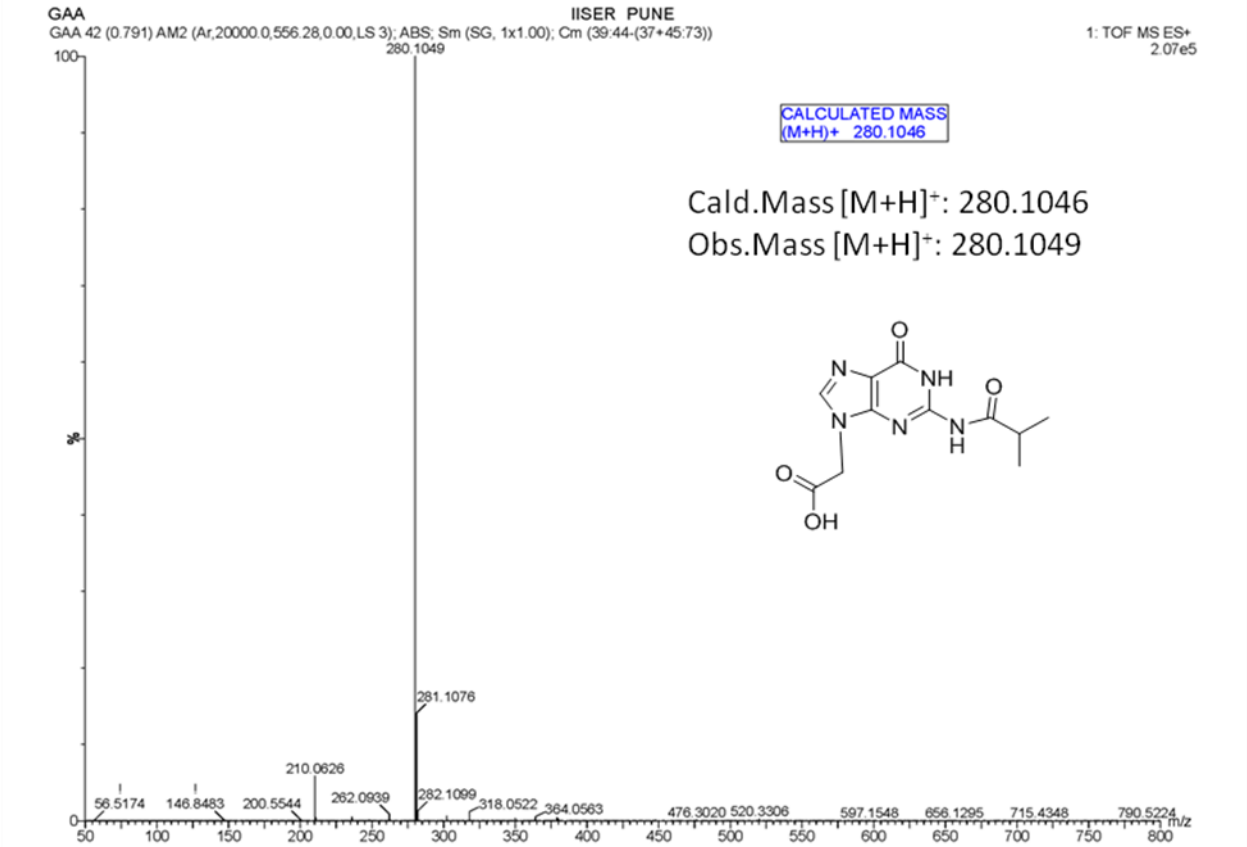


Figure S8: HRMS data of **1**

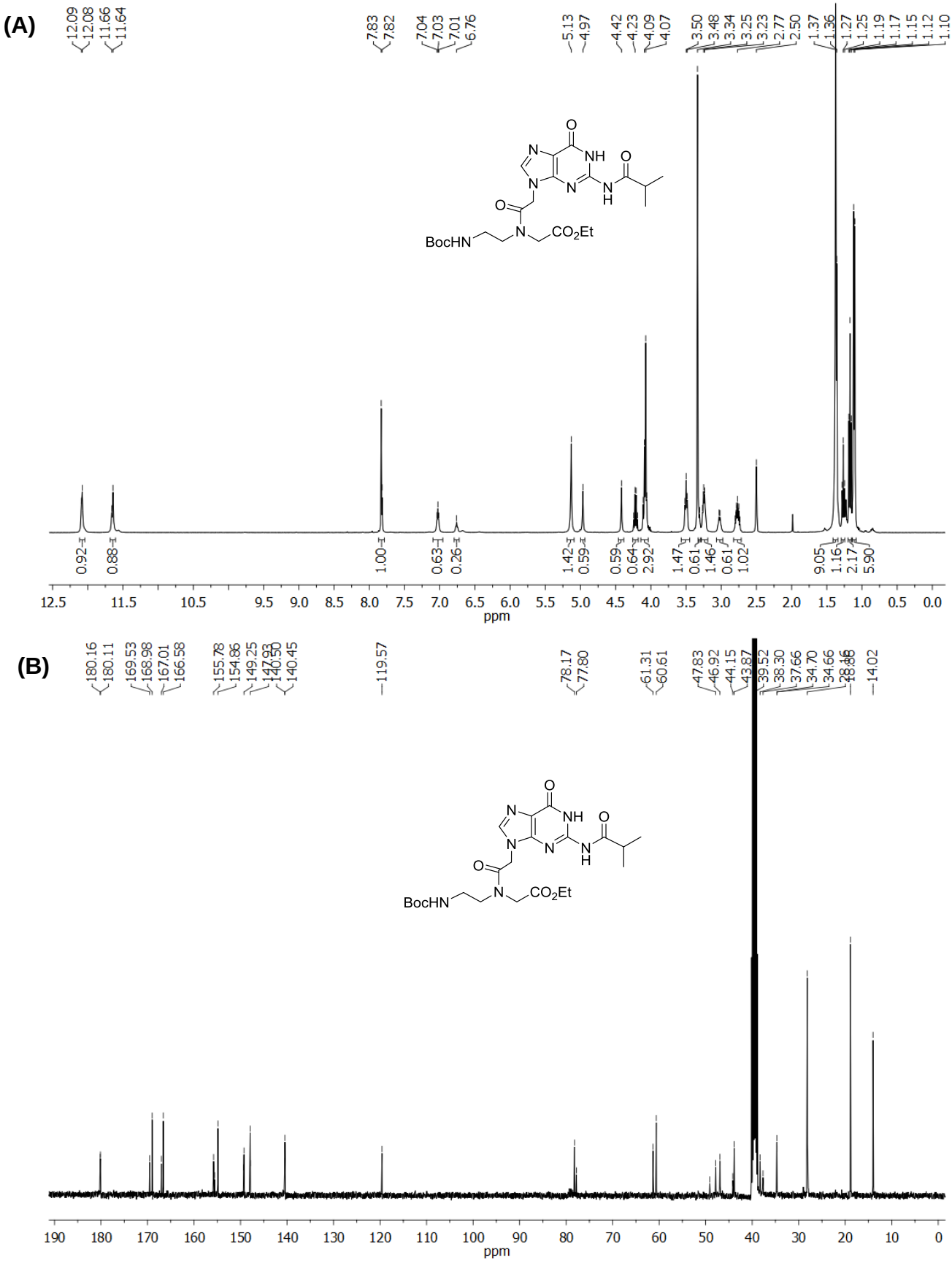


Figure S9: (A) ^{1}H and **(B)** ^{13}C NMR of **8**

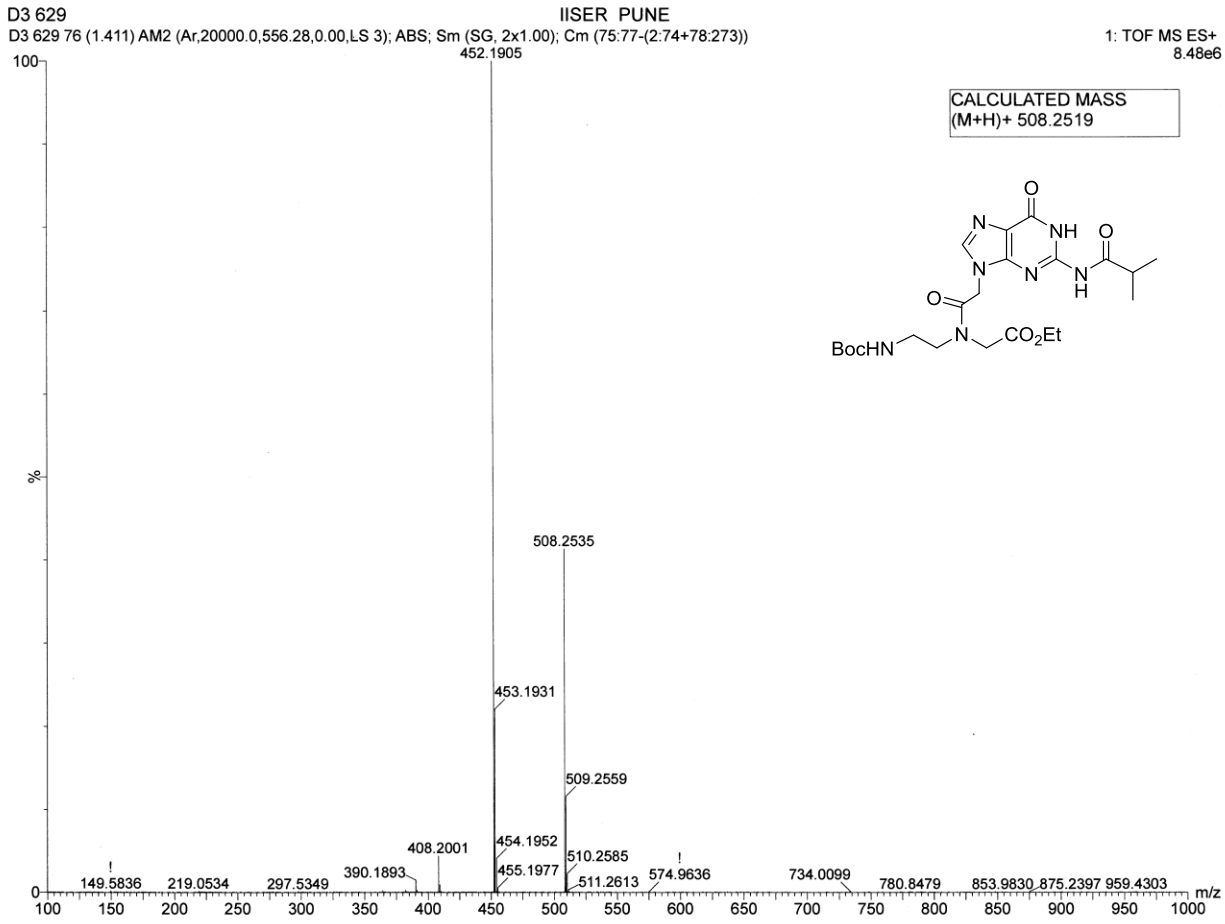


Figure S10: HRMS data of **8**

