SUPPLEMENTARY INFORMATION

Streptomyces aureorectus DSM 41692 and *Streptomyces virens* DSM 41465 are producers of the antibiotic nucleocidin and 4'-fluoroadenosine is identified as a co-product

Yawen Chen,[a] Qingzhi Zhang,[a] Xuan Feng,[a] Marta Wojonowska[a] and David O'Hagan*[a]

[a.]School of Chemistry, University of St Andrews, North Haugh, St Andrews, Fife, KY16 9ST, UK.

E-mail: do1@st-andrews.ac.uk

General experimental procedures.

All reagents were purchased from Sigma Aldrich or Alfa Aesar and used without further purification. All evaporations and concentrations were performed under reduced pressure by Büchi Rotavapor R-200.

^{19}F-NMR spectra were recorded on a Bruker 400, Bruker 500 or Bruker 700 instrument. D_2O was used as solvent. Chemical shifts are reported in parts per million (ppm) and coupling constants (J) are reported in Hertz (Hz).

LC-MS analysis was performed on the LCQ fleet ion-trap mass spectrometer (Thermo Fisher Scientific) with the electrospray ionisation (ESI) and the default method. Mobile phase A is MiliQ water with 0.05% (V/V) formic acid, and mobile phase B is acetonitrile with 0.05% (V/V) formic acid. Phenomenex Luna C-18 reverse phase column (4.5 X 50 mm) was used. The LC method was 0-1 min, 0% mobile phase B; 10 min, 95% mobile phase B; 12 min, 95% mobile phase B; 12.5 min, 0% mobile phase B; 14.5 min, 0% mobile phase B, 0.2 mL/min. The column was incubated at 40 °C. Detection wavelengths were 254, 262, 220 nm. The default tune was used for the mass detector.

Semipreparative HPLC purification was performed on a Shimadzu Prominence HPLC system fitted with a SIL-20A HT autosampler, LC-20 AT solvent delivery system, SPD-20 UV/vis detector using a Phenomenex Luna 5 µm, C-18 100A (250 × 10.00 mm) column and a guard cartridge. Mobile phase: 10 mM ammonia + 10 mM ammonium bicarbonate in water (solvent A) and acetonitrile (solvent B); Step-wise linear gradient: 0% B in 2 min, 48% B in 18 min, 95% in 25 min followed by equilibration of the column with initial condition; Flow rate of 3mL/min; Detection: 260 nm.

Growth of *Streptomyces aureorectus*, *Streptomyces virens* and *Streptomyces sulfonofaciens*

Spores of *Streptomyces aureorectus* DSM 41692 and *Streptomyces virens* DSM 41465 were grown on ISP4 agar plates (Per litre:10 g soluble starch, 1 g K_2HPO_4, 1 g $MgSO_4 \cdot 7H_2O$, 1 g NaCl, 1 g $(NH_4)_2SO_4$, 1 g $CaCO_3$, 20 g agar) and spores of *Streptomyces sulfonofaciens* were grown on ISP2 agar plate (Per litre: 4 g yeast extract, 10 g malt extract, 4 g dextrose, 20 g agar). The plates were incubated at 30 °C for 6 days. Pre-cultures of *Streptomyces aureorectus* DSM 41692 and *Streptomyces virens* DSM 41465 were prepared in 50 ml TSBY media (Per litre: 30 g trypticase soy broth, 3 g yeast extract, 0.25 g NH_4Cl, 13 g NaCl, 4 g $MgCl_2 \cdot 7H_2O$, 3.45 g $MgSO_4 \cdot 7H_2O$, 0.34 g KCl, 0.14 g $CaCl_2 \cdot 2H_2O$) in 250 mL conical flasks shaking at 30 °C at 180 rpm for 2 days. Pre-culture of *Streptomyces sulfonofaciens* DSM 41679 was prepared in 50 ml ISP2 media (Per litre: 4 g yeast extract, 10 g malt extract, 4 g dextrose).

Fermentation of *Streptomyces aureorectus*, *Streptomyces virens* and *Streptomyces sulfonofaciens*

An aliquot (2ml) each pre-culture was transformed into 500 ml flasks containing 100 ml production media (Per litre: 12.5 g corn steep liquor, 10 g mannitol, 2 g NaCl, 2 g $(NH_4)_2PO_4$, 1.5 g KH_2PO_4, 0.25 g MgSO4 · 7H2O, 1ml Hoagland's salt solution, 7.5 ml 0.5 M KF). The Hoagland's salt solution contained deionised water (1 L), manganese(II) chloride tetrahydrate (0.389 g), phosphorous acid (0.611 g), copper(II) sulfate (0.056 g), ammonium molybdate tetrahydrate (0.056 g), nickel(II) sulfate hexahydrate (0.056 g), zinc sulfate heptahydrate (0.056 g), aluminium sulfate (0.056 g), stannous chloride dihydrate (0.028 g), cobalt(II) nitrate hexahydrate (0.056 g), titanium dioxide (0.056 g), lithium chloride (0.028 g), potassium iodide (0.028 g) and potassium bromide (0.028 g). [1] Sterilised by autoclaving. The flasks were incubated in the incubator shaker at 30 °C at 180 rpm.

Extraction of fluorometabolites

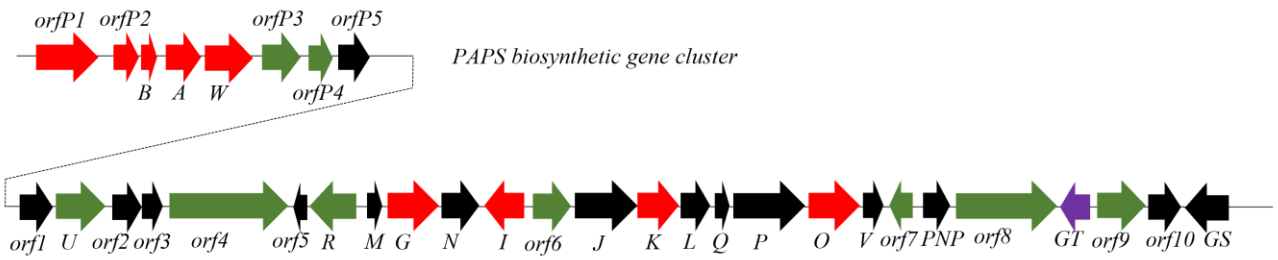
The cultures were harvested after 6 to 8 days' incubation. After centrifugation, the supernatant was extracted with 20% n-butanol. The n-butanol layer was concentrated and re-dissolved in D_2O and analysed directly by $^{19}F\{^1H\}$-NMR.

Add-mix experiments

Each add-mix experiment contained two components: (1) extract of *S. aureorectus* DSM 41692; (2) synthetic 4'-fluoroadenosine. The mixture was analysed directly by $^{19}F\{^1H\}$-NMR.

NucGT assays

NucGT assays were carried out in 50 mM Tris-HCl buffer, pH = 8.0, with 10 mg/mL UDP-glucose, 100 mM $MgCl_2$, 1 mM substrate and 1.5 nM glucosyltransferase. The reactions were incubated on a heatblock at 30 °C and the enzyme was denaturized by adding chloroform. The reaction was analysed directly by $^{19}F\{^1H\}$-NMR.[1]



Nucleocidin biosynthetic gene cluster in *S. calvus*.

OrfP1	**Ferredoxin-sulfite reductase**	***nucN***	**Amidinotransferase**
OrfP2	**PAPS reductase**	***nucI***	**Sulfatase**
nucB	**Adenylyl-sulfate kinase**	***Orf6***	**StrR-transcriptional regulator**
nucA	**Sulfate adenylyltransferase CysD**	***nucJ***	**Radical SAM/B_{12} domain superfamily**
nucW	**Sulfate adenylyltransferase subunit**	***nucK***	**Sulfotransferase domain**
OrfP3	**Sulfonate ABC transporter**	***nucL***	**SAM-Me-transferase**
OrfP4	**ABC transporter ATP-binding protein**	***nucQ***	**Rubrerythrin**
OrfP5	**ABC transporter permease**	***nucP***	**SAM-Methyltransferase**
Orf1	**Oxidoreductase**	***nucO***	**SecC motif- protein sulfotransferase**
nucU	**Cation: H^+ antiporter**	***nucV***	**Aden phosphoribosyl-transferase**
Orf2	**Hypothetical protein**	***Orf7***	**LuxR family transcription regulator**
Orf3	**Histidine phosphatase family protein**	***PNP***	**5'-SMe-adenosine phosphorylase**
Orf4	**Transcriptional regulatory protein**	***Orf8***	**Lycopene cyclase**
Orf5	**Aminoglycoside phosphotransferase**	***nucGT***	**Glucosyltransferase**
nucR	**Metabolite transport protein YhjE**	***Orf9***	**Hypothetical protein**
nucM	**Hypothetical protein**	***Orf10***	**Protein kinase domain- protein**
nucG	**Sulfatase**	***nucGS***	**β-glucosidase**

Figure S1. The biosynthetic gene cluster of nucleocidin in *S. calvus T-3018* with annotations.

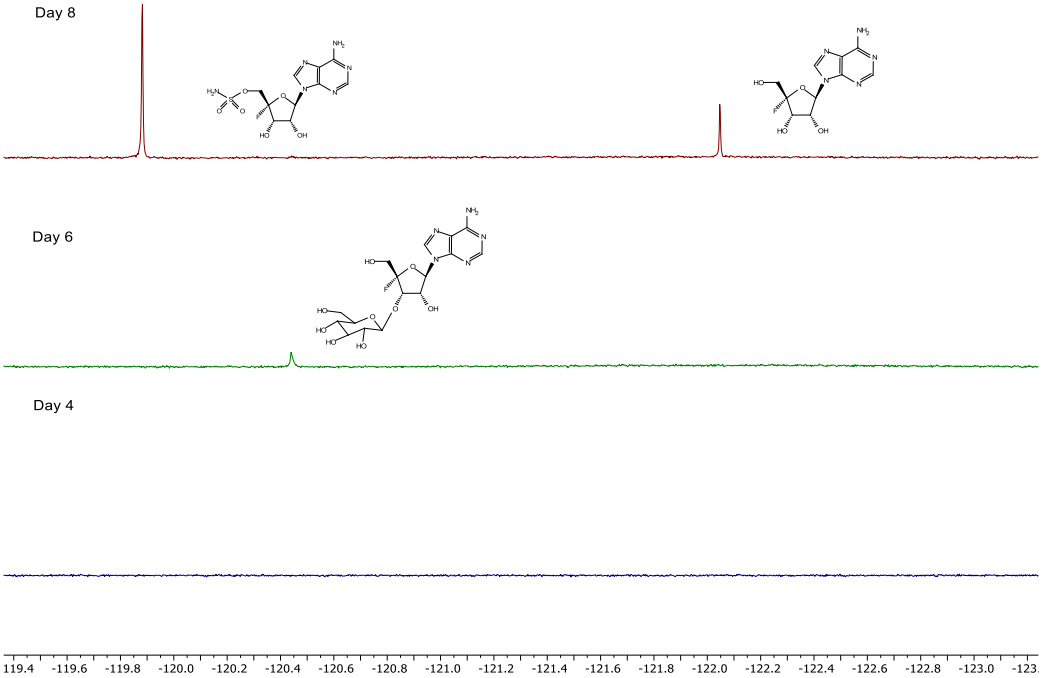


Figure S2. $^{19}F\{^{1}H\}$-NMR time course of fluorometabolites production in *S. aureorectus* DSM 41692 over 8 days (376 MHz, Deuterium Oxide).

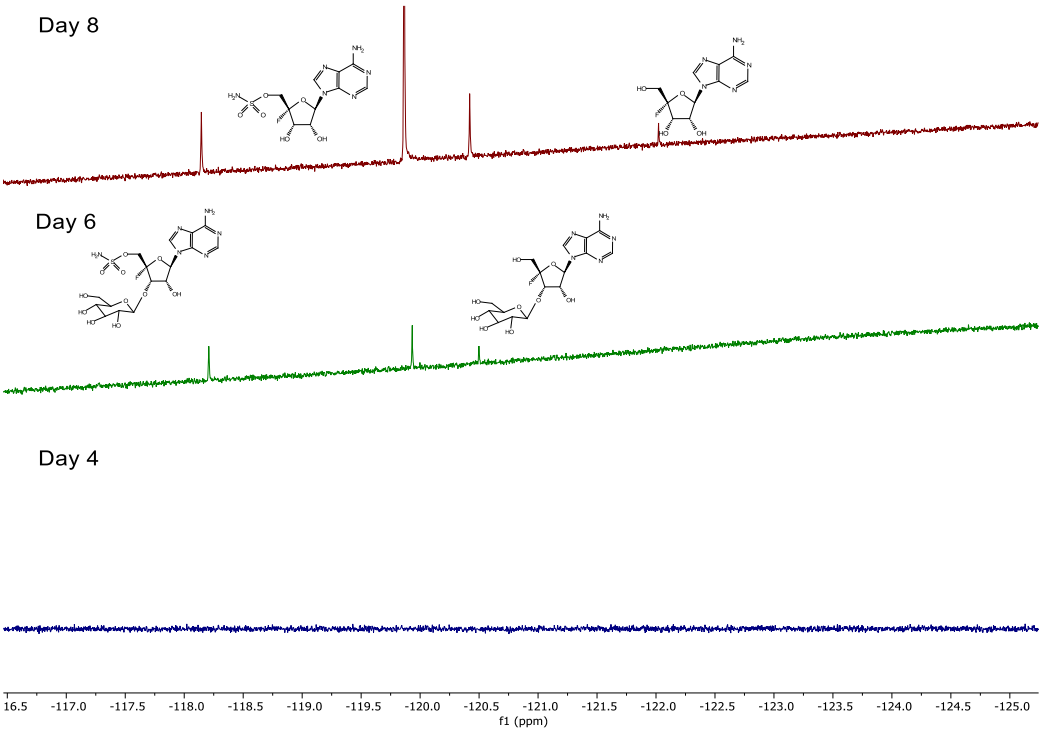


Figure S3. $^{19}F\{^{1}H\}$-NMR time course of fluorometabolite production in *S. virens* DSM 41465 over 8 days (376 MHz, Deuterium Oxide).

