Supporting Information

Deciphering the sugar biosynthetic pathway and tailoring steps of nucleoside antibiotic A201A unveils a GDP-L-galactose mutase

Qinghua Zhu[a, 1], Qi Chen[a, 1], Yongxiang Song[a], Hongbo Huang[a], Jun Li[a], Junying Ma[a], Qinglian Li[a], and Jianhua Ju[a, b, 2]

[a]Chinese Academy of Sciences (CAS) Key Laboratory of Tropical Marine Bioresources and Ecology, Guangdong Key Laboratory of Marine Materia Medica, Research Network for Applied Microbiology Center for Marine Microbiology, South China Sea Institute of Oceanology, CAS, Guangzhou 510301, China; and [b]College of Earth Sciences, University of Chinese Academy of Sciences, Beijing 10049, China

Author contributions: J.J. designed research; Q.Z., Q.C., Y.S., H.H. and J.L. performed research; J.M., Q.L., and J.J. analyzed data; and Q.Z., Q.C. and J.J. wrote the paper.

[1]Q.Z. and Q.C. contributed equally to this research.
[2]To whom correspondence should be addressed. E-mail: jju@scsio.ac.cn

Table of Contents

Materials and Methods

1. General Experimental Procedures.

Bacterial strains and plasmids used in this work are listed in Table S2. *E. coli* strains were grown on LB medium at 28 °C or 37 °C; the medium was supplemented with kanamycin (Kan), ampicillin, chloromycetin or apramycin (Apr) when necessary. *M. thermotolerans* SCSIO 00652 was cultured on the modified ISP4 medium (M-ISP4, ISP4 medium supplemented with 0.05 % yeast extract, 0.1 % tryptone and 3 % sea salt) plates at 28 °C. Tryptic Soy Broth (TSB) broth was adapted for the spore suspension of *M. thermotolerans* SCSIO 0652; *Streptomyces lividans* TK64 were propagated on SMA medium (soybean flour 2 %, mannitol 2 %, agar 2 %). For heterologous production of A201A, *S. lividans* TK64 mutants were grown in M-ISP4 medium.

Primers were synthesized at Sangon Biotech Company (Shanghai, China). EasyTaq™ DNA Polymerase, TransStart™ Fast-Pfu DNA Polymerase for polymerase chain reactions (PCR) were purchased from TransGen Biotech Company (Beijing, China). DNA Sequencing was accomplished at Beijing Genomics Institute (BGI) (Shenzhen, China). Restriction enzymes and DNA ligase were purchased from Takara Biotechnology Co. Ltd. (Dalian, China). Plasmid, gel extraction and cycle-pure kits were acquired from Omega Bioteck Inc. (GA, USA).

1-^{13}C-labeled D-mannose was purchased from Cambridge Isotope Laboratories Inc. (Hong Kong, China). L-galactose and GDP-α-D-mannose were purchased from Sigma-Aldrich (St. Louis, USA). GDP-β-L-galactose was purchased from Carbosynth Limited (Compton, UK). Other chemical solvents (analytical grade) were all purchased from standard commercial sources.

^{1}H, ^{13}C, and 2D (COSY, HMQC, HMBC and NOESY) NMR spectra were recorded at 25 °C with an Avance 500 MHz spectrometer instruments (Bruker). Low resolution and high resolution mass spectra were obtained with an Amazon SL ion trap instrument and a Maxis quadrupole-time-of-flight mass spectrometer (Bruker), respectively. Optical rotations were obtained with an MCP-500 polarimeter (Anton Paar).

Unless otherwise stated, the solvent system of analytical and semi-preparative HPLC consisted of solvent A (0.1 % AcOH and 15 % CH_3CN in ddH_2O) and solvent B (0.1 % AcOH and 85 % CH_3CN in ddH_2O). To analyze the metabolite profiles of *M. thermotolerans* SCSIO 00652, *S. lividans* TK64, and their mutants or engineered strains, the analytical HPLC program was carried out using a 210 solvent delivery module and a 335 photodiode array detector (Varian). Analytical HPLC was performed with a Phenomenex Prodigy ODS (150×4.60 mm, 5 μm) eluted with a linear gradient of 0 % to 70 % solvent B over 20 min, followed by 70 % to 100 % solvent B in 1 min, and then eluted with 100 % solvent B in 5 min, at a flow rate of 1.0 mL/min using UV detection at both 215 nm and 275 nm.

Semi-preparative HPLC was accomplished with a Hitachi Model D2000 Elite Chromatography Data Station (Hitachi, Japan) equipped with the Hitachi pump and diode array detector, using an YMC-Pack ODS column (YMC, 250×10 mm, 5 μm). Samples were eluted at 2.0 mL/min with a linear gradient from 40 % to 80 % solvent B for 20 min, 80 % to 100 % solvent B for 2 min, followed by holding at 100 % B for 5 min, and then eluted with 100 % solvent A for 3 min; UV detection was at both 215 nm and 275 nm.

2. Genomic Library Screening and Annotation of open reading frames.

We had already constructed the genomic library of *M. thermotolerans* SCSIO 00652 using the SuperCos1 vector (1, 2). The annotation of open reading frames (orfs) and predictions of their functions were accomplished using the orf finder program (http://www.ncbi.nlm.nih.gov/gorf/gorf.html), FramePlot 4.0 beta program (http://nocardia.nih.go.jp/fp4/), and Blast program (http://blast.ncbi.nlm.nih.gov/). A pair of primers mtdMF and mtdMR (Table S3) was designed and used for library screening. The positive clones were further validated with the other two pairs of primers (Table S3). In the end, 9 overlapping positive cosmids, designated 47A, 81A, 142H, 147B, 1912A, 191B, 184D, 49C and 209A, were screened.

3. Heterologous Expression of Cosmid 142H in *Streptomyces lividans* TK64.

Cosmid 142H harboring the whole sequence from *mtdC* to *orf3* was selected for heterologous expression. This cosmid was suffered from λ-RED-mediated recombination by replacing the kanamycin resistance gene within the SuperCos1 vector with a fragment excised from our modified pSET152AB vector containing the apramycin resistance gene and elements necessary for conjugation and site specific recombination (*oriT*, integrase gene and ϕC31 site) (3). The resulting cosmid, termed 142H-pSET152AB was transferred into *E. coli* ET12567/pUZ8002 and then introduced into *S. lividans* TK64 via conjugation to generate the *S. lividans* TK64/142H strain. The engineered strain *S. lividans* TK64/142H was fermented using the same medium as that of wild-type *M. thermotolerans* SCSIO 00652.

4. Gene Inactivations.

We have developed a genetic system for *M. thermotolerans* SCSIO 00652 using the λ-RED mediated PCR-targeting mutagenesis method (1, 2). Primers designed for inactivation of each gene are listed in Table S4. Cosmids 142H, 49C, and 1912A and were each transformed into *E. coli* BW25113/pIJ790 for gene inactivations. Gene disruption cassette *aac(3)IV-oriT* was amplified using a fragment from plasmid pIJ773 that was digested with *Eco*RI and *Hin*dIII. PCR products of *aac(3)IV-oriT* cassette for each disrupted gene were electro-transformed into *E. coli* BW25113/pIJ790 containing one of the above cosmids for λ-RED-mediated recombination to yield recombinant cosmids pJu3003~pJu3020. These recombinant cosmids were transformed into *E. coli* ET12567/pUZ8002, and suffered from conjugation with *M. thermotolerans* SCSIO 00652 wild strain. Double crossover mutants were first selected on the basis of the Kanamycin^{S}Apr^{R} phenotype and then further confirmed by PCR using primers listed in Table S5 (for gel analysis, see Fig. S1–S18). Finally, the 18 mutant strains of *M. thermotolerans* SCSIO

00652, Δ*mtdB*, Δ*mtdC*, Δ*mtdD*, Δ*mtdM*$_1$, Δ*mtdH,* Δ*mtdG*$_1$, Δ*mtdJ*, Δ*mtdK*, Δ*mtdL*, Δ*mtdM*, Δ*mtdG*$_2$, Δ*mtdM*$_2$, Δ*mtdM*$_3$, Δ*mtdW*, Δ*mtdM*$_4$, Δ*mtdWM*$_4$, Δ*orf1* and Δ*orf2* were generated (Table S6).

5. Construction of the Δ*mtdG*$_1$*WM*$_4$ and Δ*mtdM*$_2$*M*$_3$ Mutant Strains.

To construct the Δ*mtdG*$_1$*WM*$_4$ mutant strain, a three-step PCR-targeting technology via λ-Red-mediated recombination was used (4). The process is shown schematically in Fig.S19. The result of PCR verification of the mutant is shown in Fig. S20.

Firstly, the forward primer WdelAprF (5'- ctcattctcccgatcagggaaagcgcctccgatcgcaagACTAGT ATTCCGGGGATCCGTCGACC-3', underlined sequences represent *Spe*I site, small letters represent the 39 nt homologous to the region directly upstream of *mtdW*) and the reverse primer M4delAprR (5'-cgggatctcctcacgttccgacccgtcatcctggatcgcACTAGTTGTAGGCTGGAGCTGCTTC -3', underlined sequences represent *Spe*I site, small letters represent the 39 nt homologous to the region directly downstream of *mtdM*$_4$) were used to amplify the *acc(3)IV-oriT* resistance cassette from plasmid pIJ773 that had been digested with *Eco*RI-*Hin*dIII. The PCR product was used to replace the dual *mtdWM*$_4$ genes in the 142H cosmid, which had no *Spe*I recognition sites, to generate cosmid pJu3021. For excision of the resistance cassette, pJu3021 was digested with *Spe*I to delete the *acc(3)IV-oriT* fragment and self-ligated by T4 ligase overnight at 14 °C then ligation products were transformed into *E. coli* DH 5α and grown overnight on LB plates (with Amp and Kan) at 37 °C. The clones with Apr^{S}Kan^R phenotype were analyzed by restriction enzyme digestion and gel electrophoresis, then confirmed by amplifying a ~657 bp fragment using primers WtF2 (5'-GAGGATCTGGCTGACGGTAC-3') and M4tR (5'-GGAAGTTCTTGGTCCC GTCG-3'). The generated mutant cosmid was designated as pJu3022.