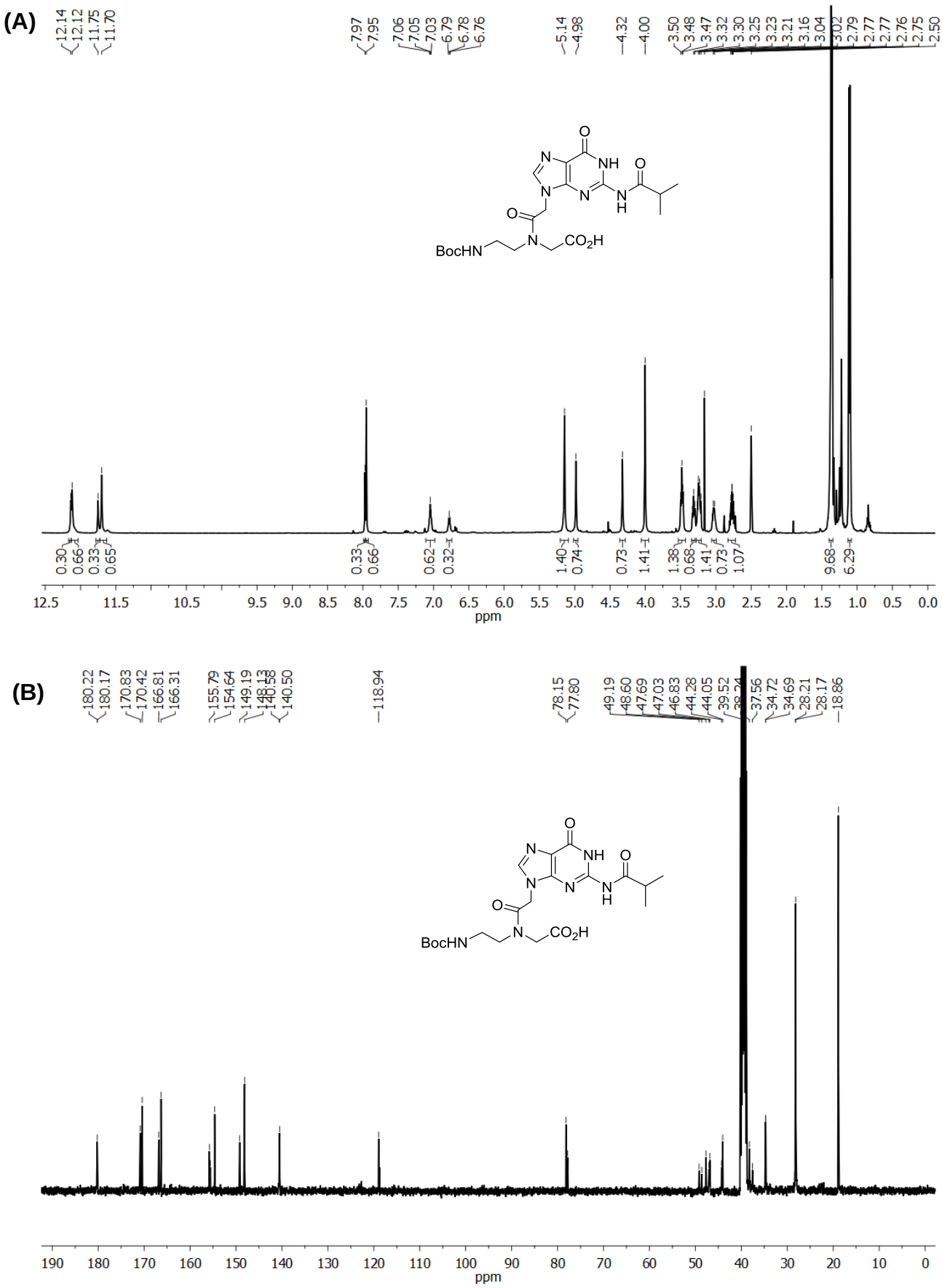


Figure S11: (A) ^{1}H and **(B)** ^{13}C NMR of **9**

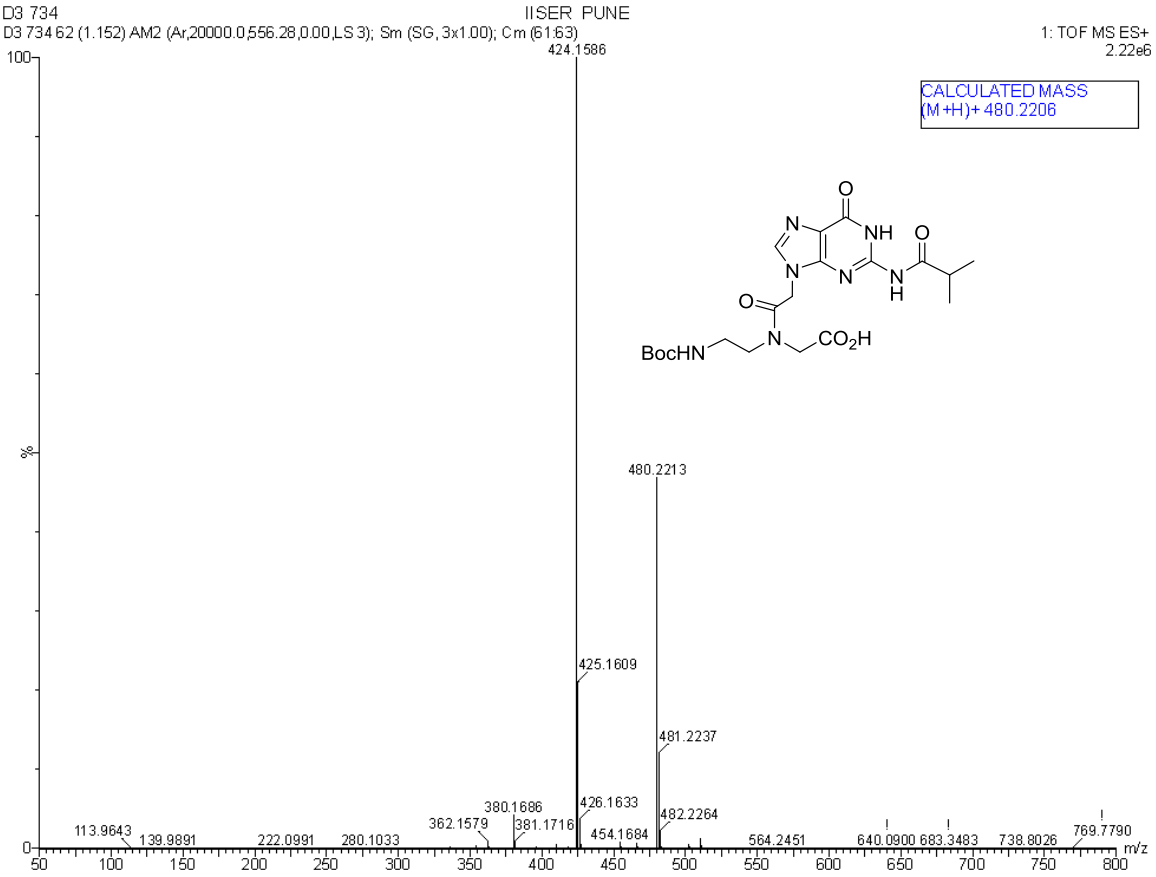


Figure S12: HRMS data of **9**

Two Dimensional (2D) NMR spectroscopy: 2D NMR spectroscopy study was carried out by using Bruker 400 MHz spectrometer in DMSO-d_6 solvent. Resonance assignments were obtained by NOESY analysis. 2D data was collected in phasesensitive mode by using the time-proportional phase incrementation (TPPI) method. Sets of 2048 and 512 data points were used in the t_2 and t_1 dimensions, respectively. A spectral width of 9014.42 Hz was used in both dimensions. Spin-lock time used was 250 ms to obtain NOESY spectra. Zero filling was carried out to finally yield a data set of 2 K x 1 K. A shifted square–sine–bell window was used before processing.