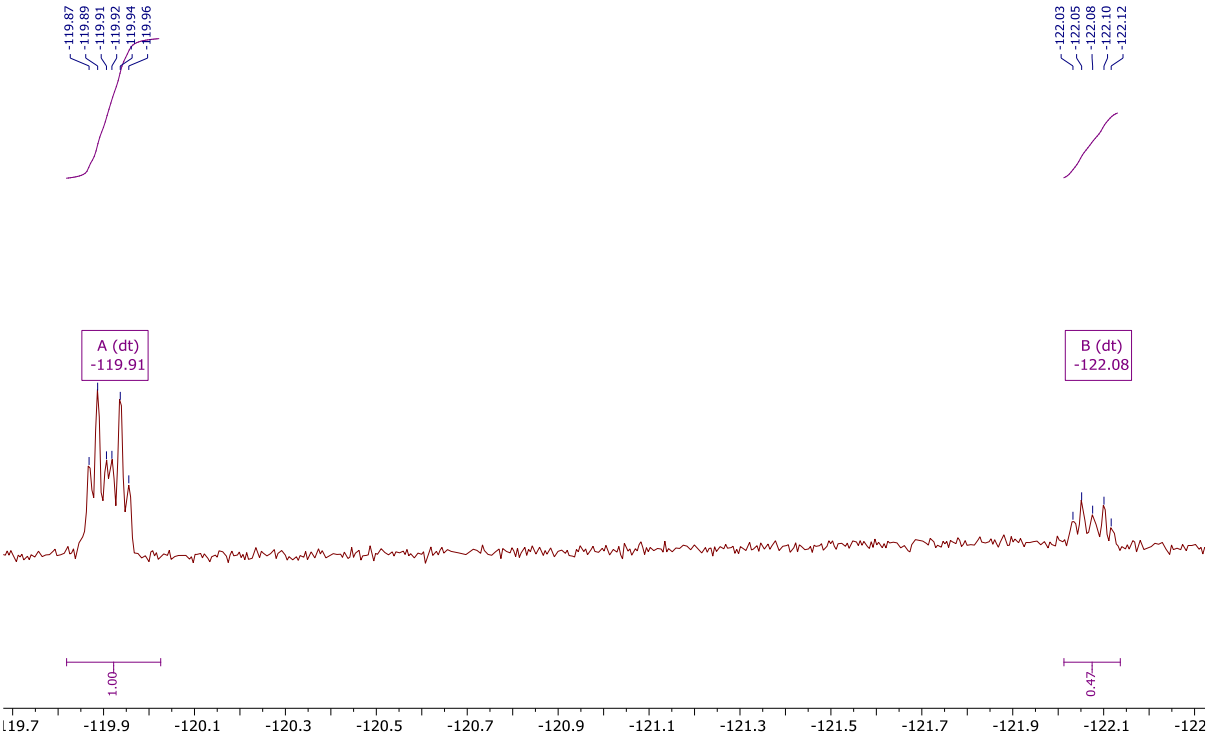


Figure S4. ^{19}F NMR of the products of *S. aureorectus* DSM 41692 on day 8. (376 MHz, Deuterium Oxide) δ -119.91 (dt, $^{3}J_{HF}$ = 18.5, 7.1 Hz), -122.08 (dt, $^{3}J_{HF}$ = 15.4, 6.6 Hz).

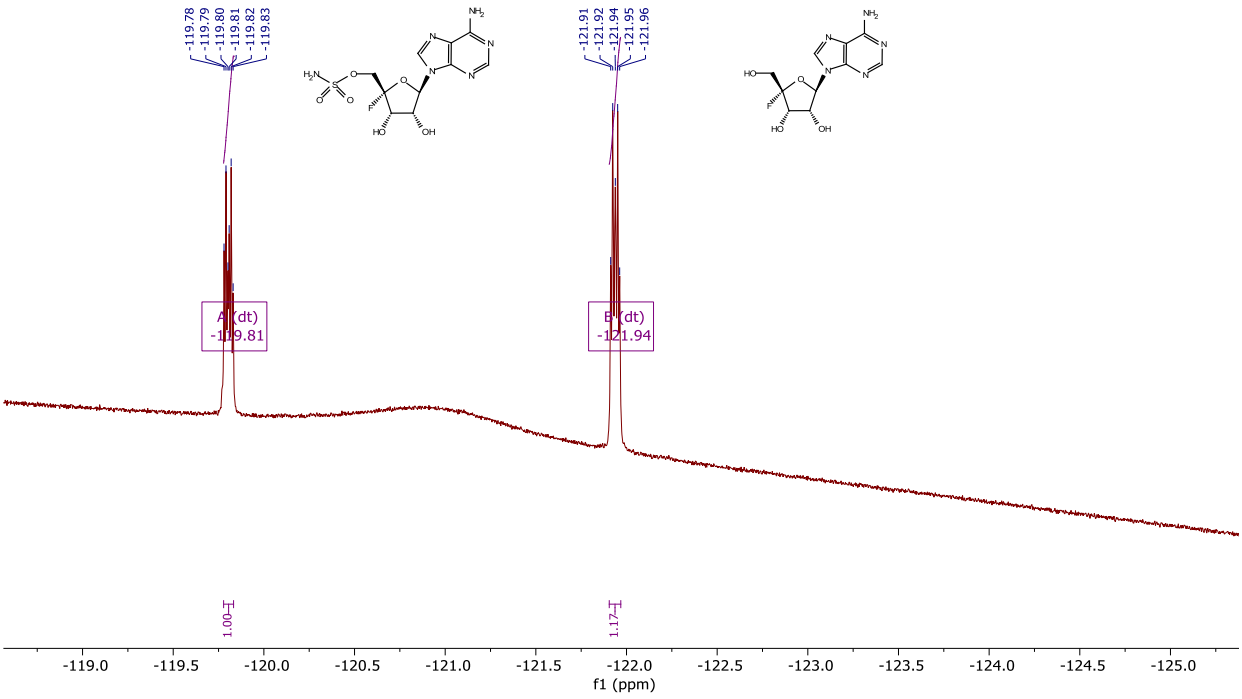


Figure S5. ^{19}F-NMR of the products after spike-in experiment. (659 MHz, Deuterium Oxide) δ -119.81 (dt, $^{3}J_{HF}$ = 18.7, 7.2 Hz), -121.94 (dt, $^{3}J_{HF}$ = 15.8, 7.3 Hz).

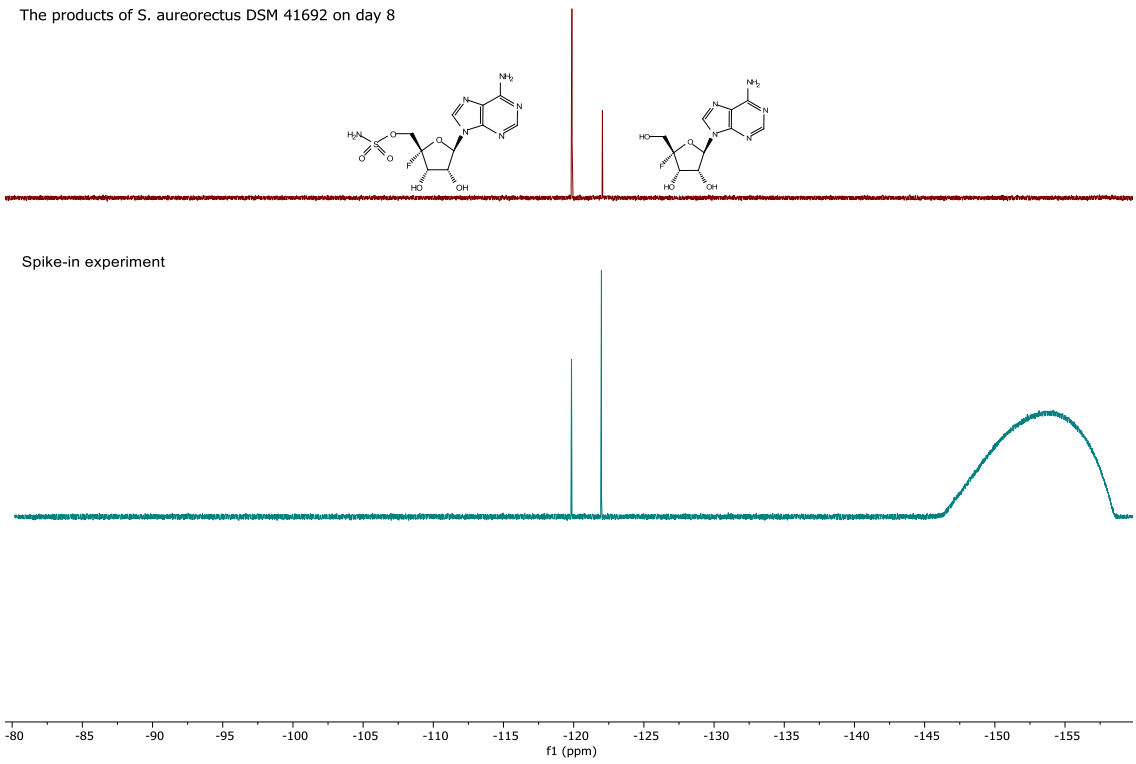


Figure S6. Stacked ^{19}F NMR of the products of *S. aureorectus* DSM 41692 on day 8 (Top - Red) and the products after spike-in experiment (Bottom - blue).

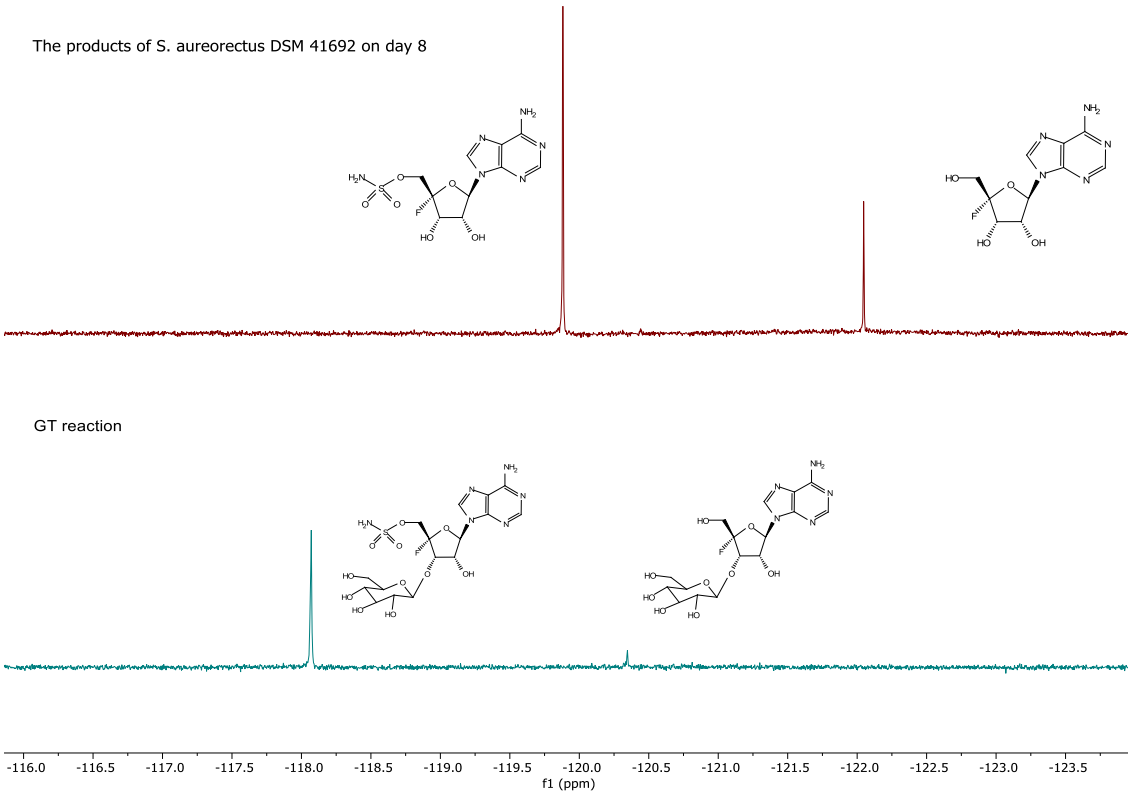
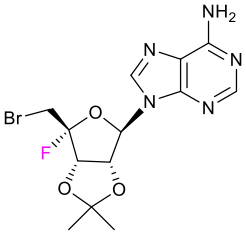


Figure S7. Stacked $^{19}F\{^{1}H\}$-NMR of the products of *S. aureorectus* DSM 41692 on day 8 (Top - Red) and the products after GT assay (Bottom - blue).