Secondly, another set of primers was designed, being pIJ773forw (5'-CAAGAGACAGGAT GAGGATCGTTTCGCATGATTCCGGGGATCCGTCGACC-3') and pIJ773rev (5'-CGGTCATTT

CGAACCCCAGAGTCCCGCTCATGTAGGCTGGAGCTGCTTC-3'; for each primer the sequences underlined are homologous to the kanamycin resistance sequence in SuperCos 1 vector). This pair of primers was used to amplify the *acc(3)IV-oriT* cassette. The amplified PCR product was transformed into *E. coli* BW25113/pIJ790 harboring pJu3022 for λ-RED mediated recombination to replace the kanamycin resistance sequence to generate pJu3023. Mutant cosmid pJu3023 was transformed into *E. coli* ET12567/pUZ8002 and conjugated with *M. thermotolerans* SCSIO 00652 wild-type as described. Single-crossover exconjugates showing apramycin resistance were first selected. The single-crossover mutants were cultured for further recombination with no antibiotic addition. The apramycin sensitive strains were then selected as a mixture of double-crossover mutants and strains back to wild-type, which could be further differentiated by PCR using primers WtF2 and M4tR and *Spe*I digestion of the PCR fragment. The mutant clones of Ju3023, in which *mtdWM*$_4$ dual genes were in-frame deleted were thus selected.

Thirdly, gene *mtdG*$_1$ was replaced by the *aac(3)IV-oriT* fragment from pJu3022 using primers G1dF and G1dR (Table S4) to generate cosmid pJu3024. Mutant cosmid pJu3024 was transformed into *E. coli* ET12567/pUZ8002 and conjugated with mutant Ju3023 spores as described. Single-crossover mutants with the Kan^{R}Apr^R phenotype were cultured on modified-ISP4 plates successively for four generations and double-crossover mutants Ju3024 were then selected from the Kan^{S}Apr^R phenotype. The mutant clones of Ju3024, in which *mtdWM*$_4$ dual genes were in-frame deleted and *mtdG*$_1$ was replaced by *aac(3)IV-oriT* fragment were further confirmed by PCR using primers G1tF and G1tR (Fig. S20).

The procedure used to construct Δ*mtdM*$_2$*M*$_3$ mutant is similar as above for the Δ*mtdG*$_1$*WM*$_4$ mutant. Finally, the *mtdM*$_2$ gene was in-frame deleted and *mtdM*$_3$ was replaced with the *aac(3)IV-oriT* resistance cassette to yield the Δ*mtdM*$_2$*M*$_3$ mutant (Ju3025), which was confirmed by PCR (Fig. S21).

6. Large Scale Fermentation, Isolation and Structural Elucidation of Metabolites 2–11 from *M. thermotolerans* Mutant Strains.

To isolate the A201A congeners from the mutants, a two-step fermentation process was adopted. A general procedure was described herein. First, a suitable portion of spore and mycelium (~1 cm^2) from solid M-ISP4 medium plate was used to inoculate 50 mL M-ISP4 medium in a 250-mL flask as a seed culture; the flask was cultured at 28 °C and 200 rpm for 36 h. Then, the seed culture (50 mL) was transferred to 200 mL M-ISP4 medium in a 1000 mL flask; the flask was cultured at 28 °C and 200 rpm for an additional 7–9 d. Multiple flasks were used. At last, the culture broth was centrifuged to yield a supernatant and a mycelium cake. The supernatant was extracted by equal volume of butanone three times and evaporated to dryness; the mycelium was extracted with 1.5 L acetone three times and evaporated to dryness; the two organic extracts were combined to yield a residue. The residue was dissolved in a 1:1 mixture of $CHCl_3$-MeOH and mixed with an appropriate amount of silica gel for normal phase silica gel column chromatography, eluted with a gradient elution of $CHCl_3$/MeOH mixture from 100/0, 98/2, 96/4, 94/6, 92/8, 90/10, 80/20 and 50/50 to yield eight fractions (Fr. A1-Fr. A8). The fractions were each analyzed by HPLC-UV. The fraction/s that contain/s the corresponding target compound was/were evaporated to dryness and subjected to another normal phase silica gel column chromatography, eluted with a gradient elution of EtOAc/MeOH mixture from 100/0, 95/5, 9/1, 85/15, and 8/2 to yield five fractions (Fr. B1- Fr. B5).

The fraction containing the targeted compound as judged by HPLC analysis was dissolved in MeOH, filtered with a 0.45 micron filtration membrane, and finally purified by semi-preparative HPLC to give compounds **2** (15 mg) from an 8-L scale fermentation of Δ*mtdM* mutant, **3** (2 mg) from a 24-L scale fermentation of Δ*mtdG$_1$WM$_4$* mutant, **4** (12 mg) from an 8-L scale fermentation of Δ*mtdJ* mutant, **5** (15 mg) from an 8-L scale fermentation of Δ*mtdH* mutant, **6** (2 mg) from an

8-L scale fermentation of Δ*mtdM*$_1$ mutant, **7** (11 mg) from an 8-L scale fermentation of Δ*mtdM*$_2$ mutant, **8** (12 mg) from an 8-L scale fermentation of Δ*mtdM*$_3$ mutant, **9** (8 mg) from an 8-L scale fermentation of Δ*mtdW* mutant, **10** (7 mg) from an 8-L scale fermentation of Δ*mtdWM*$_4$ mutant, and **11** (2 mg) from an 8-L scale fermentation of Δ*mtdM*$_2$*M*$_3$mutant, respectively.

Physicochemical Properties and Structure Elucidation of Compounds 2–11.

Compound **2** (from Δ*mtdK*, Δ*mtdL*, Δ*mtdM* or Δ*mtd*G$_2$ mutant) had a molecular formula of $C_{22}H_{26}N_6O_5$, as determined by HRESIMS. The ^{1}H and ^{13}C NMR spectroscopic data of **2** indicated that the signals for the rhamnose and hexfuranose units of the parent compound A201A (**1**) were missing in **2**. The MS, ^{1}H and ^{13}C NMR spectroscopic data of **2** are fully consistent with the hydrolysis product of **1** in the literature (5). Thus, the structure of **2** was established.

Compound 2 (from Δ*mtdK*, Δ*mtdL*, Δ*mtdM* or Δ*mtd*G$_2$ mutant): white solid; UV (MeOH) λ_{max} (log ε) 219 (4.20), 283 (4.34) nm; ^{1}H (500 MHz, CD_3OD) and ^{13}C NMR (125 MHz, CD_3OD), spectra see Table S8; (+)-HRESIMS *m/z* 455.2069 (calcd. for $C_{22}H_{27}N_6O_5$, 455.2037); NMR spectra, see Figs. S37 and S38.

Compound **3** (from Δ*mtdG*$_1$*WM*$_4$ mutant strain) had a molecular formula of $C_{28}H_{36}N_6O_{10}$ as provided by HRESIMS, corresponding to a hexose unit (162 Da) attached to **2**. Analysis of the ^{1}H, COSY, HSQC, and HMBC NMR data of **3** allowed full assignment of the NMR signals of **3**. The HMBC correlations from H-1' (5.60) to C-4 (156.2) of the *p*-hydroxy-α-methylcinnamic acid moiety, and C-4' (83.9) of the hexose moiety, and from H-4' (3.85) to C-1' (100.3) confirmed that the sugar unit is in furanose form attached to C-4 of the cinnamic acid moiety. The HPLC-UV-MS patterns of the two samples were fully in agreement. Hence, the hexanose unit in **3** was determined to be L-galatofuranose (see 7. Absolute Stereochemistry of the Galactose in Compound **3**).

Compound 3 (from Δ*mtdG*$_1$*WM*$_4$ mutant): white solid; ^{1}H and ^{13}C NMR data, see Table S8; (+)-HRESIMS *m/z* 617.2562 (calcd. for $C_{28}H_{37}N_6O_{10}$, 617.2566); NMR spectra, see Figs. S39–S43.

Compound **4** (from Δ*mtdG*$_1$ *or* Δ*mtdJ* mutant) had the molecular formula of $C_{29}H_{36}N_6O_{10}$, which was established by HRESIMS. The ^{1}H and ^{13}C NMR spectra for **4** showed less signals than those for A201A. Detailed comparison of the 1D NMR spectroscopic data of **4** with those of A201A revealed that these signals ascribed to the rhamnose moiety were absent in **4**, suggesting that the rhamnose was missing in **4**. The ^{1}H and ^{13}C NMR spectroscopic data of **4** were identical with those reported in literature as the minor product A201D (5). Thus, compound **4** was identified as A201D.

Compound 4 (from Δ*mtdG*$_1$ *or* Δ*mtdJ* mutant): white solid; $[\alpha]_D^{25}$ = -171 (*c* = 0.10, MeOH); UV (MeOH) λ_{max} (log ε) 218 (4.24), 281 (4.35) nm; ^{1}H and ^{13}C NMR data, see Table S9; (+)-HRESIMS *m/z* 629.2565 (calcd. for $C_{29}H_{37}N_6O_{10}$, 629.2560); NMR spectra, see Figs. S44–S45.

The molecular formula of compound **5** (from Δ*mtdH* mutant) was determined to be $C_{37}H_{50}N_6O_{15}$ on the basis of HRESIMS; **5** therefore was characterized by one more oxygen atom than found in A201A. The ^{1}H and ^{13}C NMR spectra for **5** were similar to those of A201A, except that the signals representative of the methyl in the rhamnose unit (5-Me) were missing, whereas additional signals for one oxygen-bearing methylene (δ_H 3.84, 3.72; δ_C 62.5 ppm) were clearly present. Furthermore, the ^{13}C NMR signal of C-4 in the rhamnose upfield shifted from δ_C 83 ppm in A201A to δ_C 77.2 ppm in **5**, and the ^{13}C NMR chemical resonance of C-5 in the rhamnose downfield shifted from δ_C = 69 to 73.9 ppm. These changes indicated that the 5-Me in rhamnose was replaced by a hydroxymethyl group. This structural elucidation was further confirmed on the basis of analysis of the ^{1}H-^{1}H COSY and HMBC spectra of **5** (Table S9).