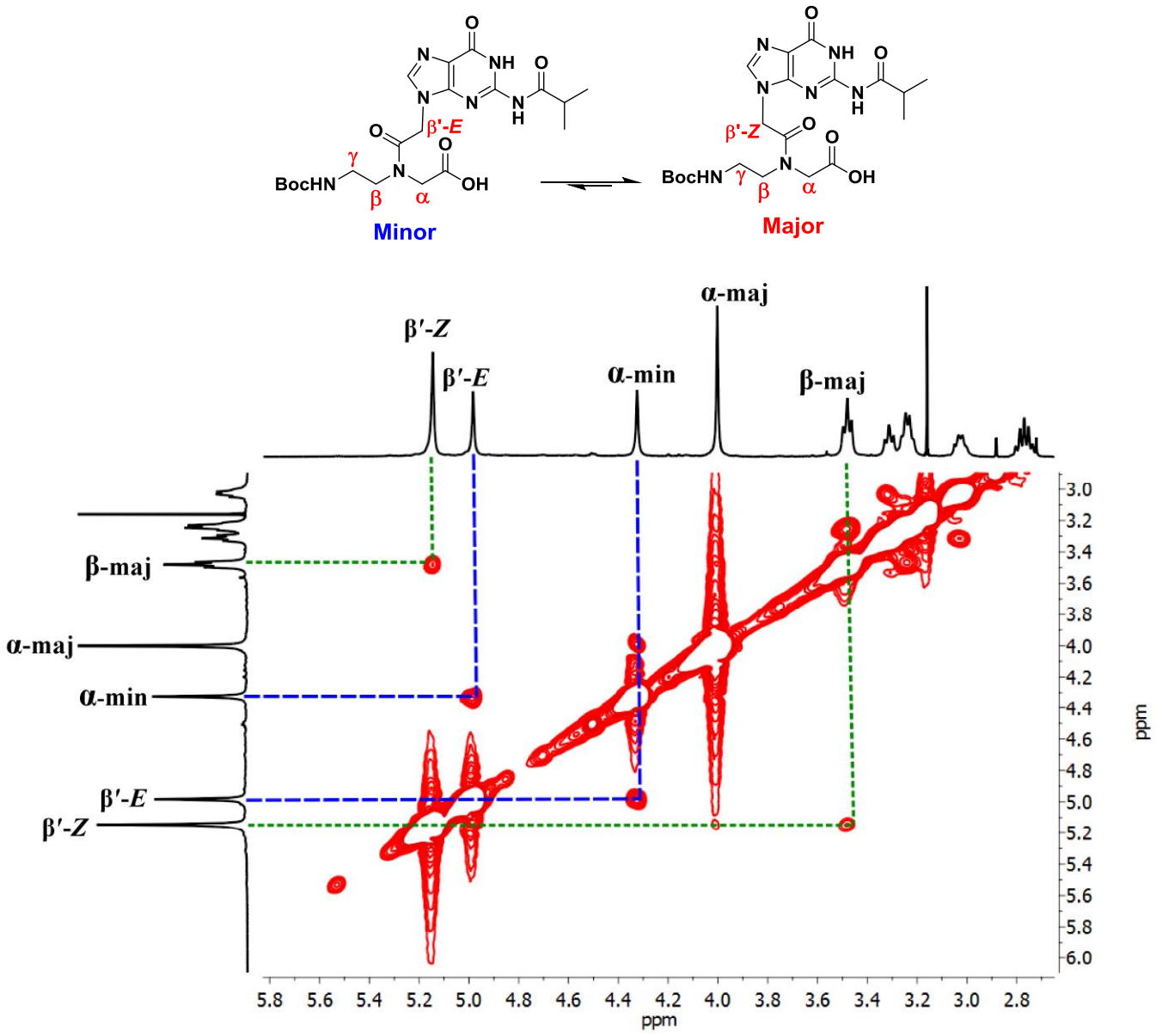


Figure S13: Partial NOESY of **9** in 400 MHz NMR shows major Z-rotamer in DMSO-d_6