5'-Bromo-4'-fluoro-2',3'-*O*-isopropylideneadenosine 11. 5'-Iodo-4'-fluoro-2',3'-*O*-isopropylideneadenosine **10** (186 mg, 0.43 mmol), prepared according to literature[1], and sodium bromide (442 mg, 4.30 mmol, 10 eq) was added to anhydrous DMF (10 mL). The mixture was stirred at 97 °C for 48 h. After removal of DMF under reduced pressure, the residue was participated between water and dichloromethane. The aqueous layer was extracted with dichloromethane for another three times. The combined organic layers were dried over anhydrous magnesium sulfate and purified by column chromatography (silica gel, 5% MeOH in DCM) followed by preparative TLC (DCM: acetone 90:10) to give the product as pale white amorphous solid (86 mg, 51%). ^{1}H NMR ($CDCl_3$, 400 MHz) 8.36 (s, 1H, H-2), 7.89 (s, 1H, H-8), 6.34 (s, 1H, H-1'), 5.79 (br s, 2H, NH_2), 5.50 (dd, $^{3}J_{FH}$ =12.0 Hz, $^{3}J_{HH}$ 6.5 Hz, H-3'), 5.28 (d, $^{3}J_{HH}$ 6.5 Hz, H-2'), 3.74 (dd, $^{2}J_{HH}$ = 12.0 Hz, $^{3}J_{FH}$ 15.8 Hz, H-5'a), 3.74 (dd, $^{2}J_{HH}$ = 12.0 Hz, $^{3}J_{FH}$ =13.0 Hz, H-5'b), 1.68 (s, 3H, CH_3), 1.44 (s, 3H, CH_3). ^{13}C NMR ($CDCl_3$, 125 MHz) 155.8 (C-6), 153.3 (C-2), 149.1 C-4), 139.6 (C-8), 120.4 (C-5), 117.0 (C of isopropylidene), 114.9 (d, $^{1}J_{CF}$ = 239 Hz, C-4'), 88.9 (C-1'), 83.3 (C-2'), 82.0 (d, $^{2}J_{CF}$ = 20.5 Hz, C-3'), 31.6 (d, , $^{2}J_{CF}$ = 36.0 Hz, C-5'), 25.8 (CH_3), 25.9 (CH_3) . ^{19}F NMR (376 MHz, $CDCl_3$) δ_F -107.2 (ddd, $^{3}J_{HF}$ = 13.0, 15.5, 15.8 Hz); HRMS (ESI$^+$) 388.0406 $[M+H]^+$, $C_{13}H_{16}BrFN_5O_3$ requires 388.0415.

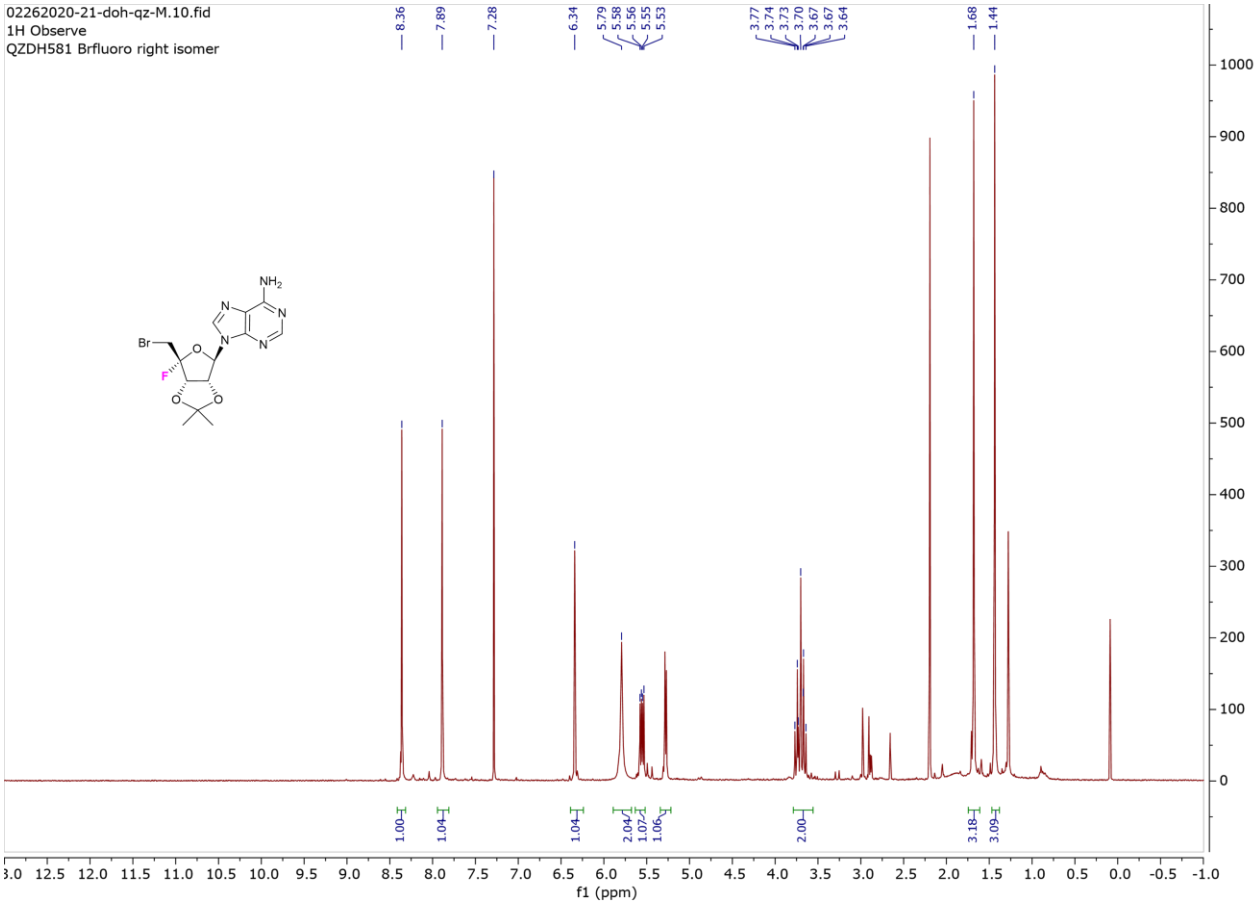


Figure S8. ^{1}H NMR of 5'-Bromo-4'-fluoro-2',3'-*O*-isopropylideneadenosine **11**.

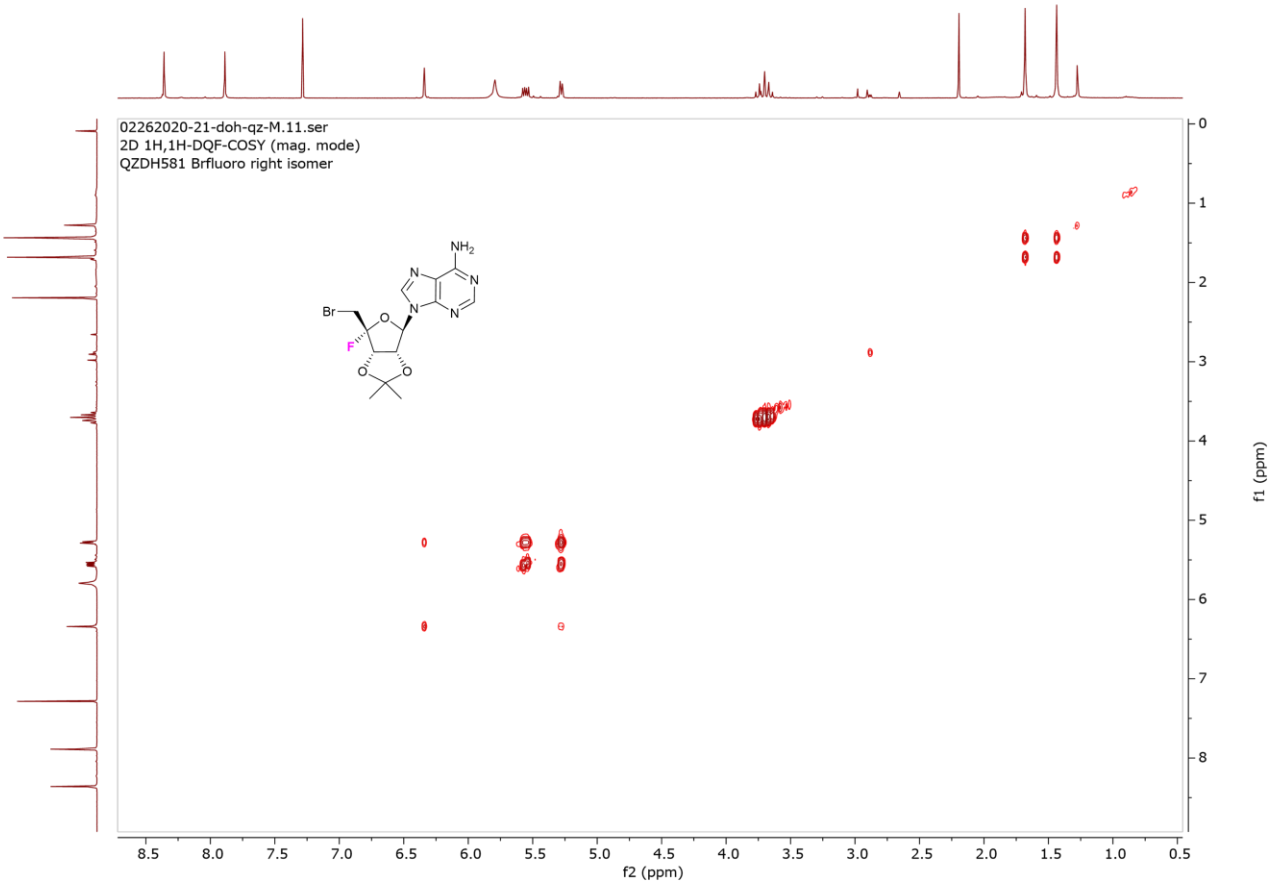


Figure S9. ^{1}H-^{1}H-COSY of 5'-Bromo-4'-fluoro-2',3'-*O*-isopropylideneadenosine **11**.

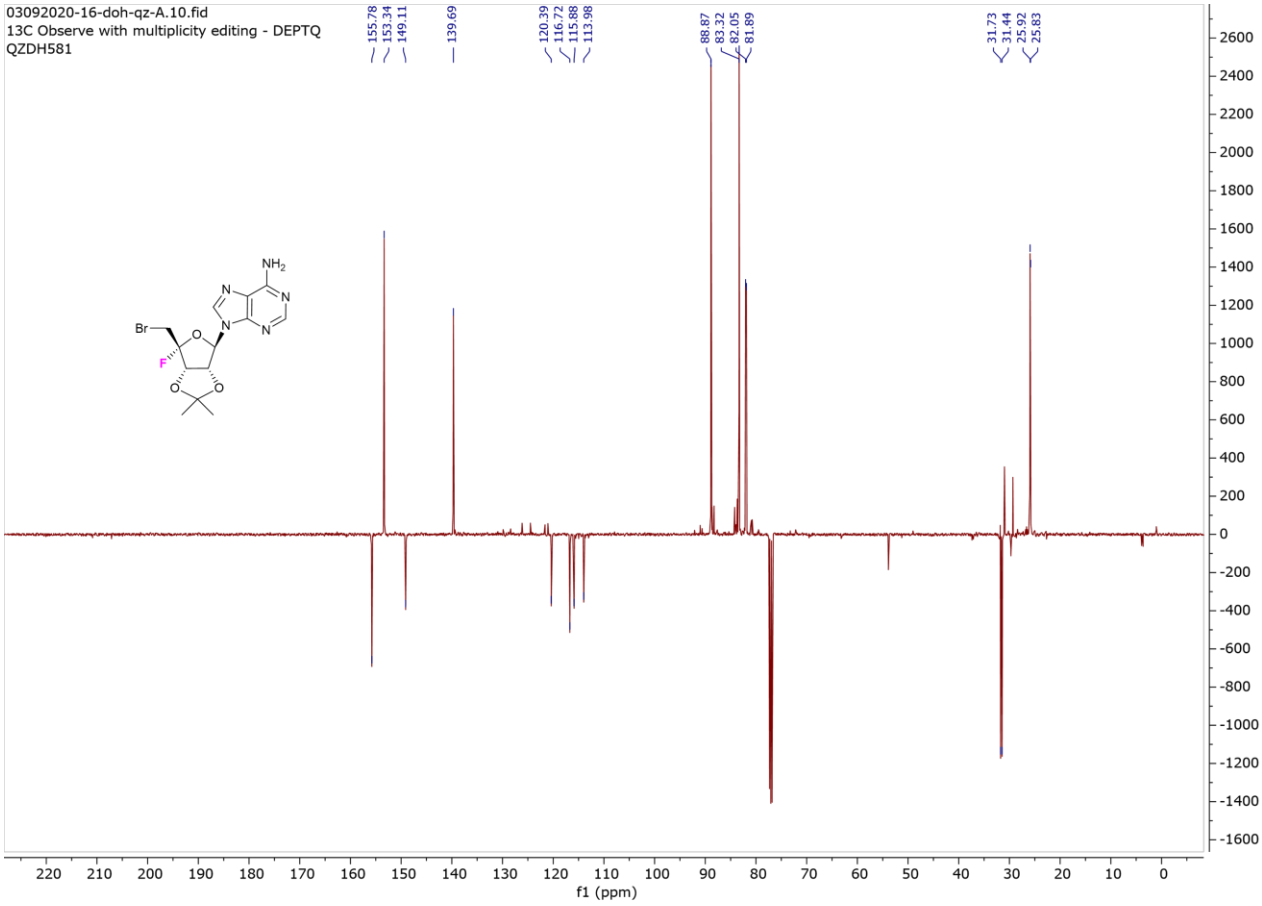


Figure S10. ^{13}C NMR of 5'-Bromo-4'-fluoro-2',3'-*O*-isopropylideneadenosine **11**.