Compound 5 (from Δ*mtdH* mutant): white solid; $[\alpha]_D^{25}$ = -107 (*c* = 0.22, MeOH); UV (MeOH) λ_{max} (log ***ε***) 217 (4.43), 281 (4.46) nm; ^{1}H and ^{13}C NMR data, see Table S9; (+)-HRESIMS *m/z* 819.3435 (calcd. for $C_{37}H_{51}N_6O_{15}$, 819.3407); NMR spectra, see Figs. S46–S50.

Compound **6** (from Δ*mtdM*$_1$ mutant) was isolated as a white solid. The molecular formula of **6** was established to be $C_{35}H_{46}N_6O_{14}$ by HRESIMS, 28 mass units smaller than that of A201A. Careful examination of the ^{1}H and ^{13}C NMR spectra with those of A201A revealed good similarity, except that the signals for the two methyl groups (δ_H 3.37, δ_C 37.7 ppm) of the adenine moiety were missing, indicating the two methyl groups were substituted by hydrogen atoms. The subsequent HMQC and HMBC experiments supported this structural elucidation and allowed the complete assignment of the ^{1}H and ^{13}C NMR spectroscopic data (Table S9).

Compound 6 (from Δ*mtdM*$_1$ mutant): white solid; $[\alpha]_D^{25}$ = -106 (*c* = 0.66, MeOH); UV (MeOH) λ_{max} (log ***ε***) 215 (4.49), 270 (4.39) nm; ^{1}H and ^{13}C NMR data, see Table S9; (+)-HRESIMS *m/z* 775.3152 (calc. for $C_{35}H_{47}N_6O_{14}$, 775.3145); NMR spectra, see Figs. S51–S55.

Compounds **7** (from Δ*mtdM*$_2$ mutant) and **8** (from Δ*mtdM*$_3$ mutant) shared the same molecular formula $C_{36}H_{48}N_6O_{14}$, which was one CH_2 unit smaller than that of A201A, according to HRESIMS. The ^{1}H and ^{13}C NMR spectra for **7** and **8** resembled those of A201A, except that one methoxy group signals were missing in **7** and **8**, respectively. Additionally, the ^{1}H and ^{13}C NMR signals for the rhamnose unit in **7** and **8** were different with those in A201A. The ^{1}H-^{1}H COSY correlations of H-1/H-2/H-3/H-4/H-5/H_3-6 showed the presence of the rhamnose skeleton in both **7** and **8** with the aid of an HMQC experiment. The structural difference of **7** and **8** was only in the location of the methoxy group on the rhamnose. The HMBC correlation from the methoxy protons (δ_H 3.43 ppm, 3-OMe) to C-3 (δ_C 82.1 ppm) linked the methoxy at C-3 of the rhamnose in **7**. However, in **8**, the HMBC correlation from the methoxy protons (δ_H 3.56 ppm, 4-OMe) to C-4 (δ_C 84.5 ppm) located the methoxy at C-4 on the rhamnose. Thus, the structures of **7** and **8**

were identified. The ^{1}H and ^{13}C NMR spectroscopic data of **7** and **8** were assigned by ^{1}H-^{1}H COSY, HMQC and HMBC experiments, respectively (Tables S9 and S10).

***Compound* 7** (from Δ*mtdM*$_2$ mutant): white solid; $[\alpha]_D^{25}$ = -102 (*c* = 0.20, MeOH); UV (MeOH) λ_{max} (log ε) 217 (4.42), 281 (4.45) nm; ^{1}H and ^{13}C NMR data, see Table S9; (+)-HRESIMS *m/z* 789.3296 (calcd. for $C_{36}H_{49}N_6O_{14}$, 789.3315); NMR spectra, see Figs. S56–S60.

***Compound* 8** (from Δ*mtdM*$_3$ mutant): white solid; $[\alpha]_D^{25}$ = -212 (*c* = 0.13, MeOH); UV (MeOH) λ_{max} (log ε) 217 (4.84), 279 (4.80) nm; ^{1}H and ^{13}C NMR data, see Table S10; (+)-HRESIMS *m/z* 789.3312 (calcd. for $C_{36}H_{49}N_6O_{14}$, 789.3315); NMR spectra, see Figs. S61–S65.

The molecular formula of **9** (from Δ*mtdW* mutant) was determined to be $C_{37}H_{52}N_6O_{14}$ on the basis of HRESIMS, which has two mass units greater than that of A201A. Comparisons of the ^{1}H and ^{13}C NMR spectroscopic data with those of A201A revealed that the double bond (δ_C 143.1, 132.0) signals of the hexofuranose moiety were absent, whereas two sets of oxygen-bearing methine signals were observed at δ_H 3.90, δ_C 83.5 and δ_H 3.37, δ_C 83.0 ppm, indicating that the olefinic bond was saturated in **9**. In the ^{1}H-^{1}H COSY spectrum, the correlation of the protons at δ_H 3.90 and 3.37 ppm suggested the presence of X-(O)CH-CH(O)-X′ fragment. HMBC correlations from H-4 to C-1, C-3, and from H-5 to C-4, 5-OMe confirmed the saturated hexofuranose unit. Thus, compound **9** has the same structure as the reported A201E, which was previously isolated as a minor product but the stereochemistry of the hexofuranose has not heen solved (5). The hexofuranose unit in compound **9** was deduced to be α-L-galactofuranose as in **3** on the basis of ^{1}H and ^{13}C NMR data comparisons with those of **10** and **3**; considerations of common biosynthetic machineries also support this conclusion.

Compound 9 (from Δ*mtdW* mutant): white solid; $[\alpha]_D^{25}$ = -61 (*c* = 0.21, MeOH); UV (MeOH) λ_{max} (log ε) 216 (4.26), 281 (4.36) nm; ^{1}H and ^{13}C NMR data, see Table S10; (+)-HRESIMS *m/z* 805.3625 (calcd. for $C_{37}H_{53}N_6O_{14}$, 805.3614); NMR spectra, see Figs. S66–S70.

Compound **10** **(**from Δ*mtdWM4* mutant) possessed the molecular formula of $C_{36}H_{50}N_6O_{14}$, which was one CH_2 unit smaller than that of **9**, as determined by HRESIMS. The ^1H and ^{13}C NMR spectra of **10** were very similar with those of **9**, except for the absence of the methoxy group signals at δ_H 3.33, δ_C 59.7 ppm in **10**. Furthermore, the ^{13}C resonance for C-5 of the hexofuranose unit shifted upfield from δ_C 83.0 ppm in **10** to δ_C 72.0 ppm in **9**. These observed changes indicated that the methoxy group at C-5 of the hexofuranose unit was substituted by a hydroxy group. The hexofuranose unit in compound **10** was deduced to be α-L-galactofuranose as in **3**. The ^1H and ^{13}C NMR spectroscopic data were assigned on the basis of the analysis of HMQC and HMBC spectra (Table S10).

Compound 10 **(**from Δ*mtdWM4* mutant): white solid; $[\alpha]_D^{25}$ = -90 (*c* = 0.11, MeOH); UV (MeOH) λ_{max} (log *ε*) 218 (4.38), 282 (4.50) nm; ^1H and ^{13}C NMR data, see Table S10; (+)-HRESIMS *m/z* 791.3485 (calcd. for $C_{36}H_{51}N_6O_{14}$, 791.3458); NMR spectra, see Figs. S71–S75.

Compound **11** was determined to have the molecular formula $C_{35}H_{46}N_6O_{14}$ on the basis of HRESIMS, two CH_2 units less than that of **1**. A full set of 1D (^1H and ^{13}C) and 2D (COSY, HSQC, and HMBC) NMR data for **11** were acquired, thus allowing for full assignment of the ^1H and ^{13}C signals. In comparison with those of **1**, the two methyl groups within the rhamnose moiety were missing in **11**. Analysis of the HMBC correlations of **11** confirmed this conclusion.

Compound 11: white solid; $[\alpha]_D^{25}$ = -69 (*c* = 0.10, MeOH); ^1H and ^{13}C NMR data, see Table S10; (+)-HRESIMS *m/z* 775.3114 (calcd. for $C_{35}H_{47}N_6O_{14}$ $[M + H]^+$, 775.3145); NMR spectra, see Figs. S76–S80.

7. Absolute Stereochemistry of the Galactose in Compound 3.

Compound **3** (1.0 mg) was dissolved in 1.0 mL of 1 N HCl in 90 % MeOH, and the solution was kept at 80 °C for 12 h in a sealed tube. The reaction mixture was dried under vacuum, and partitioned between $CHCl_3$ and H_2O/MeOH (8:2). The aqueous layer was concentrated to

dryness in vacuo, and then benzoylated with 150 μL of benzoyl chloride in 1 mL of pyridine at 25 °C for 12 h. The reaction was quenched with 1 mL of MeOH. The reaction mixture was dried under vacuum. The residue was subjected to SiO_2 CC and eluted with $CHCl_3$ to afford fraction A. Fraction A was then purified by HPLC with a phenomenex ODS luna column (150×4.60 mm, 5 μm) using an elution system consisting of solvent A (0.1 % HOAc/15 % CH_3CN in H_2O) and solvent B (0.1 % HOAc/85 % CH_3CN in H_2O). Elution was done at 1 mL/min with a linear gradient from 100 % to 15 % A over the course of 20 min, and from 15 % A to 100 % B over the course of 5 min, then hold for 5 min. The peak at t_R 24 min was gathered and dried by air to give compound **12**. This compound was identified as methyl tetra-*O*-benzoyl-α-L-galactopyranoside on the basis of comparisons of its MS, ^{1}H-NMR and CD spectrum data with those of **13** (6, 7).