Synthesis of PNA on Solid Phase: Boc-G-PNA monomer incorporated into peptide sequence by solid phase synthesis on MBHA resin having 0.25 mmol/g loading value. Deprotection of Boc group from the N-terminus of the resin bound lysine with 50% TFA in DCM (3 × 15 min) was followed by washing with DCM and DMF (3 × 10 mL) to give a TFA salt of amine which was neutralized using 10% DIPEA in DCM (3 × 10 min) to generate free amine. After washing with DCM and DMF (3 × 10 mL), the free amine was coupled with carboxylic acid of incoming monomer in dry DMF (0.5 mL) using HOBt (3 equiv), HBTU (3 equiv), and DIPEA (3 equiv). The reagents were then removed by filtration, and the resin was washed with DMF. This protocol was repeated until the desired length of peptides obtained. The resin bound peptide was cleaved from solid support by triflouroacetic acid (TFA) and trifluoromethanesulfonic acid (TFMSA). After evaporation of volatile matters, chilled diethyl ether was added to precipitate the peptide which was collected after decanting off the supernatant solvent. The residue thus obtained was dissolved in 30% aqueous ammonia (1.5 mL) and stirred for 30 h to afford the desired peptide **pG$_5$**.

N. B. The coupling sequence here is initiated with L-lysine to increase the peptide solubility in water.

Purification of pG$_5$ by Reverse-Phase HPLC: For the purification of the PNA, a semi-preparative BEH130 C18 (10 × 250 mm) column was used. Purification of PNA oligomers was performed with the gradient elution method: A to 100 % B in 33min; A = 0.1% TFA in CH_3CN/H_2O (5:95); B = 0.1 % TFA in CH_3CN / H_2O (1:1) with a flow rate of 2 mL / min. Oligomer was monitored at 220 and 260 nm wavelengths during purification.

After purification peptide MALDI spectra were measured using Bruker Daltonics AutoflexII TOF/TOF spectrometer. The samples were prepared using 2,5-dihydroxybenzoic acid (DHB) as a matrix.