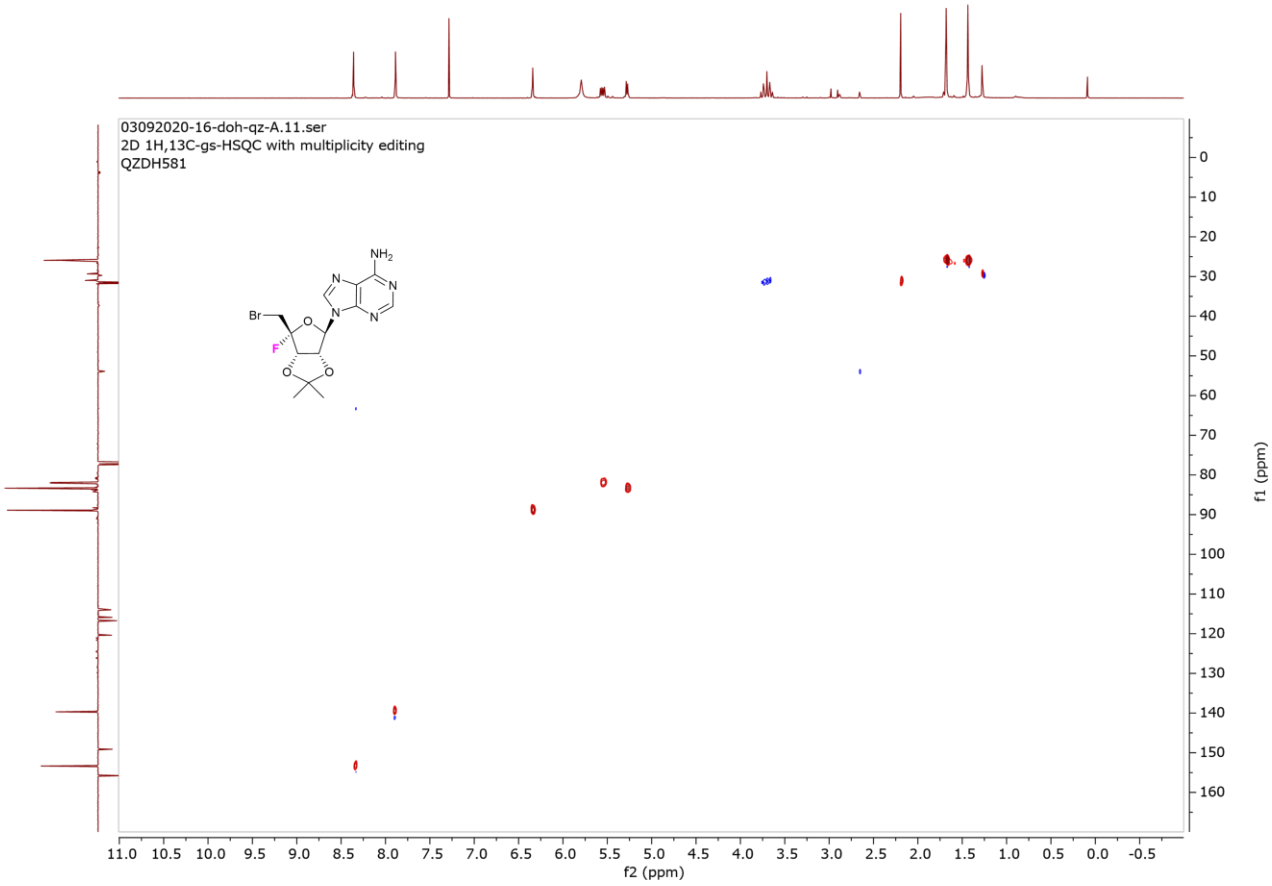


Figure S11. ^{1}H-^{13}C HSQC of 5'-Bromo-4'-fluoro-2',3'-*O*-isopropylideneadenosine **11**.

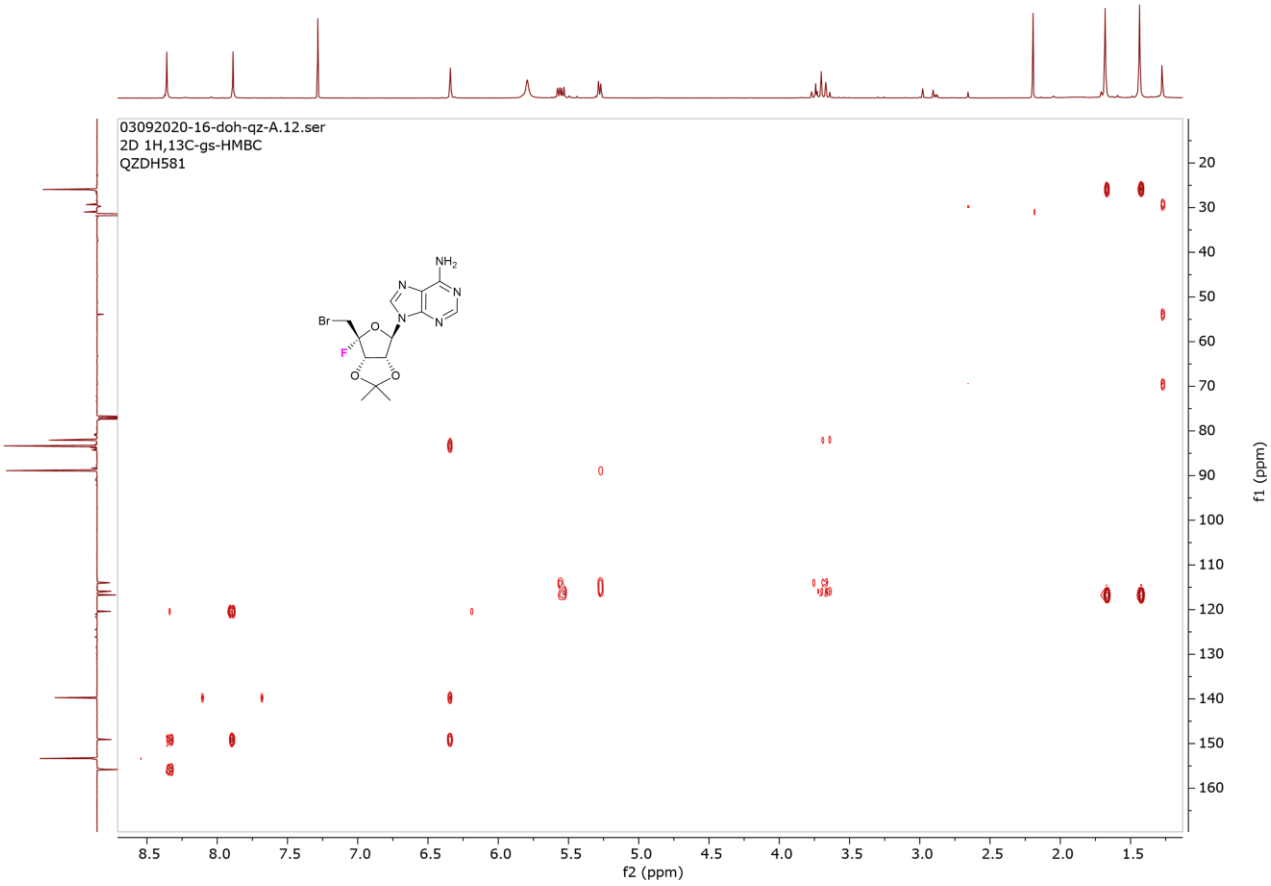


Figure S12. ^{1}H-^{13}C HMBC of 5'-Bromo-4'-fluoro-2',3'-*O*-isopropylideneadenosine **11**.

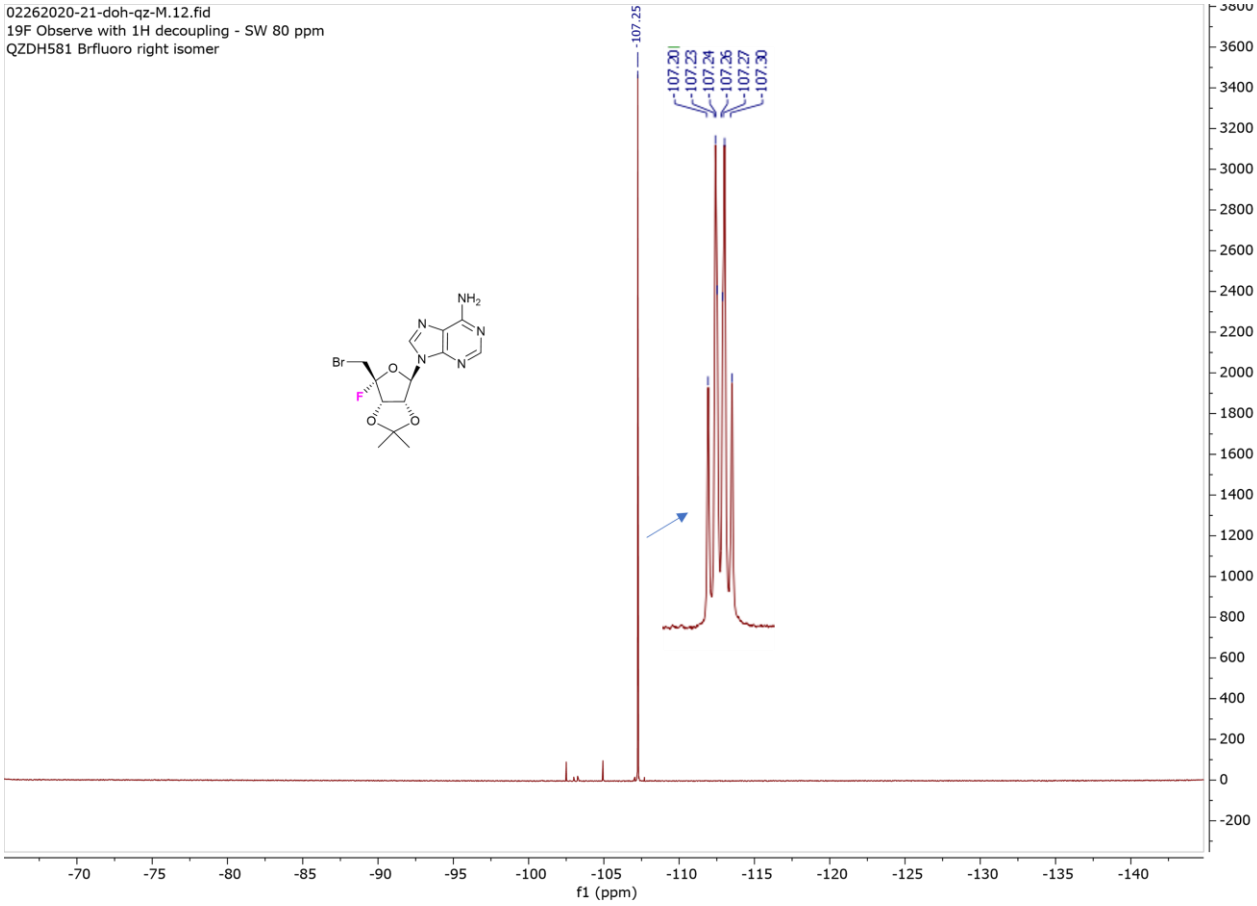
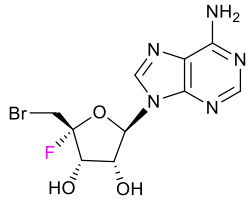


Figure S13. ^{19}F NMR of 5'-Bromo-4'-fluoro-2',3'-*O*-isopropylideneadenosine **11**.