Methyl Tetra-*O*-benzoyl-α-L-galactopyranoside (13).

3.0 mg of L-galactose was dissolved in 1.5 mL of 1 N HCl in 90 % MeOH, and the solution was kept at 80 °C for 12 h in a sealed tube. The reaction mixture was benzoylated as noted above. The residue was subjected to HPLC purification using the same conditions as described above to yield compound **13**. HR-ESI-MS (*m/z*) 611.1909 ([M + H]$^+$, calcd for 611.1912, $C_{35}H_{31}O_{10}$). ^{1}H NMR ($CDCl_3$): δ = 8.10 (2H, d, 8.0, benzoyl proton), 8.03 (2H, d, 8.0 Hz, benzoyl proton), 7.99 (2H, d, 8.0 Hz, benzoyl proton), 7.79 (2H, d, 8.0 Hz, benzoyl proton), 7.62-7.24 (12H, overlapping signals, benzoyl proton), 6.04 (1H, d, 3.3 Hz, H-4), 6.01 (1H, dd , 10.8 and 3.3 Hz, H-3), 5.69 (1H, dd , 10.8 and 3.5 Hz, H-2), 5.33 (1H, d, 3.5 Hz, H-1), 4.62 (2H, overlapping signals, H-5 and H-6a), 4.41 (1H, dd, 8.8 and 8.8 Hz, H-6b), 3.49 (3H, s, OMe) ppm. CD spectrum see Fig. S27.

Methyl Tetra-*O*-benzoyl-α-D-galactopyranoside (14).

5.0 mg of D-galactose was subjected to acidic methanolysis followed by benzoylation as described above. HPLC purification using the same conditions as above described to attain pure **14**. HR-ESI-MS (*m/z*) 611.1902 ([M + H]$^+$, calcd for 611.1912, $C_{35}H_{31}O_{10}$). ^{1}H NMR ($CDCl_3$):

δ = 8.10 (2H, d, 8.0, benzoyl proton), 8.03 (2H, d, 8.0 Hz, benzoyl proton), 7.99 (2H, d, 8.0 Hz, benzoyl proton), 7.80 (2H, d, 8.0 Hz, benzoyl proton), 7.63-7.25 (12H, overlapping signals, benzoyl proton), 6.03 (1H, d, 3.3 Hz, H-4), 6.01 (1H, dd , 10.7 and 3.3 Hz, H-3), 5.70 (1H, dd , 10.7 and 3.5 Hz, H-2), 5.33 (1H, d, 3.5 Hz, H-1), 4.62 (2H, overlapping signals, H-5 and H-6a), 4.43 (1H, dd, 8.8 and 8.8 Hz, H-6b), 3.49 (3H, s, OMe) ppm. CD spectrum see Fig. S27.

8. Feeding Experiment with [1-^{13}C] D-mannose into *M. thermotolerans* SCSIO 00652 and Purification of ^{13}C Enriched A201A.

A two-step fermentation process was adopted for the production of A201A. A volume of 100 μL spores of *M. thermotolerans* SCSIO 00652 was inoculated into a 250 mL Erlenmeyer flask containing 50 mL M-ISP4 medium (three flasks were used); the flasks were incubated at 28 °C, 200 rpm on the rotary shaker for 36 h. Then, each of the 50 mL seed cultures was transferred into each of the three 1000 mL Erlenmeyer flask each containing 200 mL M-ISP4 medium; the three flasks were incubated at 28 °C, 200 rpm on the rotary shaker for another 24 h. Subsequently, 200 mg of [1-^{13}C] D-mannose were added to each of the three flasks. After an additional 7 d of cultivation at 28 °C, 200 rpm, the fermentation broth (0.75 L) was centrifuged to yield a supernatant and a mycelium cake. The supernatant was extracted with 0.75 L butanone three times; the mycelium was extracted with 0.5 L acetone three times. Both organic extracts were evaporated to dryness and combined to yield a residue. The residue was dissolved in a 1:1 mixture of $CHCl_3$-MeOH and mixed with an appropriate amount of silica gel for normal phase silica gel column chromatography, eluted with a gradient elution of $CHCl_3$/MeOH mixture from 100/0, 98/2, 96/4, 94/6, 92/8 and 90/10 to yield six fractions (Fr.1-Fr.6). Fr.6 was dissolved in MeOH and filtered with a 0.45 micron filtration membrane and purified by semi-preparative HPLC to give ^{13}C-enriched A201A (**1**) (~0.6 mg) and then subjected to ^{13}C NMR data acquisition (Figs. S84 and S85).

9. Overexpression and Purification of MtdM and MtdL.

Overexpression of *mtdM* and *mtdL* in *E.coli* BL21 (DE3) is described below. The *mtdM* was PCR-amplified from cosmid 142H with primer pairs of MtdMexpF (CATATGTCAACAGGAA TCAAACGCGC, the underlined sequences represent *Nde*I site) and MtdMexpR (ACTAGTAAGCTTTCACGGCCGGACATGCGGCT, the underlined sequences represent *Hin*dIII site); the *mtdL* was PCR-amplified from cosmid 142H with primer pairs of MtdLexpF (CATATGTCCGGCCGTGACATCTCGAC, the underlined sequences represent *Nde*I site) and MtdLexpR (GGATCCTCACGTTCCCGCCACCGTG, the underlined sequences represent *Bam*HI site). The PCR products of *mtdM* and *mtdL* were each recovered from agarose gel with a gel extraction kit, modified by way of extension using EasyTaq DNA polymerase at 72 °C for 20 min, recycled by cycle-pure kit, ligated into pCR 2.1 TA cloning vector, and finally confirmed by DNA sequencing. The corresponding fragments were then excised from pCR 2.1/*mtdM* with *Nde*I/*Bam*HI and pCR 2.1/*mtdL* with *Nde*I/*Hin*dIII and cloned into the same site of pET/28a(+) vector to yield plasmids pET28a(+)/*mtdM* and pET28a(+)/*mtdL*, which were transformed into *E. coli* BL21(DE3) to yield strains *E.coli* BL21(DE3)/pET28a(+)/*mtdM* and *E.coli* BL21(DE3)/ pET28a(+)/*mtdL* for protein expression, respectively.

Each of the above two strains was cultured at 28 °C and 200 rpm to OD_{600}=0.6. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was then added to a final concentration of 0.05 mM to induce the expression of *mtdM* and *mtdL*. After cultivation at 25°C for an additional 12 h, the cells were collected by centrifuge, washed with 50 mM Tris·HCl buffer (pH 8.0) twice, resuspended in the binding buffer (50 mM phosphate buffer, 500 mM NaCl, and 5 mM imidazole, pH 8.0), sonicated (0 °C) and centrifuged. The supernatant was loaded onto 1 mL Ni affinity column packed by Ni-NTA His·Bind Resin, washed by 3 mL washing buffer I (50 mM phosphate buffer, pH 8.0, 500 mM NaCl, 30 mM imidazole, 10% glycerol) and 3 mL washing buffer II (50 mM Tris·HCl buffer, pH 8.0, 500 mM NaCl, 50 mM imidazole, 10% glycerol), eluted by 2.5 mL elution buffer I

(50 mM Tris·HCl buffer, pH 8.0, 500 mM NaCl, 250 mM imidazole, 10% glycerol) and 1 mL elution buffer II (50 mM Tris·HCl buffer, pH 8.0, 500 mM NaCl, 1 M imidazole, 10% glycerol). The fractions eluted by elution buffer I containing MtdM or MtdL were then desalted by PD-10 desalting column, concentrated by filtration on a 10K Amicon Ultra-15 centrifugal filters, finally dissolved in a storage buffer (25 % glycerol, 50 mM phosphate buffer, pH 8.0) and stored at -80 °C for further experiments. All of the protein purification steps were conducted at 4 °C. The concentration of purified MtdM and MtdL was determined using Bio-Rad protein assay dye reagent according to the protocol.

10. Site-directed Mutagenesis and Purification of MtdL Mutant Proteins.

Site-directed mutagenesis of MtdL was conducted following the Fast Mutagenesis System (TransGen, Beijing, China) manual protocol. The expression and purification of mutant proteins was carried out using the same procedures described above. All primers used here are listed in Table S7.

11. In vitro Biochemical Activities of MtdM and MtdL.

To validate the in vitro biochemical activities of MtdM and MtdL, we overexpressed and purified the two enzymes as soluble N-terminus–His_6-tagged proteins from *E. coli* (Fig. S89). MtdM enzymatic activity was tested in 50 μL volume containing 0.2 mM GDP-α-D-mannose (or GDP-β-L-galactose), 0.5 mM NAD^+, 2 μM MtdM, in 50 mM Tris·HCl (pH 8.0), at 37 °C for 40 min. The reaction was quenched with 5 μL formic acid. After centrifugation to remove the protein, the supernatant was analyzed with analytical HPLC using a 210 solvent delivery module, a 335 photodiode array detector (Varian) and a Waters Symmetry (250×4.60 mm, 5μm) column. Samples were eluted with an isocratic elution of 50 mM triethylammonium acetate (TEAA) containing 1.5 % CH_3CN in 35 min, at a flow rate of 0.7 mL/min using UV detection at 254 nm.

The reaction mixture was also subjected to LC-MS analysis to confirm the correct molecular weight of each peak (Fig. S88).

Enzymatic activities of MtdL and its site-directed mutants were tested in 50 μL volume containing 0.2 mM GDP-β-L-galactose, 5 mM Mg^{2+} or Mn^{2+}, 2 μM MtdL, in 50 mM PBS (pH 8.0), at 37 °C for 20 min. After quenched with 5 μL formic acid, the supernatant was for HPLC analysis. The detection for MtdL-catalyzed reactions was conducted using a Waters Symmetry column (250×4.60 mm, 5 μm) connected with another Kromasil 100-5C18 (250×4.60 mm, 5 μm) column. The rest of the procedure was carried out as noted above.