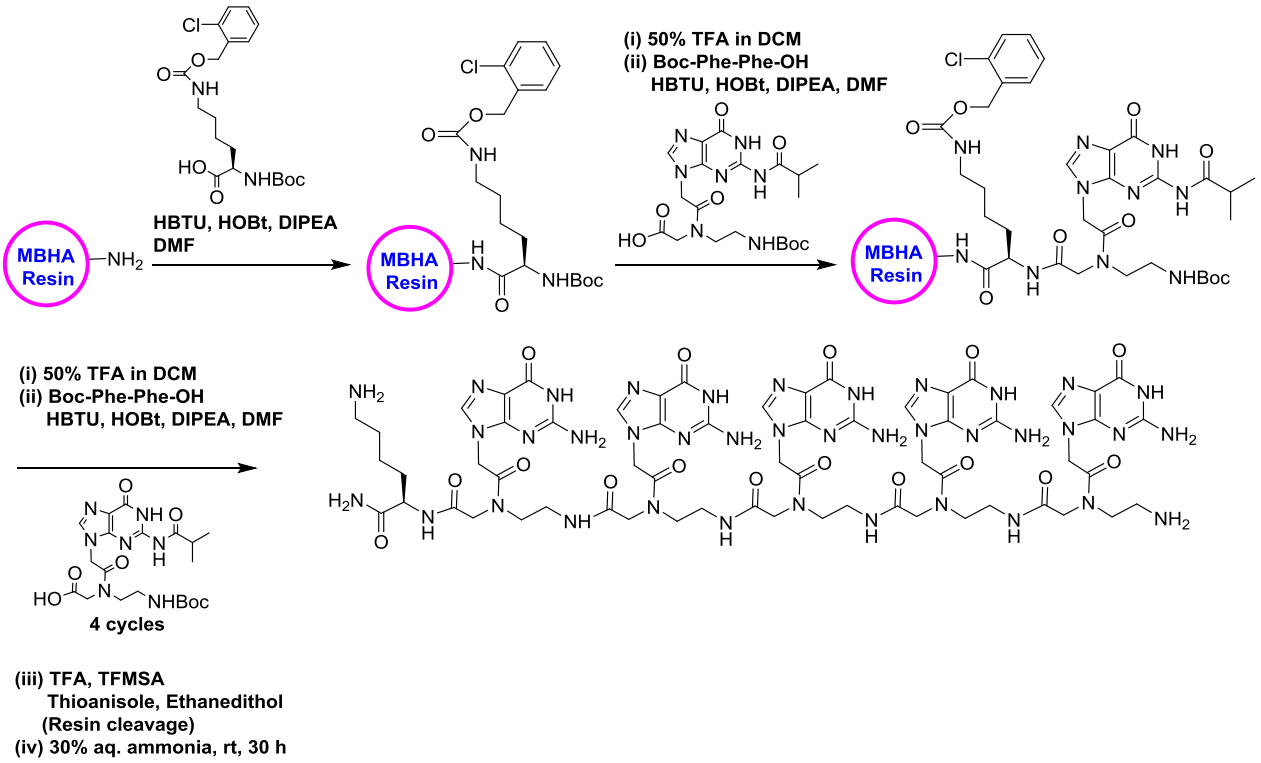
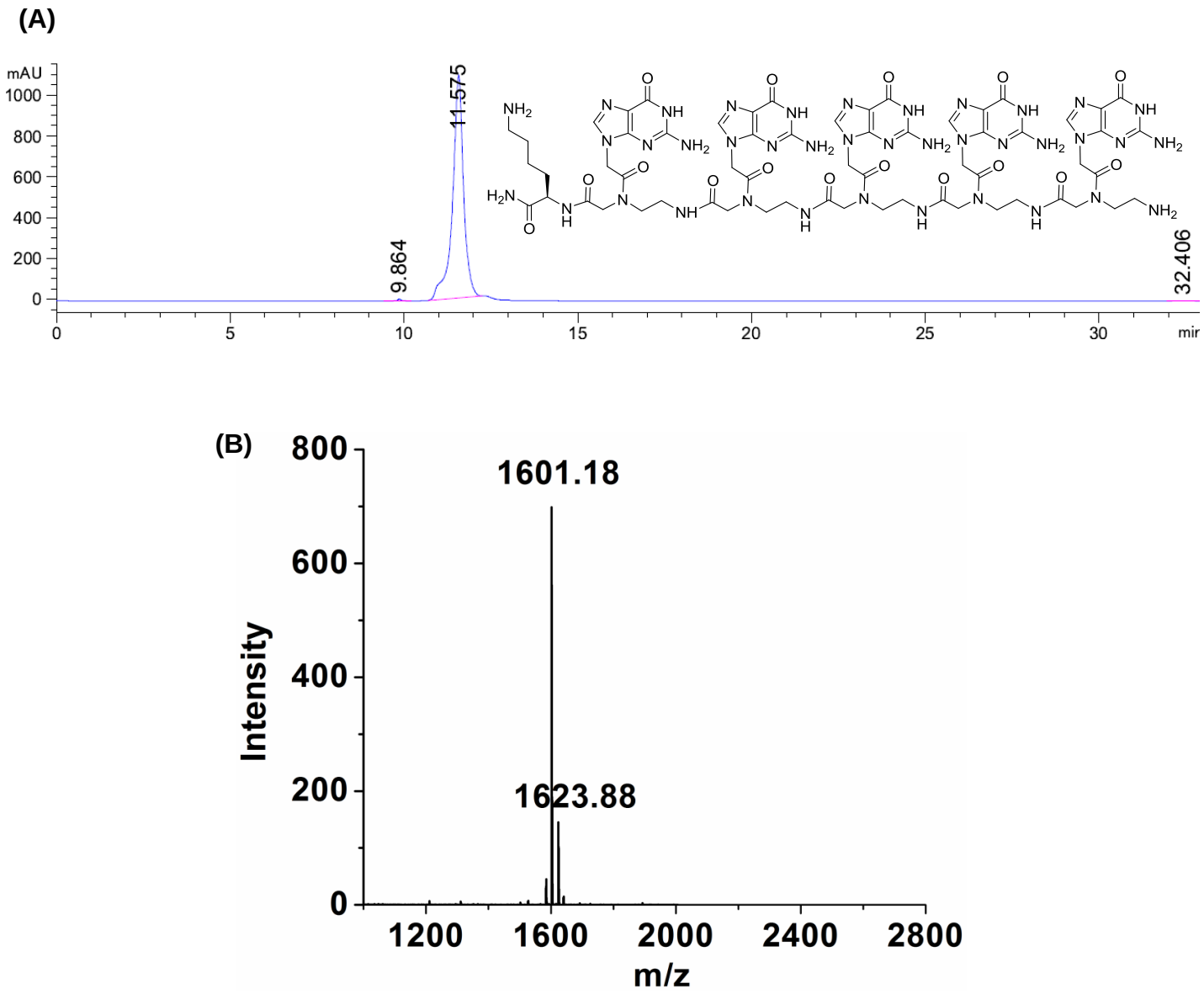


Figure S14: Schematic representation of PNA synthesis using *N*9-G-PNA monomer **9**.