5'-Bromo-4'-fluoroadenosine 12. 5'-Bromo-4'-fluoro-2',3'-*O*-isopropylideneadenosine (26 mg, 0.067 mmol) was dissolved in 1 ml of 60% trifluoracetic acid in water at 0°C. The reaction mixture was allowed to warm up to rt and stirring was continued until all starting material consumed as monitored by TLC, ca 4h. The solvent was removed under reduced pressure. The residue was purified by C18 reversed phase cartridge to give the product as a white solid (15 mg, 64%). ^{1}H NMR (MeOD, 400 MHz) 8.28 (s, 1H, H-8), 8.23 (s, 1H, H-2), 6.30 (d, 1H, $^{3}J_{HH}$ = 2.4 Hz, H-1'), 5.03 (dd, $^{3}J_{FH}$ =17.6 Hz, $^{3}J_{HH}$ 6.7Hz, H-3'), 4.81 (d, $^{3}J_{HH}$ 6.7, 2.4 Hz, H-2'), 3.80 (dd, $^{2}J_{HH}$ = 11.6 Hz, $^{3}J_{FH}$ 8.0 Hz, H-5'a), 3.74 (dd, $^{2}J_{HH}$ = 11.6 Hz, $^{3}J_{FH}$ =14.4 Hz, H-5'b). ^{13}C NMR (MeOD, 100 MHz) 156.0 (C-6), 152.8 (C-2), 149.0 (C-4), 140.3 (C-8), 119.2 (C-5), 115.3 (d, $^{1}J_{CF}$ = 232.4 Hz, C-4'), 91.2 (C-1'), 72.1 (C-2'), 71.1 (d, $^{2}J_{CF}$ = 20.3 Hz, C-3'), 29.9 (d, $^{2}J_{CF}$ = 38.6 Hz,C-5'). ^{19}F NMR (376 MHz, MeOD) δ_F -115.1 (ddd, $^{3}J_{HF}$ = 8.0, 14.4, 17.6 Hz); HRMS (ESI^{+}) 348.0098 [M+H]$^{+}$, $C_{10}H_{11}{}^{79}BrFN_5O_3$ requires 348.0102.

Figure S14. ^{1}H NMR of 5'-Bromo-4'-fluoroadenosine **12**.

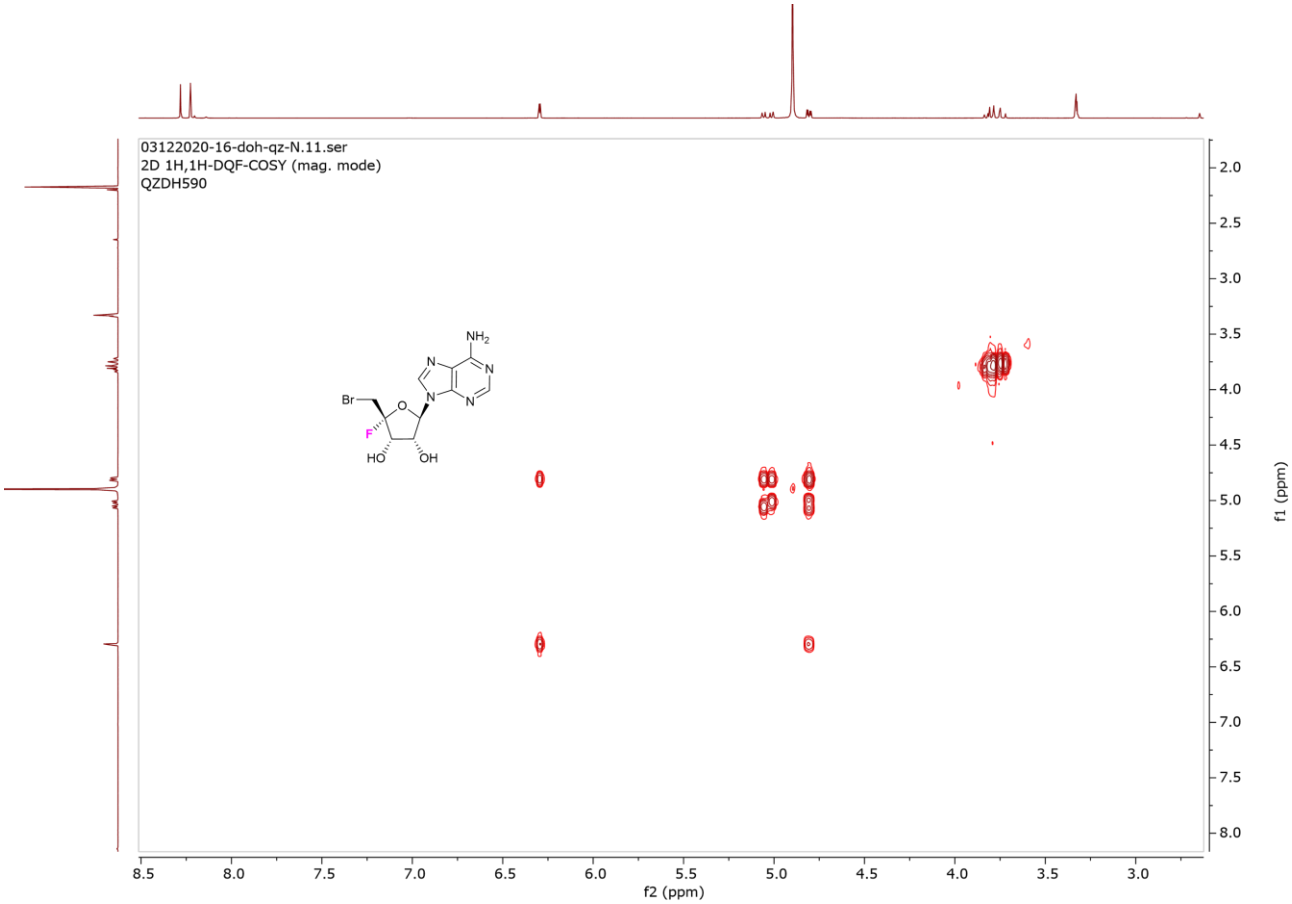


Figure S15. ^{1}H-^{1}H-COSY of 5'-Bromo-4'-fluoroadenosine **12**.

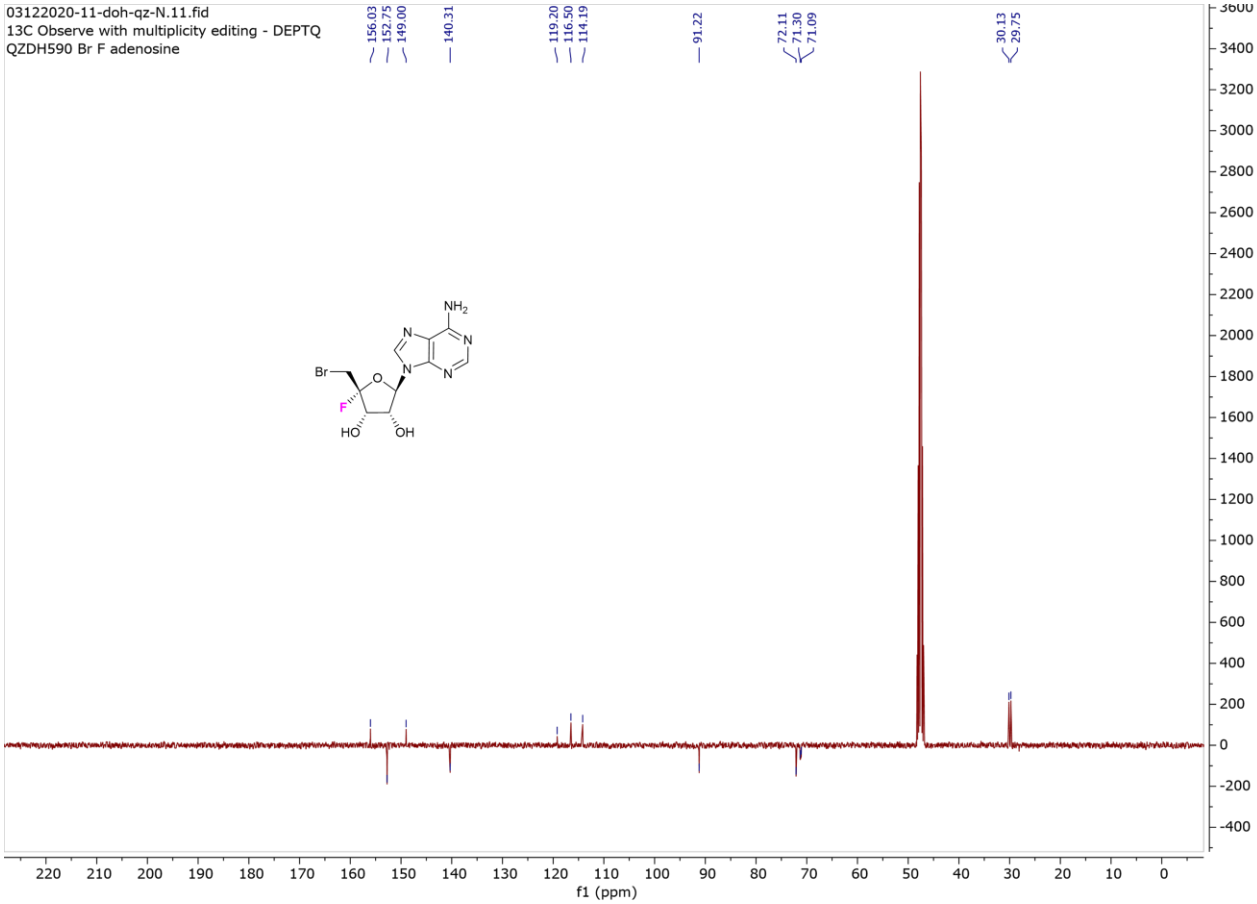


Figure S16. ^{13}C NMR of 5'-Bromo-4'-fluoroadenosine **12**.

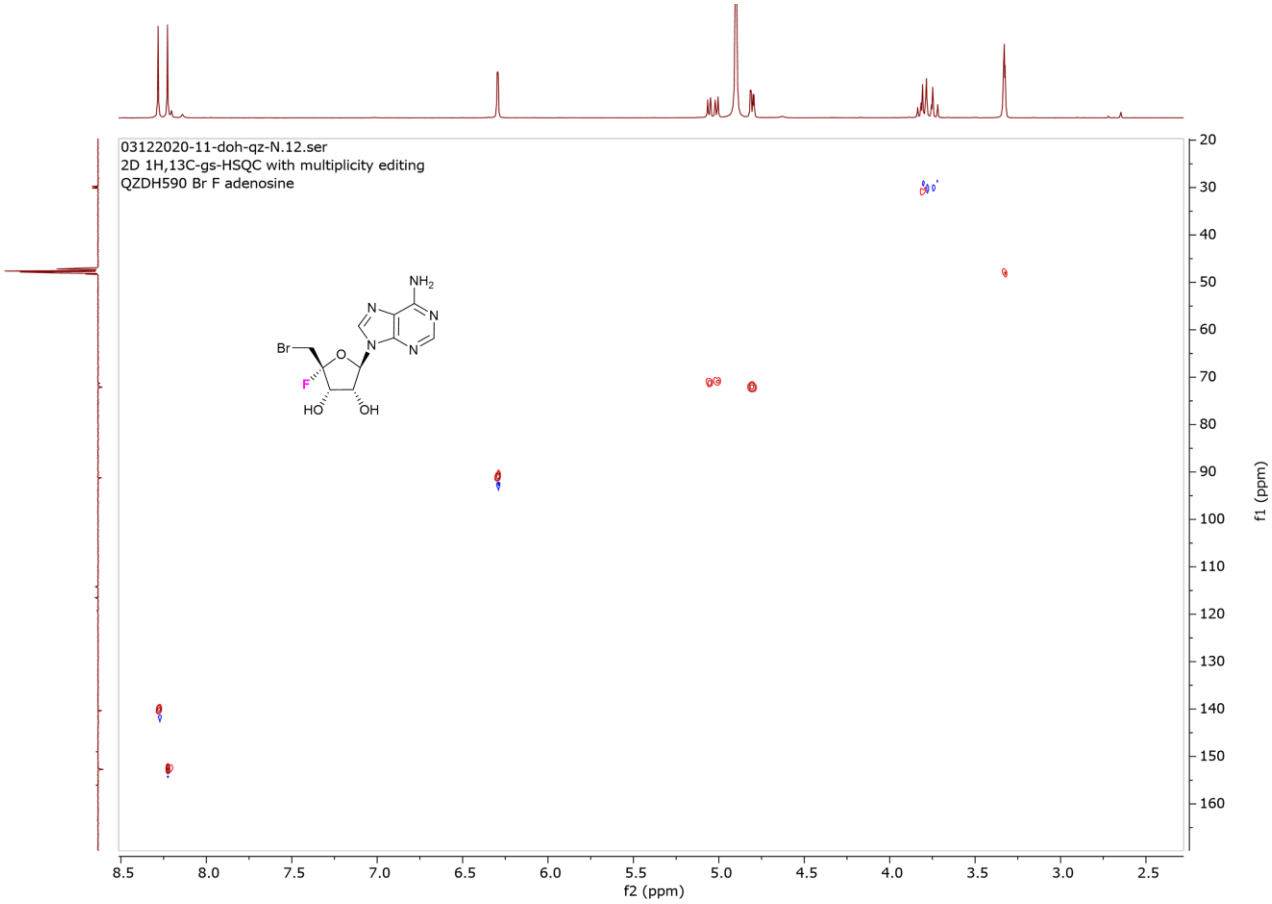


Figure S17. ^{1}H-^{13}C HSQC of 5'-Bromo-4'-fluoroadenosine **12**.