12. Large-scale MtdL Enzymatic Reaction and Characterization of the Product *in situ*.

All the solvents were prepared with 50 mM PBS (dissolved with D_2O). The Tris·HCl buffer of MtdL protein was exchanged with 50 mM PBS (pH 8.0) three times using Amicon® Ultra-0.5 Centrifugal Filter Devices. 2 mg of GDP-β-L-galactose was dissolved in 300 μL 50 mM PBS (pH 8.0) for ^1H NMR and H-H COSY analysis. The GDP-β-L-galactose solution was used for the large-sized reaction. A total of 500 μL reaction containing 5 mM Mg^{2+} and excess MtdL was tested. After incubation at 37 °C for 20 min, the reaction system was filtered through an Amicon® Ultra-0.5 Centrifugal Filter Device to remove the enzyme. The supernatant containing both GDP-β-L-galactose and MtdL enzymatic products was used to carry out ^1H NMR and ^1H-^1H COSY analyses.

13. Antibacterial Activities.

The antibacterial activities of compounds **1**, **2**, and **4**–**11** were assessed using 2-fold serial dilutions of antibacterial agents in MH broth, according to previously reported standard methods provided by Clinical and Laboratory Standards Institute (CLSI) (8). These compounds were

tested for their antibacterial activities against *Staphyloccocus aureus* ATCC 29213, *Escherichia coli* ATCC 25922, *Acinetobacter baumannii* ATCC 19606, *Aeromonas hydrophila* ATCC 7966, *Micrococcus luteus*, methicillin-resistant *Staphylococcus aureus* (clinical isolate shhs-A1 from Shanghai Huashan hospital), and methicillin-resistant *Staphylococcus epidermidis* (clinical isolate shhs-E1 from Shanghai Huashan hospital) using a broth dilution method. Each of the reported MIC values in Table S11 is the lowest concentration of antimicrobial agent that completely inhibits growth of the organism in microdilution wells as detected by the unaided eyes.

14. The Water Solubility of Compound 1 and 6.

The water solubility of compounds A201A (**1**) and analog des-*N*, *N*-dimthyl A201A (**6**) were calculated using the Water Solubility Module in Percepta software of Advanced Chemistry Development, Inc. At pH 7.32, the water solubility unit of **1** is –2.95 (log(mol/L)) while **6** is –2.90 (log(mol/L)).

Supplementary Tables S1–S11

Table S1. The gene cluster for biosynthesis of A201A in *M. thermotolerans* SCSIO 00652 and proposed functions.

Protein	size[a]	protein ID and proposed function	*ata* homolog (protein ID), identity/positive	*pur*/*hyg* homolog (protein ID), identity/positive
MtdA	244	AET51839.1, GntR-family transcriptional regulator	—	—
MtdB	675	AET51840.1, eptidase S15	—	—
MtdC	428	AET51841.1, putative uncharacterized protein	—	—
MtdD	111	AET51842.1, putative uncharacterized protein	—	—
MtdE	148	AET51843.1, NTP-pyrophosphohydrolase	AtaP7 (CAD27645.1), 70/81	Pur7 (CAA63159.1), 61/71
MtdF	384	AET51844.1, oxireductase	AtaP10 (CAD27646.1), 67/78	Pur10 (CAA63160.1), 51/68
MtdG	429	AET51845.1, aminotransferase	AtaP4 (CAD27647.1), 80/87	Pur4 (CAA63162.1), 73/80
$MtdM_1$	234	AET51846.1, *N*-methyltransferase	AtaP5 (CAD27648.1), 75/82	Pur5 (CAA63163.1), 66/76
MtdH	354	AET51847.1, GDP-D-manose-4, 6-dehydratase	Ata12 (CAD27644.1), 76/84	—
MtdI	404	AET51848.1, acyltransferase	AtaPKS1 (CAD27643.1), 60/69	Hyg22 (ABC42559.1), 50/62
$MtdG_1$	375	AET51849.1, glycosyltransferase	Ata13 (CAD62186.1), 77/83	—
MtdJ	311	AET51850.1, oxidoreductase	Ata14 (CAD62187.1), 49/56	—
$MtdR_1$	570	AET51851.1, ABC transporter	Ard1 (CAA59109.1) , 53/63	—
MtdK	452	AET51852.1, glucokinase	Ata15 (CAD62188.1), 53/.63	—
MtdL	376	AET51853.1, transglycosylase	Ata16 (CAD62189.1), 72/82	Hyg20 (ABC42557.1), 61/74
MtdM	342	AET51872.1, NAD-dependent epimerase/dehydratase	Ata17 (CAD62190.1), 74/81	—
MtdN	115	AET51854.1, acyl carrier protein	AtaPKS2 (CAD62191.1), 57/71	Hyg9 (ABC42546.1), 30/41
MtdO	382	AET51855.1, ketoacyl synthase 1	AtaPKS3 (CAD62192.1), 64/73	Hyg10 (ABC42547.1), 53/64
MtdP	268	AET51856.1, uncharacterized protein	AtaPKS4 (CAD62193.1), 55/66	Hyg11 (ABC42548.1), 51/65
MtdQ	484	AET51857.1, CoA-ligase	Ata18 (CAD62194.1), 61/71	Hyg12 (ABC42549.1), 47/60
MtdS	94	AET51858.1, ACP	Ata19 (CAD62195.1), 63/81	Hyg13(ABC42550.1), 43/63
MtdT	339	AET51859.1, 3-ketoacyl ACP dehydratase	Ata2 (CAD62196.1), 72/80	Hyg14 (ABC42551.1), 43/58
$MtdR_2$	194	AET51860.1, phosphotransferase	Ard2 (CAD62197.1), 75/83	Hyg21 (ABC42558.1), 59/74
MtdU	249	AET51861.1, 3-ketoacyl ACP reductase	Ata4 (CAD62198.1), 73/82	Hyg15 (ABC42552.1), 55/69
$MtdG_2$	426	AET51862.1, glycosyltransferase	Ata5 (CAD62199.1), 75/81	Hyg16 (ABC42553.1), 63/72
$MtdM_2$	354	AET51863.1, methyltransferase	Ata6 (CAD62200.1), 71/78	—
MtdV	186	AET51864.1, putative chorismate pyruvate-lyase	Ata7 (CAD62201.1), 57/65	Hyg4 (ABC42541.1), 18/31
$MtdM_3$	242	AET51865.1, methyltransferase	Ata8 (CAD62202.1), 69/80	Hyg6 (ABC42543.1), 28/40
$MtdR_3$	429	AET51866.1, transmembrane protein	Ata9 (CAD62203.1), 67/78	Hyg19 (ABC42556.1), 50/66
MtdW	529	AET51867.1, oxidoreductase	Ata10 (CAD62204.1), 65/73	—
$MtdM_4$	240	AET51868.1, methyltransferase	Ata11 (CAD62205.1), 70/78	Hyg6 (ABC42543.1), 31/40
Orf1	260	AET51869.1, inositol-phosphate phosphatase	AtaP3 (CAD27649.1), 61/73	Pur3 (CAA63164.1), 61/72
Orf2	101	AET51870.1, putative gas vesicle synthesis protein	—	—

[a]Size in units of amino acids (aa); *ata*: biosynthetic gene cluster of A201A in *Saccharothrix mutabilis* subsp. *capreolus*; *pur:* biosynthetic gene cluster of puromycin in *Streptomyces alboniger; hyg:* biosynthetic gene cluster of hygromycin A in *Streptomyces hygroscopicus* NRRL 2388; ABC: ATP-binding cassete; CoA: coenzyme A; ACP: acyl carrier protein.

Table S2. Bacteria and plasmids used in this study.

Strains	Description	Reference or source
E. coli		
DH5α	Host strain for general clone	Stratagene
ET12567	*dam, dcm, hsdS, cat, tet*	(4)
BW25113	K12 derivative: *araBAD, rhaBAD*	(4)
BL21(DE3)	F-*ompT hsdS gal dcm* (DE3)	Novagen
Plasmids		
pCR2.1	Amp^R, Kan^R, general clone vector	Invitrogen
pET28a(+)	Kan^R, expression vector	Novagen
pIJ790	Cml^R, including λ-RED (*gam, bet, exo*) for PCR-targeting	(4)
pIJ773	*aac(3)IV* (Apr^R), *oriT*	(4)
pUZ8002	*tra, neo, RP4*	(4)
142H	*M. thermotolerans* SCSIO 00652 genomic library cosmid	This study
49C	*M. thermotolerans* SCSIO 00652 genomic library cosmid	This study
1912A	*M. thermotolerans* SCSIO 00652 genomic library cosmid	This study
pJu3003	49C cosmid derivative where *mtdB* was disrupted by *aac(3)IV-oriT* fragment using primers BdF and BdR	This study
pJu3004	49C cosmid derivative where *mtdC* was disrupted by *aac(3)IV-oriT* fragment using primers CdF and CdR	This study
pJu3005	49C cosmid derivative where *mtdD* was disrupted by *aac(3)IV-oriT* fragment using primers DdF and DdR	This study
pJu3006	49C cosmid derivative where $mtdM_1$ was disrupted by *aac(3)IV-oriT* fragmentusing primers M1dF and M1dR	This study
pJu3007	49C cosmid derivative where *mtdH* was disrupted by *aac(3)IV-oriT* fragment using primers HdF and HdR	This study
pJu3008	49C cosmid derivative where $mtdG_1$ was disrupted by *aac(3)IV-oriT* fragment using primers G1dF and G1dR	This study
pJu3009	142H cosmid derivative where *mtdJ* was disrupted by *aac(3)IV-oriT* fragment using primers JdF and JdR	This study
pJu3010	142H cosmid derivative where *mtdK* was disrupted by *aac(3)IV-oriT* fragment using primers KdF and KdR	This study
pJu3011	142H cosmid derivative where *mtdL* was disrupted by *aac(3)IV-oriT* fragment using primers LdF and LdR	This study
pJu3012	142H cosmid derivative where *mtdM* was disrupted by *aac(3)IV-oriT* fragment using primers MdF and MdR	This study
pJu3013	1912A cosmid derivative where $mtdG_2$ was disrupted by *aac(3)IV-oriT* fragment using primers G2dF and G2dR	This study
pJu3014	1912A cosmid derivative where $mtdM_2$ was disrupted by *aac(3)IV-oriT* fragment using primers M2dF and M2dR	This study
pJu3015	1912A cosmid derivative where $mtdM_3$ was disrupted by *aac(3)IV-oriT* fragment using primers M3dF and M3dR	This study
pJu3016	1912A cosmid derivative where *mtdW* was disrupted by *aac(3)IV-oriT* fragment using primers WdF and WdR	This study