Sequence of oligomer	Mol. Formulae	Calcd. Mass (M+H)	Obsd. Mass (M+H)
pG_5 (H-G-G-G-G-G-Lys-NH_2)	$C_{61}H_{80}N_{38}O_{16}$	1601.67	1601.18

Figure S15: (A) HPLC trace, and (C) MALDI Tof mass of desired peptide **pG_5**.

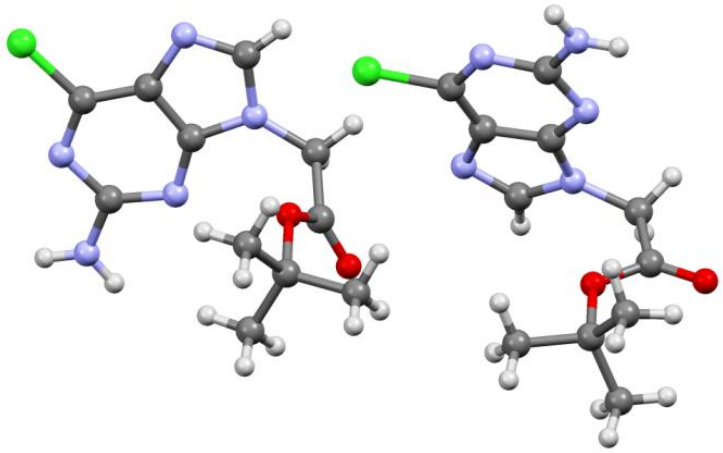


Figure S16: Crystal structure of *tert*-butyl(2-amino-6-chloropurin-9-yl)acetate **6** showing N9 isomer (CCDC 1880599).

Crystallographic data of **6**

Bond precision: C-C = 0.0060 A Wavelength=0.71073

Cell: a=10.9111(12) b=21.694(2) c=11.4134(13)
alpha=90 beta=94.773(4) gamma=90

Temperature: 296 K

	Calculated	Reported
Volume	2692.2(5)	2692.3(5)
Space group	P 21/c	P 21/c
Hall group	-P 2ybc	-P 2ybc
Moiety formula	C11 H14 Cl N5 O2	-
Sum formula	C11 H14 Cl N5 O2	C11 H14 Cl N5 O2
Mr	283.72	283.72
Dx,g cm-3	1.400	1.400
Z	8	8
Mu (mm-1)	0.290	0.290
F000	1184.0	1184.0
F000'	1185.59	
h,k,lmax	14,28,15	14,28,15
Nref	6723	6723

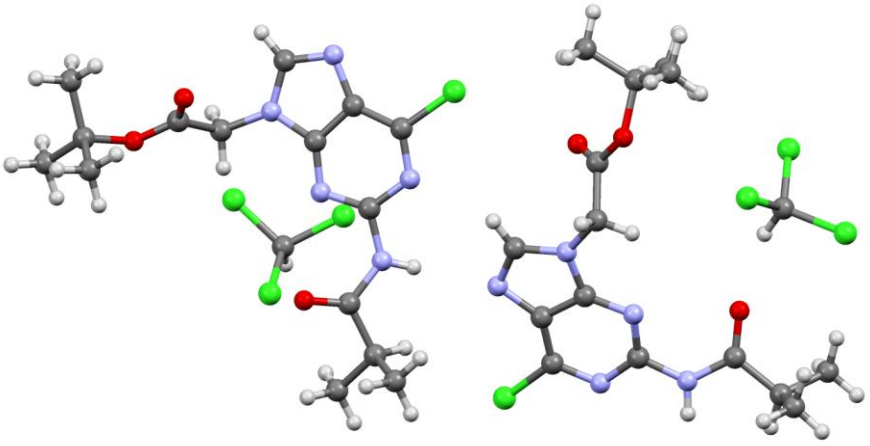


Figure S17: Crystal structure of *tert*-butyl[2-(isobutyryl)amino-6-chloropurin-9-yl]acetate **7** showing N9 isomer (CCDC 1880600).

Crystallographic data of **7**

Bond precision:	C-C = 0.0042 A		Wavelength=0.71073
Cell:	a=10.7765(11)	b=14.4165(12)	c=15.4237(15)
	alpha=104.294(3)	beta=108.872(3)	gamma=90.842(3)
Temperature:	296 K		

	Calculated	Reported
Volume	2185.5(4)	2185.5(4)
Space group	P -1	P -1
Hall group	-P 1	-P 1
Moiety formula	C15 H20 Cl N5 O3, C H Cl3	-
Sum formula	C16 H21 Cl4 N5 O3	C16 H21 Cl4 N5 O3
Mr	473.18	473.18
Dx,g cm-3	1.438	1.438
Z	4	4
Mu (mm-1)	0.568	0.568
F000	976.0	976.0
F000'	978.68	
h,k,lmax	14,19,20	14,19,20
Nref	10912	10912