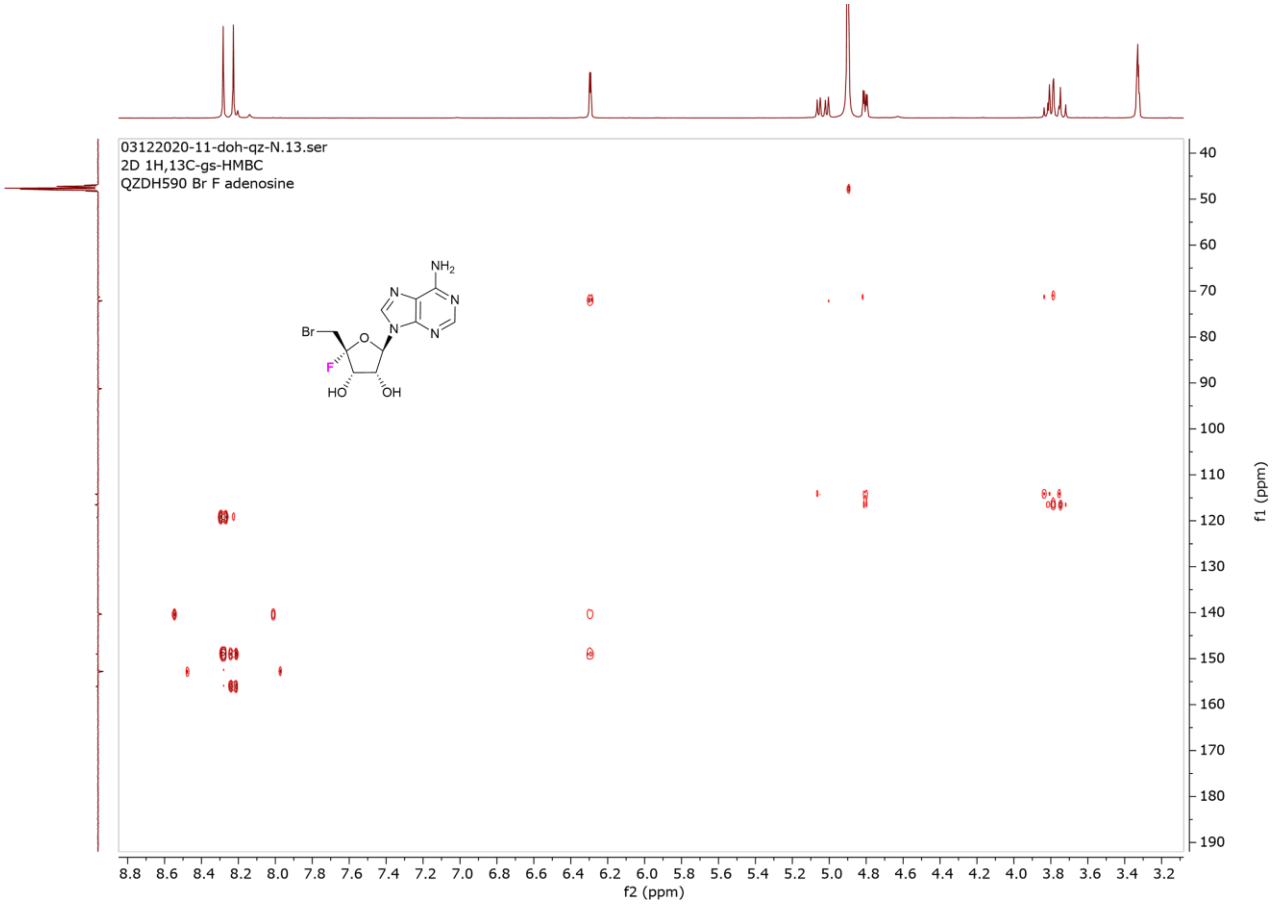


Figure S18. ^{1}H-^{13}C HMBC of 5'-Bromo-4'-fluoroadenosine **12**.

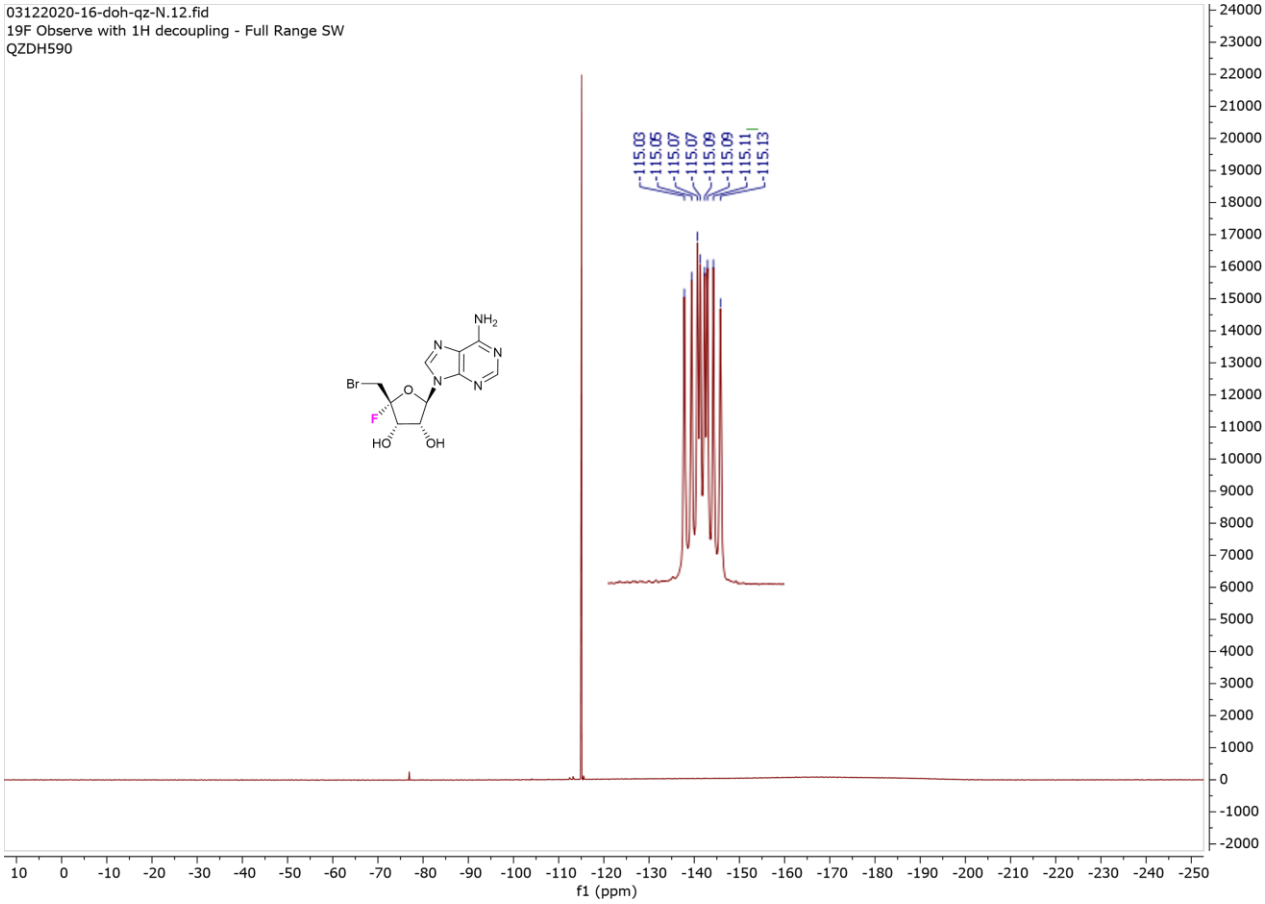
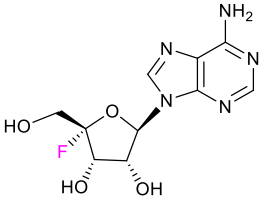


Figure S19. ^{19}F NMR of 5'-Bromo-4'-fluoroadenosine **12**.



4'-Fluoroadenosine 9. 5'-Bromo-4'-fluoro-adenosine **12** (10 mg, 0.029 mmol) and potassium acetate (200 mg, 2.04 mmol, 70 eq) was suspended in 3 ml of anhydrous DMF. The mixture was heated at 97 °C for 48 h, cooled to room temperature, and filtered through a pad of celite followed by washing with acetonitrile. The filtrate was rotary evaporated under reduced pressure. The residue was taken up in MeOD and treated with sodium carbonate (31 mg, 0.29 mmol, 10 eq) by stirring at rt for 30 min. The mixture was then separated by preparative HPLC to give the titled compound **9** (1.6 mg, 20%). ^{1}H NMR (MeOD, 700 MHz) 8.33 (s, 1H, H-8), 8.22 (s, 1H, H-2), 6.33 (d, 1H, $^{3}J_{HH}$ = 2.9 Hz, H-1'), 4.78 (dd, $^{3}J_{FH}$ =15.7 Hz, $^{3}J_{HH}$ 6.3 Hz, H-3'), 4.66 (d,1H, $^{3}J_{HH}$ = 6.3, 2.9 Hz, H-2'), 3.79 (2 s, overlapped, H-5'a and 5'b). ^{13}C NMR (MeOD, 175 MHz) 156.1 (C-6), 152.6 (C-2), 148.8 (C-4), 139.9 (C-8), 119.2 (C-5), 117.3 (d, $^{1}J_{CF}$ = 224.0 Hz, C-4'), 90.9 (C-1'), 72.7 (C-2'), 69.5 (d, $^{2}J_{CF}$ = 19.9 Hz, C-3'), 60.7 (d, $^{2}J_{CF}$ = 42.8 Hz,C-5'). ^{19}F NMR (657 MHz, MeOD) δ_F -124.2 (dt, $^{3}J_{HF}$ = 15.7, 4.8 Hz); HRMS (ESI^{+}) 286.0947 [M+H]$^{+}$, $C_{10}H_{13}FN_5O_4$ requires 286.0949.

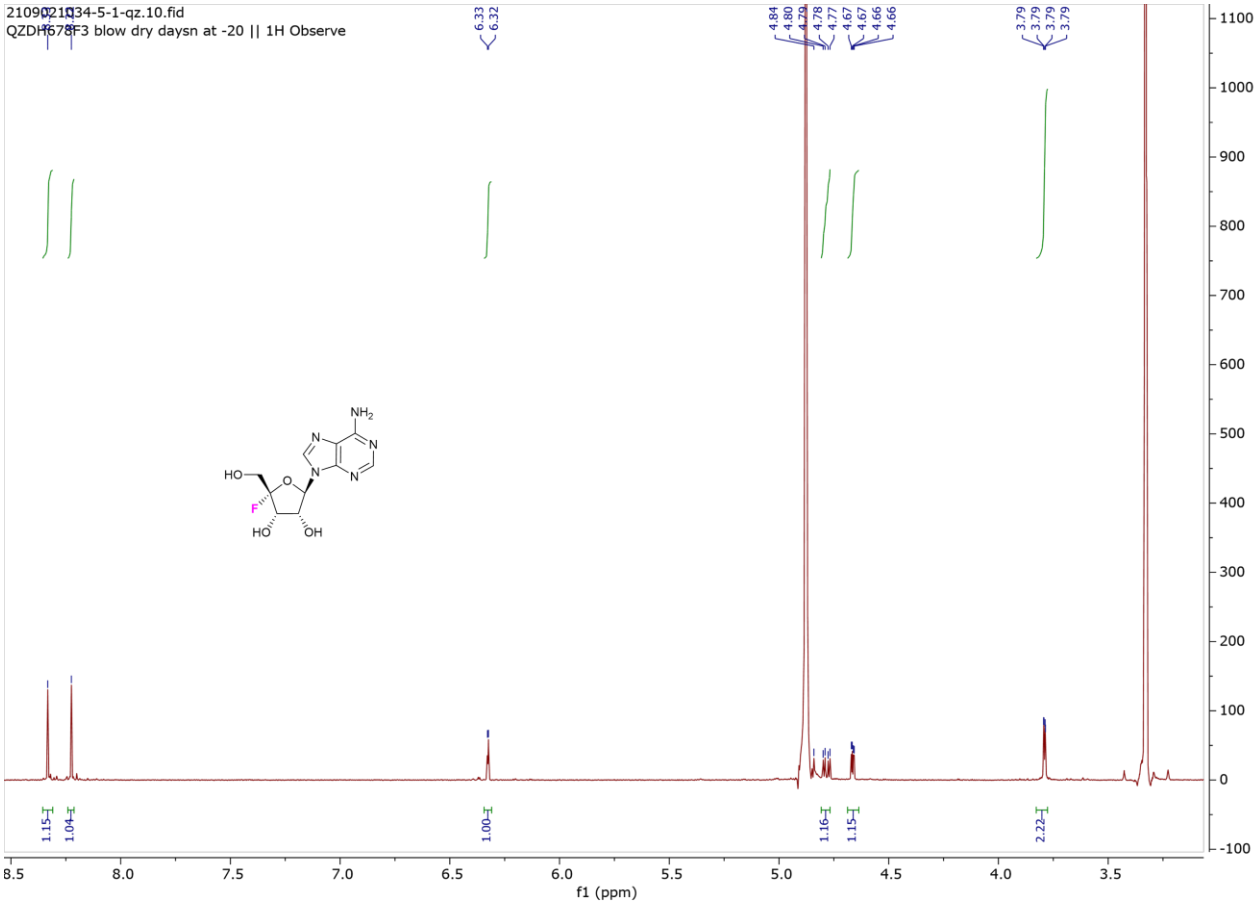


Figure S20. ^{1}H NMR of 4'-Fluoroadenosine **9**.