pJu3017	1912A cosmid derivative where *mtdM*$_4$ was disrupted by *aac(3)IV-oriT* fragmentusing primers M4dF and M4dR	This study
pJu3018	1912A cosmid derivative where *mtdWM*$_4$ was disrupted by *aac(3)IV-oriT* fragment using primers WdF and M4dR	This study
pJu3019	1912A cosmid derivative where *orf1* was disrupted by *aac(3)IV-oriT* fragment using primers orf1dF and orf1dR	This study
pJu3020	1912A cosmid derivative where *orf2* was disrupted by *aac(3)IV-oriT* fragment using primers orf2dF and orf2dR	This study
pJu3021	142H cosmid derivative where *mtdWM*$_4$ was disrupted by *aac(3)IV-oriT* fragment using primers WdelAprF and M4delAprR	This study
pJu3022	pJu3021 cosmid derivative where *mtdWM*$_4$ was in-frame deleted	This study
pJu3023	pJu3022 cosmid derivative where *mtdWM*$_4$ was in-frame deleted and kanamycin resistant gene was replaced by *aac(3)IV-oriT* fragment using primers pIJ773forw and pIJ773rev	This study
pJu3024	pJu3022 cosmid derivative where *mtdWM*$_4$ was in-frame deleted and *mtdG*$_1$ was disrupted by *aac(3)IV-oriT* fragment using primers G1dF and G1dR	This study
pJu3025	142H cosmid derivative where *mtdM*$_2$ was in-frame deleted using primers mtdspeIAprM2F and mtdspeIAprM2R, and *mtdM*$_3$ was disrupted by *aac(3)IV-oriT* fragment using primers M3dF and M3dR	This study
S. lividans TK64	*SLP2*$^-$ *SLP3*$^-$ *pro-2 str-6*	(9)

Table S3. Primer pairs used in screening for the biosynthetic gene cluster of A201A in the genomic library of *Marinactinospora thermotolerans* SCSIO 00652.

Locus	Primer code	Sequences (5′-3′)	Length of deduced product
mtdG	mtdGF	CTGTTCGCCGCCTATTTCG	514bp
	mtdGR	CGAACCCGCTCACCCAGTAG	
mtdM	mtdMF	GGTGCTCGGTGCCGATGTTG	508bp
	mtdMR	TTCATCTCCGCCAACCACG	
mtdW	mtdWF	TGACGGTTGGACCCATCGC	645bp
	mtdWR	TTGGCCTCGCCGTCAGACTC	

Table S4. Primer pairs used for mutant producer strain construction.

Gene target	Primer code	Primer pairs used for inactivation (5′-3′)
mtdB	BdF	tgggccgacgggtactccaacgcggtcttccgtctcatgATTCCGGGGATCCGTCGACC
	BdR	accgatgtcgtcgaaacggaccgtgcccagatccttgacTGTAGGCTGGAGCTGCTTC
mtdC	CdF	atcttcgtcctgggactggacgagcacaaccgcgagaccATTCCGGGGATCCGTCGAC
	CdR	gacgcaccggtcgtacttgtccttcagctcctcctcgtcTGTAGGCTGGAGCTGCTTC
mtdD	DdF	ctgatctcctggctcaacggcgtctacgccatgcagcggATTCCGGGGATCCGTCGAC
	DdR	cgacctgctcgggggcgtaatcggccagggcgtccttgaTGTAGGCTGGAGCTGCTTC
mtdM$_1$	M1dF	gccacggaggtcatggcgtcggcccttccctaccctgtcATTCCGGGGATCCGTCGAC
	M1dR	gtccccggccagccgttccagcagcgcgccgtcggcgagTGTAGGCTGGAGCTGCTTC
mtdH	HdF	cagaagagcctgaccacggccatcgaccaggtacaacccATTCCGGGGATCCGTCGACC
	HdR:	gacgtcccgcacggaatgggtgcgcccggtgccgatgacTGTAGGCTGGAGCTGCTTC
mtdG$_1$	G1dF	gaccacctcgacgcgctgtggacctcctccgagcactccATTCCGGGGATCCGTCGAC
	G1dR	ttcggcgtcgtcctcgcgtgcccggcgcaggacggcggtTGTAGGCTGGAGCTGCTTC
mtdJ	JdF	gggtgggtgctggcacgggggctcgccgagcgtgggcacATTCCGGGGATCCGTCGAC
	JdR	gacggcggccggatcggagggcaggagcggctcggtgagTGTAGGCTGGAGCTGCTTC
mtdK	KdF	ctcaccgccgacaggccggaggggttcttggagcaggtcATTCCGGGGATCCGTCGACC
	KdR	caggtggcggccataggcggacagaatggggcgaaacccTGTAGGCTGGAGCTGCTTC
mtdL	LdF	gacgtggtctgtccttccgtggcggaacagcaggatctcATTCCGGGGATCCGTCGACC
	LdR	accgttgacccgctcggcgatctcctggagggcataactTGTAGGCTGGAGCTGCTTC
mtdM	MdF	ctccacctccccgagttccgacccaccgaagccgacgagATTCCGGGGATCCGTCGACC
	MdR	gatgtcccgctcgatccagtggtacatggcgctcatcccTGTAGGCTGGAGCTGCTTC
MtdG$_2$	G2dF	agggtccctccctacctgccgttcgccgacgtggagcgcATTCCGGGGATCCGTCGAC
	G2dR	gaccacgttgggcacgccgtgttccagtgcggcgatgacTGTAGGCTGGAGCTGCTTC
mtdM$_2$	M2dF	gacggccacttcgacctgatcttctgtctcagcgtcctgATTCCGGGGATCCGTCGAC
	M2dR	attgcggtagacctcgcggaacgagtcgggctggcaggcTGTAGGCTGGAGCTGCTTC
mtdM$_3$	M3dF	cggtgggacaccatgatgctgccctggttccgcaagtacATTCCGGGGATCCGTCGAC
	M3dR	cctgaggaggacggggtcatagccccaggaacggtagagTGTAGGCTGGAGCTGCTTC
mtdW	WdF	atcaacgggctgatcgccatgtggccgatgcccgaggacATTCCGGGGATCCGTCGAC
	WdR	cgcctcgtggcagcgggccgccatccaggagtcgagttcTGTAGGCTGGAGCTGCTTC
mtdM$_4$	M4dF	ccgggcaccgtgaaggtgatgaaacgcgtggtcgaacccATTCCGGGGATCCGTCGAC
	M4dR	cgggatctcctcacgttccgacccgtcatcctggatcgcTGTAGGCTGGAGCTGCTTC
mtdWM$_4$	WdF	atcaacgggctgatcgccatgtggccgatgcccgaggacATTCCGGGGATCCGTCGAC
	M4dR	cgggatctcctcacgttccgacccgtcatcctggatcgcTGTAGGCTGGAGCTGCTTC
orf1	orf1dF	gcacgctacggggccgctgacctccaggtccgtgacaaaATTCCGGGGATCCGTCGACC
	orf1dR	ggcgatgtcccagaaaccgacgatcggctcggccgcgagTGTAGGCTGGAGCTGCTTC
orf2	orf2dF	gagcagcgcggaaccaggaagaagaagtccgcgttggcgATTCCGGGGATCCGTCGAC
	orf2dR	gtagcgacgcagcctgcggtactcgatgaggtcaccgtcTGTAGGCTGGAGCTGCTTC
mtdG$_1$*WM*$_4$	WdelAprF	ctcattctcccgatcagggaaagcgcctccgatcgcaagACTAGTATTCCGGGGATCCGTCGACC
	M4delAprR	cgggatctcctcacgttccgacccgtcatcctggatcgcACTAGTTGTAGGCTGGAGCTGCTTC
	G1dF	gaccacctcgacgcgctgtggacctcctccgagcactccATTCCGGGGATCCGTCGAC
	G1dR	ttcggcgtcgtcctcgcgtgcccggcgcaggacggcggtTGTAGGCTGGAGCTGCTTC
mtdM$_2$*M*$_3$	mtdspelAprM2F	gagtaccccctgggccgccacccacgccggtctgcgcccgACTAGTATTCCGGGGATCCGTCGAC
	mtdspelAprM2R	cagcgtccgccacctcagggcgcgttccacgaccgcctcACTAGTTGTAGGCTGGAGCTGCTTC
	M3dF	cggtgggacaccatgatgctgccctggttccgcaagtacATTCCGGGGATCCGTCGAC
	M3dR	cctgaggaggacggggtcatagccccaggaacggtagagTGTAGGCTGGAGCTGCTTC

[a] small letters were derived from targeted genes

Table S5. Primer pairs used for PCR confirmation of the double-crossover mutants.