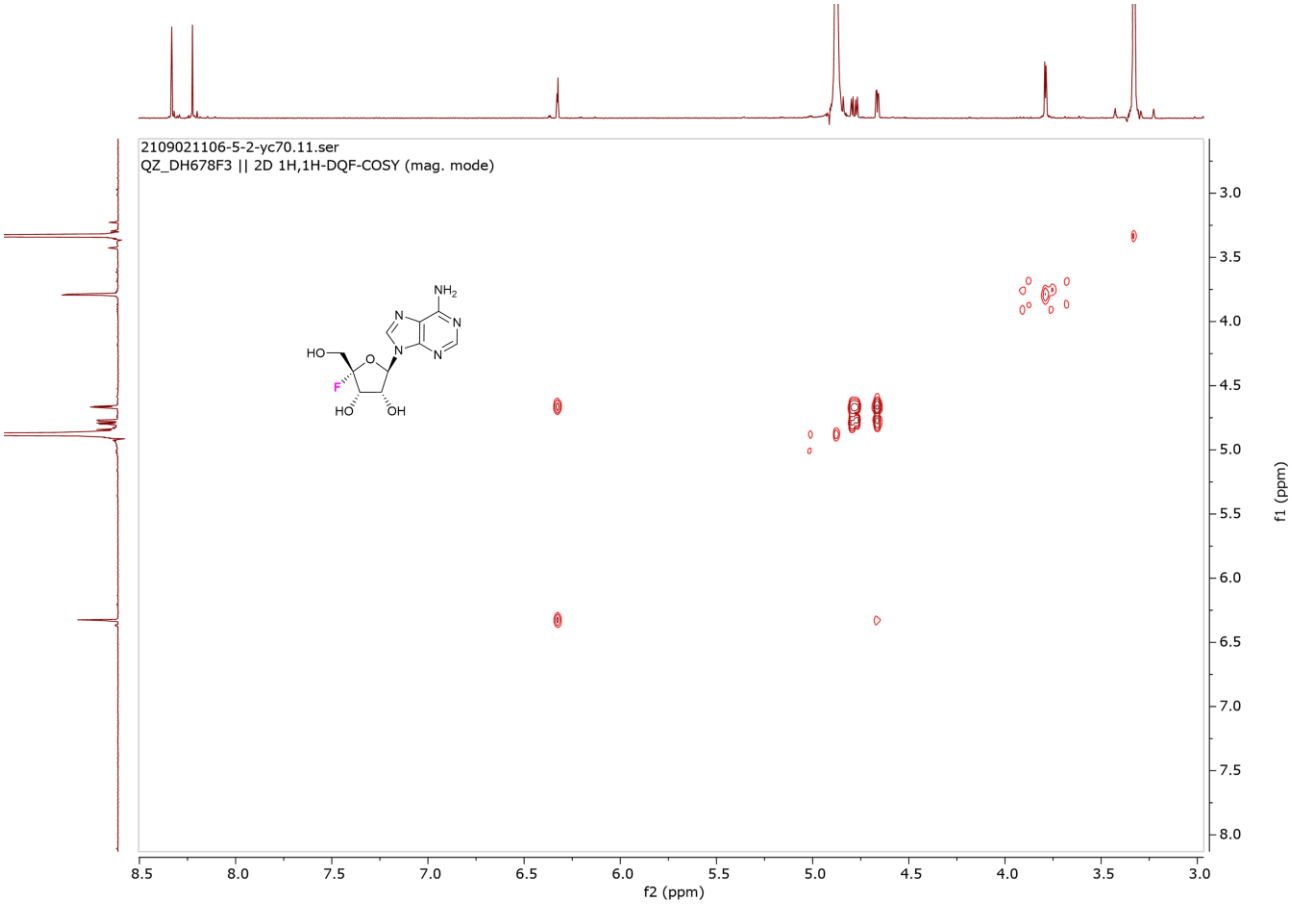


Figure S21. ^{1}H-^{1}H-COSY of 4'-Fluoroadenosine **9**.

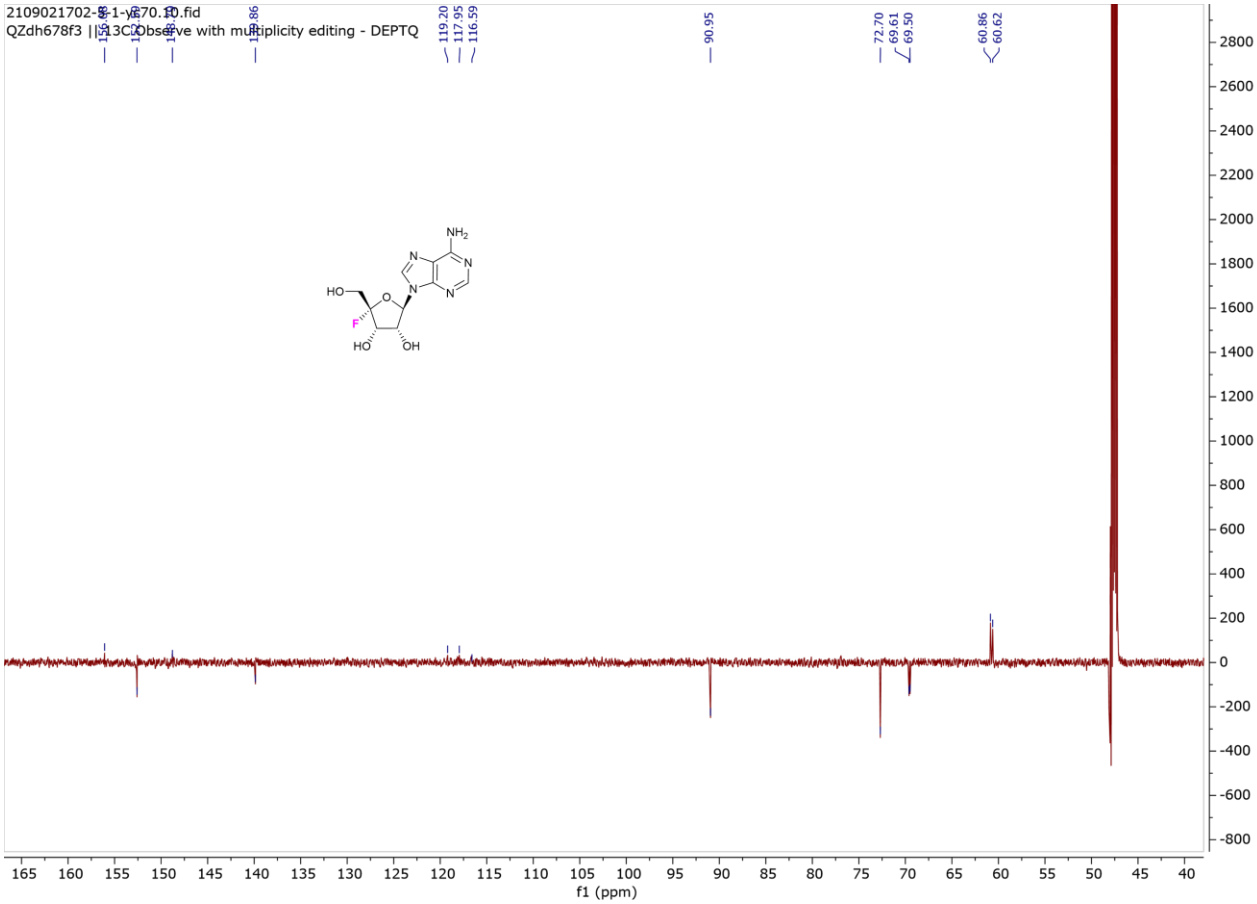


Figure S22. ^{13}C NMR of 4'-Fluoroadenosine **9**.

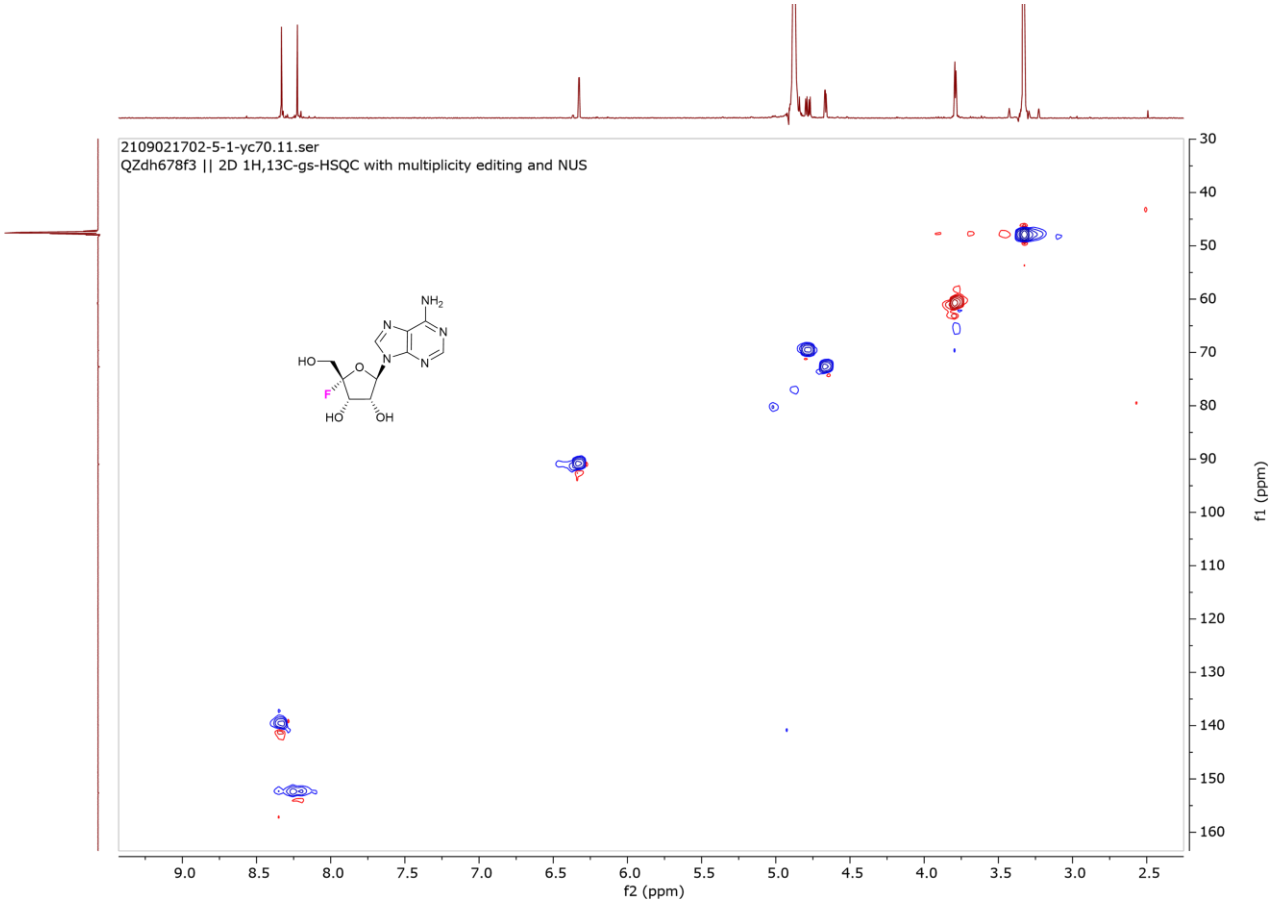


Figure S23. ^{1}H-^{13}C-HSQC of 4'-Fluoroadenosine **9**.

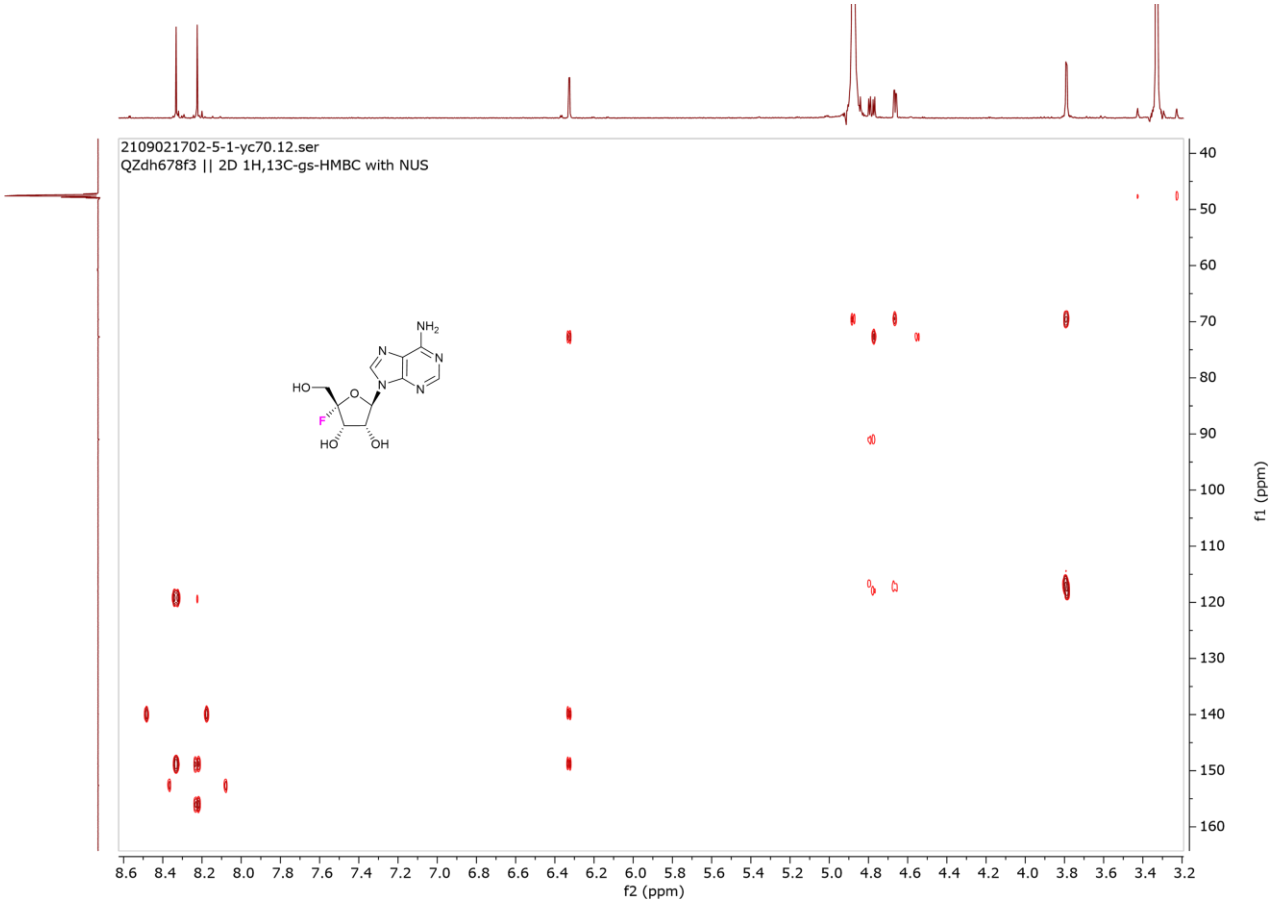


Figure S24. ^{1}H-^{13}C-HMBC of 4'-Fluoroadenosine **9**.

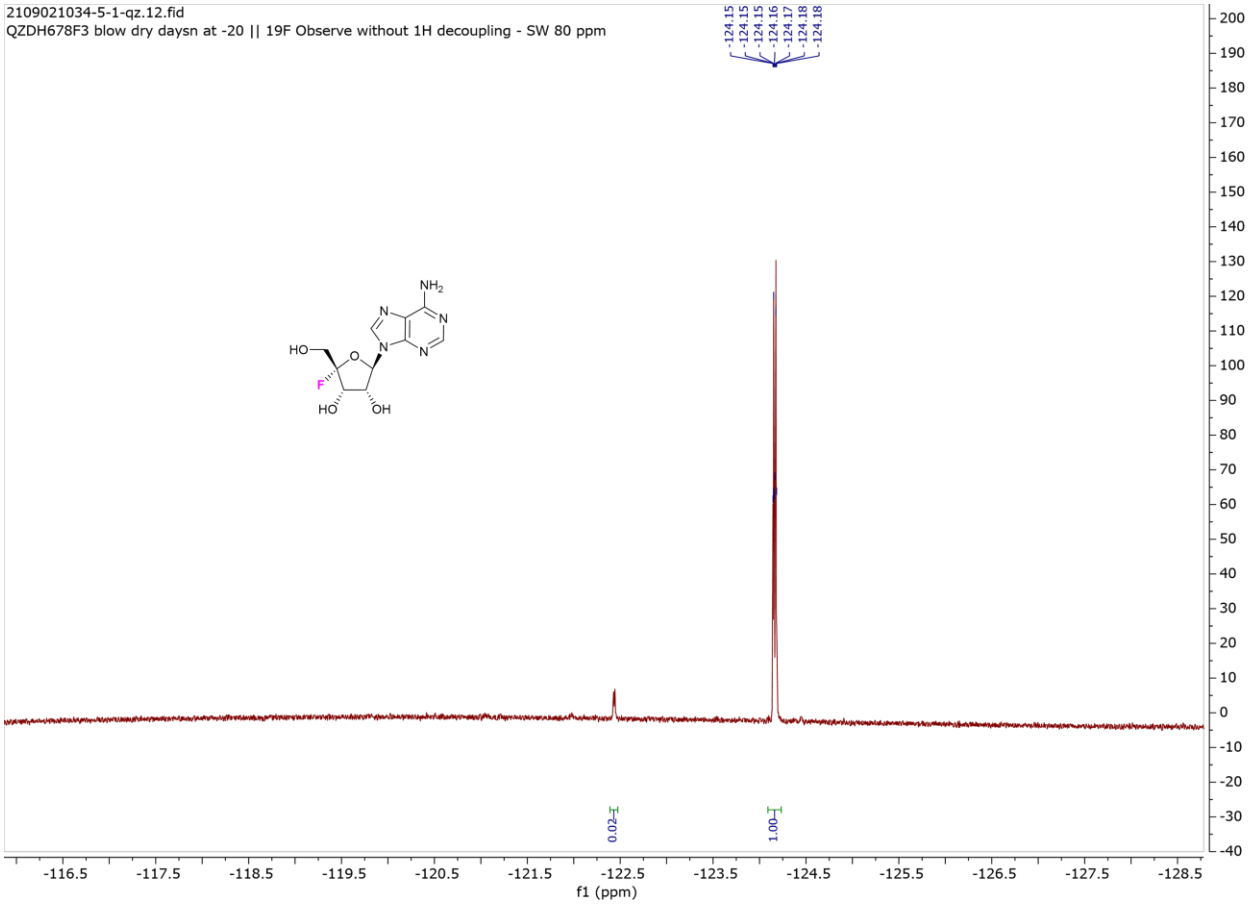


Figure S25. ^{19}F NMR of 4'-Fluoroadenosine **9**.

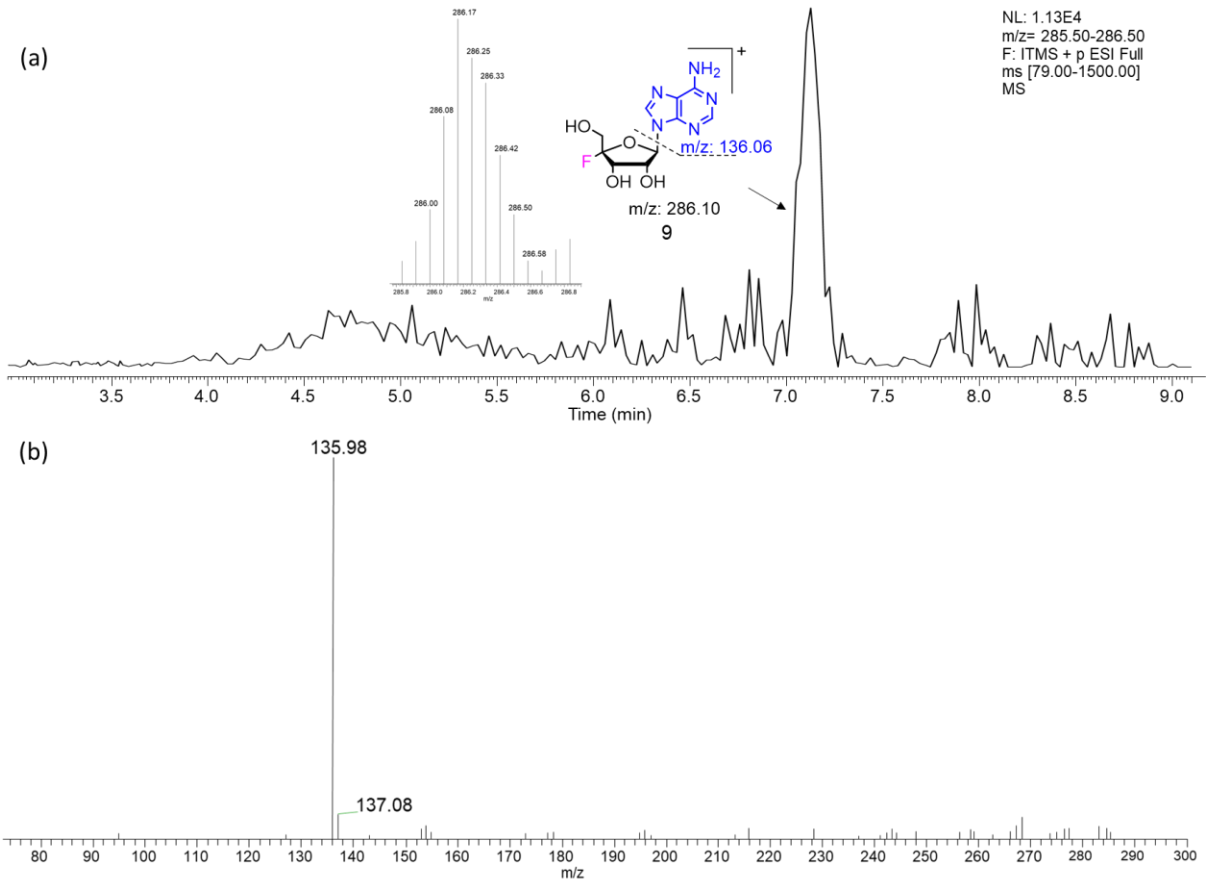


Figure S26. LC-MS and MS^2 of 4-fluoroadenosine in the extract of *S. aureorectus* DSM 41692.

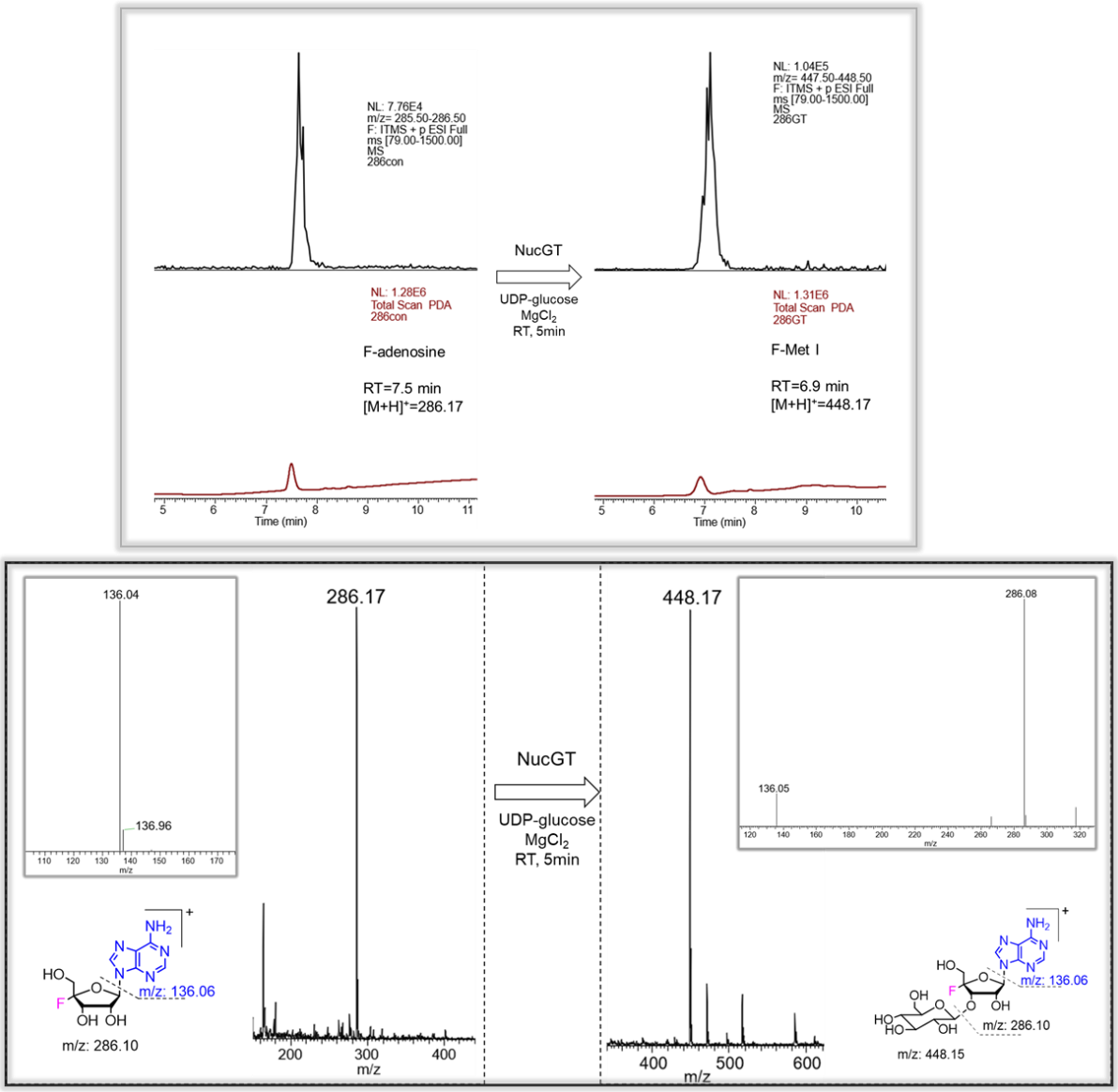


Figure S27. LC-MS and MS^2 of NucGT assay with synthetic 4-fluoroadenosine.

Reference

[1] X. Feng, D. Bello, P. T. Lowe, J. Clark, D. O'Hagan, *Chem. Sci.*, 2019, **10**, 9501.

[2] A. R. Maguire, W. Meng, S. M. Roberts, A. J. Willets, *J. Chem. Soc. Perkin Trans I*, 1993, **15**, 1795-1808.