Gene	Primer code	Primer pairs designed to verify the mutants (5′-3′)	Length of desired PCR fragments	
			Wild strain	Mutant strain
mtdB	BtF	TGACCGTCGTGACCGACCAG	1460	1857
	BtR	GTGCTCCACCGACGACCGT		
mtdC	CtF	CAGCTCCTGCTCCAGGTACTC	1892	2115
	CtR	GGACCTCTTCCATGCCGTCG		
mtdD	DtF	CGGTTGTGCTCGTCCAGTCCCA	894	2029
	DtR	CCGGACCGTCAGCGGAAAGG		
mtdM$_1$	M1tF	AGTACGACACCCGATCACGA	914	1905
	M1tR	GAGCTGGTGCGGATGATGGT		
mtdH	HtF	ACCATCATCCGCACCAGCTC	994	1835
	HtR	GGTTCACCGTGCATGACGC		
mtdG$_1$	G1tF	TGGCCGAACTCTACTGTCATGG	1108	1895
	G1tR	CGCCCTGGAGAATCTGCGTC		
mtdJ	JtF	GACGCAGATTCTCCAGGGCG	1206	1837
	JtR	ACGCTCCTCCTCGTGACCC		
mtdK	KtF	GGAAGTGTTCGTAGTCGGGT	1359	1936
	KtR	CACCTGATGCGTTCCTGGACG		
mtdL	LtF	AGGTAACGGATCTTCAGCAGG	1468	2057
	LtR	ACTGGATCGAGCGGGACATC		
mtdM	MtF	TCGGGGATCACGATCAGCCG	1208	1806
	MtR	CGCTTTCACCATCACGTTCG		
MtdG$_2$	G2tF	CGTGTTCTCTTCTTCCCGTGG	1127	1806
	G2tR	ATGTCCACCGAGGCGACGT		
mtdM$_2$	M2tF	GACCGCAACGACAACGTCAA	1169	1917
	M2tR	TTCCAGCACGTGGGCCTC		
mtdM$_3$	M3tF	ACCTACCGGATCGTCATCGAA	965	1806
	M3tR	TACGACGAGCACTAGGTCCT		
mtdW	WtF	AACCAAGGAGGAAAGAGCGA	1512	1928
	WtR	GGACGAGACGGATGTGTCATGT		
mtdM$_4$	M4tF	CGCCGACCACATGACACATCC	1062	1918
	M4tR	GGAAGTTCTTGGTCCCGTCG		
mtdWM$_4$	WtF	AACCAAGGAGGAAAGAGCGA	2544	2094
	M4tR	GGAAGTTCTTGGTCCCGTCG		
orf1	orf1tF	GTGACGACTACATCAACCTCCTG	799	1670
	orf1tR	GCGAGGACGCTGGTGATGC		
orf2	orf2tF	GTCGAAGCCGCAGGTCATCC	563	1758
	orf2tR	CGAGGTCTGCTGAACGACGG		
mtdG$_1$WM$_4$	WtF2	GAGGATCTGGCTGACGGTAC	2812	657
	M4tR	GGAAGTTCTTGGTCCCGTCG		
	G1tF	TGGCCGAACTCTACTGTCATGG	1108	1895
	G1tR	CGCCCTGGAGAATCTGCGTC		
mtdM$_2$M$_3$	M2tF	GACCGCAACGACAACGTCAA	1169	410

	M2tR	TTCCAGCACGTGGGCCTC		
	M3tF	ACCTACCGGATCGTCATCGAA	965	1806
	M3tR	TACGACGAGCACTAGGTCCT		

Table S6. Mutant strains generated in this study.

Wild strain	A201A producing strain
Ju3003	*mtdB gene* disrupted mutant of SCSIO 00652
Ju3004	*mtdC gene* disrupted mutant of SCSIO 00652
Ju3005	*mtdD gene* disrupted mutant of SCSIO 00652
Ju3006	*mtdM$_1$ gene* disrupted mutant of SCSIO 00652
Ju3007	*mtdH gene* disrupted mutant of SCSIO 00652
Ju3008	*mtdG$_1$ gene* disrupted mutant of SCSIO 00652
Ju3009	*mtdJ gene* disrupted mutant of SCSIO 00652
Ju3010	*mtdK gene* disrupted mutant of SCSIO 00652
Ju3011	*mtdL gene* disrupted mutant of SCSIO 00652
Ju3012	*mtdM gene* disrupted mutant of SCSIO 00652
Ju3013	*MtdG$_2$ gene* disrupted mutant of SCSIO 00652
Ju3014	*mtdM$_2$ gene* disrupted mutant of SCSIO 00652
Ju3015	*mtdM$_3$ gene* disrupted mutant of SCSIO 00652
Ju3016	*mtdW gene* disrupted mutant of SCSIO 00652
Ju3017	*mtdM$_4$ gene* disrupted mutant of SCSIO 00652
Ju3018	*mtdWM$_4$ gene* disrupted mutant of SCSIO 00652
Ju3019	*orf1 gene* disrupted mutant of SCSIO 00652
Ju3020	*orf2 gene* disrupted mutant of SCSIO 00652
Ju3023	mutant of SCSIO 00652 *in which mtdWM$_4$ was in frame deleted*
Ju3024	mutant of SCSIO 00652 *in which mtdWM$_4$ was in frame deleted and mtdG$_1$ was replaced by aac(3)IV-oriT* fragment
Ju3025	mutant of SCSIO 00652 *in which mtdM$_2$ was in frame deleted and mtdM$_3$ was replaced by aac(3)IV-oriT* fragment

Table S7. Primer pairs used for site-directed mutagenesis of MtdL.

Name of sequence	Sequence (5'-3')
652-mtdL-D109A	CGTCTCGATGGCGGACGACAACCTC
652-mtdL-D109Aanti	GAGGTTGTCGTCCGCCATCGAGACG
652-mtdL-D110A	CTCGATGGACGCGGACAACCTCCCC
652-mtdL-D110Aanti	GGGGAGGTTGTCCGCGTCCATCGAG
652-mtdL-D111A	CGATGGACGACGCGAACCTCCCCAC
652-mtdL-D111Aanti	GTGGGGAGGTTCGCGTCGTCCATCG
652-mtdL-R159A	GAGGTCTTCCCCGCGGGTTTCCCCTTC
652-mtdL-R159Aanti	GAAGGGGAAACCCGCGGGGAAGACCTC

Table S8. ^{1}H and ^{13}C NMR spectral data of 1 in DMSO-d_6 and CD_3OD, and 2 and 3 in CD_3OD.

		1 (DMSO-d_6)		**1** (CD_3OD)		**2**		**3**	
		δ_C	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C	δ_H
Adenine moiety	2	151.8	8.24, s	153.0	8.21, s	153.5	8.24, s	151.9	8.23, s
	4	149.6		150.6		150.7		149.5	
	5	119.6		121.6		121.7		121.0	
	6	154.2		156.1		156.2		156.6	
	8	137.8	8.48, s	139.2	8.38, s	139.3	8.41, s	138.0	8.41, s
	10-CH_3	37.7	3.37, br s	39.2	3.49, br s	39.1	3.53, br s	38.5	3.52, br s
Aminopentose moiety	1	89.3	6.05, d (3.0)	92.0	6.11, d (3.1)	92.1	6.11, d (3.0)	91.5	6.11, d (3.0)
	2	73.1	4.53, dd (5.6, 3.0)	75.1	4.73, dd (5.8, 3.1)	75.2	4.69, dd (6.0, 3.0)	74.0	4.70, m
	3	50.7	4.59, m	52.6	4.80 *	52.6	4.76, t (6.0)	51.5	4.77, t (6.3)
	3-NH		7.88, d (7.2)						
	4	82.7	4.22, m	85.2	4.30, m	85.2	4.33, m	84.0	4.33, m
	5	60.7	3.77, dd (12.1, 1.8); 3.59, dd (12.1, 3.4)	62.6	4.00, dd (12.5, 1.5); 3.90, dd (12.5, 2.6)	62.6	3.99, dd (12.5, 2.0); 3.82, dd (12.5, 3.0)	62.6	3.99, dd (12.5, 1.5); 3.82, dd (13.0, 2.5)
Aromatic acid moiety	1	129.9		131.9		130.0		130.5	
	2,6	130.7	7.41, d (8.8)	132.1	7.40, d (8.6)	132.4	6.84, d (8.5)	132.1	7.39, d (8.5)
	3,5	116.3	7.12, d (8.8)	117.8	7.20, d (8.6)	116.3	7.31, d (8.5)	118.0	7.18, d (8.5)
	4	156.0		158.0		158.9		156.2	
	7	132.2	7.27, s	135.0	7.32, s	135.9	7.32, s	135.3	7.34, s
	8	130.7		132.1		128.7		131.3	
	9	169.3		173.1		173.5		171.3	
	8-CH_3	14.3	2.05, s	14.7	2.13, s	14.6	2.15, s	14.7	2.13, s
Hexofuranose moiety	1	99.3	5.87, d (4.2)	101.4	5.88, d (4.1)			100.3	5.60, d (4.5)
	2	76.3	4.08, m	78.3	4.25, dd (5.3, 4.1)			78.7	4.18, dd (8.0, 4.5)
	3	72.7	4.57, m	74.7	4.80, m			74.5	4.32, t (7.5)
	4	143.1		144.7				83.9	3.85, dd (7.0, 5.0)
	5	132.0		134.3				74.0	3.68, dd (11.0, 5.0)
	6	62.3	4.31, d (12.5); 4.00, d (12.5)	64.3	4.43, d (12.5); 4.16, d (12.5)			63.8	3.63, d (7.0, 4.5); 3.56, d (11.0, 6.0)
	5-OMe	57.3	3.49, s	61.0	3.65, s				
Rhamnose moiety	1	99.6	4.67, br d	100.8	4.80 *				
	2	66.1	3.84, s	68.4	4.03, s				
	3	80.7	3.22, dd (9.3, 3.1)	82.4	3.41, dd (9.4, 3.1)				
	4	81.2	3.02, t (9.3)	83.1	3.16, t (9.4)				
	5	67.0	3.48, m	69.0	3.68, m				
	6	17.7	1.17, d (6.2)	18.2	1.29, d (6.5)				
	3-OMe	55.9	3.29, s	57.2	3.43, s				
	4-OMe	59.8	3.39, s	59.0	3.52, s				

* Overlapped

Table S9. ^{1}H and ^{13}C NMR spectral data of 4 in $CD_3OD/CDCl_3$ and 5–7 in CD_3OD.

		4		5		6		7	
		δ_C	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C	δ_H
Adenine moiety	2	151.9	8.25, s	153.0	8.22, s	153.8	8.24, s	152.9	8.19, s
	4	148.9		150.7		150.0		150.5	
	5	121.1		121.7		120.7		121.5	
	6	155.2		156.2		157.5		156.0	
	8	137.7	8.34, s	139.3	8.40, s	141.2	8.48, s	139.2	8.38, s
	10-CH_3	38.9	3.53, br s	39.1	3.50, s			39.1	3.45, s
Aminopentose moiety	1	91.4	6.01, d (2.9)	92.1	6.11, d (3.0)	92.2	6.13, d (3.0)	92.0	6.10, d (1.8)
	2	74.5	4.59, dd (6.0, 2.9)	75.2	4.72, dd (6.0, 3.0)	75.3	4.73, dd (5.9, 3.0)	75.1	4.72, dd (5.7, 3.0)
	3	50.7	4.64, t (6.0)	52.7	4.77, d (6.0)	52.6	4.79, m	52.5	4.77, m
	3-NH								
	4	84.9	4.30, m	85.2	4.35, br d (7.0)	85.2	4.34, m	85.1	4.34, m
	5	61.6	4.05, dd (12.9, 1.5); 3.87, dd (12.9, 2.2)	62.7	4.00, br d (12.0); 3.84, br d (12.0)	62.6	3.99, dd (12.5, 1.8); 3.83, dd (12.5, 3.0)	62.6	4.00, br d (11.3); 3.82, br d (11.3)
Aromatic acid moiety	1	130.0		131.9		131.9		131.8	
	2,6	131.3	7.35, d (8.5)	132.1	7.41, d (8.5)	132.1	7.40, dd (7.2, 2.0)	132.1	7.37, d (8.2)
	3,5	117.1	7.17, d (8.5)	117.9	7.21, d (8.5)	117.8	7.21, dd (7.2, 2.0)	117.8	7.18, d (8.2)
	4	156.9		158.1		158.1		158.0	
	7	135.0	7.36, s	135.0	7.33, s	135.1	7.34, s	135.1	7.31, s
	8	130.6		132.1		131.9		132.1	
	9	171.3		173.2		173.2		173.1	
	8-CH_3	14.2	2.13, s	14.6	2.14, s	14.6	2.14, s	14.7	2.12, s
Hexofuranose moiety	1	100.4	5.82, d (4.1)	101.5	5.89, d (4.2)	101.4	5.89, d (4.2)	101.9	5.88, d (3.7)
	2	77.0	4.23, m	78.3	4.24, t (4.2)	78.3	4.23, dd (5.4, 4.2)	78.2	4.24, m
	3	75.6	4.81, br d (5.4)	74.7	4.80, m	74.7	4.78, m	74.7	4.78, br d (8.7)
	4	141.7		144.9		144.7		144.7	
	5	136.1		134.3		134.4		134.3	
	6	58.6	4.26, d (13.4); 4.16, d (13.4)	64.2	4.45, d (12.5); 4.19, d (12.5)	64.3	4.43, d (12.5); 4.15, d (12.5)	64.2	4.46, d (12.5); 4.16, d (12.5)
	5-OMe	58.6	3.70, s	59.1	3.65, s	59.0	3.65, s	59.0	3.65, s
Rhamnose moiety	1			100.7	4.92, s	100.8	4.83, d (1.4)	101.3	4.86, br s
	2			68.4	4.02, br s	68.4	4.02, dd (3.0, 1.4)	68.1	4.05, br s
	3			82.7	3.46 *	82.5	3.40, m	82.1	3.33, m
	4			77.2	3.45 *	83.1	3.14, t (9.4)	72.8	3.45, m

5	73.9	3.58, m	69.0	3.66, m	70.0	3.68, m
6	62.5	3.72, dd (12.0, 5.0); 3.84, br d (12.0)	18.2	1.29, d (6.2)	18.1	1.29, d (6.1)
3-OMe	57.2	3.44, s	57.2	3.43, s	57.3	3.43, s
4-OMe	60.8	3.52, s	61.1	3.52, s		

* Overlapped

Table S10. ^{1}H and ^{13}C NMR spectral data of 8–11 in CD_3OD.

		8		9		10		11	
		δ_C	δ_H	δ_C	δ_H	δ_C	δ_H	δ_C	δ_H
Adenine moiety	2	153.1	8.24, s	153.0	8.24, s	153.0	8.24, s	153.1	8.25, s
	4	150.8		150.7		150.1		150.7	
	5	121.7		121.7		121.7		121.7	
	6	156.4		156.3		156.6		156.3	
	8	139.3	8.41, s	139.3	8.41, s	139.3	8.41, s	139.4	8.41, s
	10-CH_3	39.1	3.52, s	39.1	3.49, s	39.1	3.49, s	39.3	3.54, br s
Aminopentose moiety	1	92.1	6.12, d (3.4)	92.1	6.11, d (3.5)	92.1	6.11, d (3.0)	92.1	6.12, d (3.2)
	2	75.2	4.72, dd (5.9, 3.2)	75.2	4.71, m	75.1	4.70, m	75.1	4.70, dd (6.0, 3.3)
	3	52.7	4.77, m	52.6	4.77, m	52.6	4.77, t (6.5)	52.6	4.77, m
	3-NH								
	4	85.3	4.34, dt (6.7, 2.5)	85.2	4.34, m	85.2	4.33, m	85.1	4.34, dt (7.0, 2.4)
	5	62.7	4.00, dd (12.4, 1.3); 3.84, dd (12.4, 2.5)	62.6	3.96, dd (13.1, 1.7); 3.82, dd (13.1, 3.0)	62.6	3.96, dd (12.5, 2.0); 3.82, m	62.6	3.99, dd (12.5, 1.5); 3.82, dd (12.5, 2.4)
Aromatic acid moiety	1	130.6		131.5		130.5		131.9	
	2,6	132.1	7.43, d (8.6)	132.0	7.39, dd (8.7, 2.0)	132.1	7.39, d (8.5)	132.1	7.43, d (8.7)
	3,5	118.0	7.22, d (8.6)	117.9	7.19, dd (8.7, 2.0)	118.0	7.18, d (8.5)	117.9	7.22, d (8.7)
	4	158.1		158.4		158.8		158.1	
	7	135.0	7.35, s	135.3	7.30, s	135.3	7.34, s	135.0	7.35, s
	8	132.0		132.1		131.3		132.0	
	9	173.2		173.3		173.3		173.3	
	8-CH_3	14.6	2.15, s	14.7	2.13, s	14.7	2.13, s	14.6	2.15, d (1.2)
Hexofuranose moiety	1	101.6	5.89, d (4.1)	101.4	5.61, d (4.5)	100.3	5.88, d (4.0)	101.5	5.89, d (4.1)
	2	78.4	4.23, m	78.8	4.18, dd (7.5, 4.5)	78.7	4.18, dd (7.5, 4.0)	78.3	4.23, dd (5.4, 4.3)
	3	74.8	4.70, br d (10.4)	76.1	4.27, t (7.5)	75.6	4.32, t (7.5)	74.7	4.78, d (5.5)
	4	144.8		83.5	3.90, t (7.5)	83.9	3.88, dd (7.5, 4.5)	144.8	
	5	134.5		83.0	3.37, m	72.0	3.80, m	134.4	
	6	64.5	4.44, d (12.5); 4.16, d (12.5)	67.2	3.83, d (11.0); 3.47, d (11.0, 6.0)	69.3	3.75, m; 3.40, m	64.3	4.46, d (12.5); 4.15, d (12.5)
	5-OMe	59.1	3.65, s	59.7	3.33, s			59.1	3.66, s
Rhamnose moiety	1	101.0	4.79, br s	101.9	4.69, d (1.4)	101.9	4.71, d (1.5)	101.1	4.81, d (1.2)
	2	72.4	3.81, m	68.3	4.01, dd (3.1, 1.4)	68.2	4.05, dd (2.0, 1.5)	72.3	3.83, m
	3	72.6	3.76, dd (9.2, 3.4)	82.5	3.40, m	82.4	3.42, m	72.4	3.67, m
	4	84.5	3.11, t (9.2)	83.1	3.12, t (9.4)	83.1	3.13, t (9.5)	74.1	3.39, t (9.5)
	5	69.0	3.66, m	68.9	3.61, m	68.9	3.62, m	70.1	3.68, m
	6	18.3	1.30, d (6.3)	18.2	1.28, d (6.5)	18.2	1.28, d (6.5)	18.1	1.30, d (6.5)
	3-OMe			57.2	3.43, s	57.2	3.43, s		
	4-OMe	61.0	3.56, s	61.1	3.52, s	61.1	3.52, s		

Table S11. Antibacterial activities (MIC, μg/mL) of compounds 1, 2, and 4–11.

Compound	MRSA[a]	MRSE[b]	*Klebsiella pneumoniae* ATCC13883	*Escherichia coli* ATCC 25922	*Staphylococcus aureus* ATCC 29213	*Aeromonas hydrophila* ATCC 7966	*Micrococcus luteus*	*Bacillus thuringiensis*
1	2	16	> 64	> 64	2	16	8	2
2	> 64	> 64	> 64	> 64	> 64	64	> 64	> 64
4	> 64	64	> 64	> 64	> 64	> 64	> 64	> 64
5	32	32	> 64	> 64	32	> 64	> 64	16
6	4	16	> 64	> 64	4	64	4	2
7	16	> 64	> 64	> 64	16	64	64	8
8	8	> 64	> 64	> 64	8	> 64	16	4
9	32	32	> 64	> 64	32	32	> 64	8
10	64	32	> 64	> 64	> 64	64	> 64	64
11	> 64	64	> 64	> 64	64	> 64	> 64	> 64
Ampicillin	> 64	8	> 64	8	4	> 128	<0.125	8
Erythromycin	> 64	2	32	64	0.25	32	0.25	0.25
Vancomycin	1	2	> 64	> 64	1	> 64	0.5	0.5

[a] MRSA: methicillin-resistant *Staphylococcus aureus* (clinical isolate shhs-A1)
[b] MRSE: methicillin-resistant *Staphylococcus epidermidis* (clinical isolate shhs-E1)

Supplementary Figures S1–S91

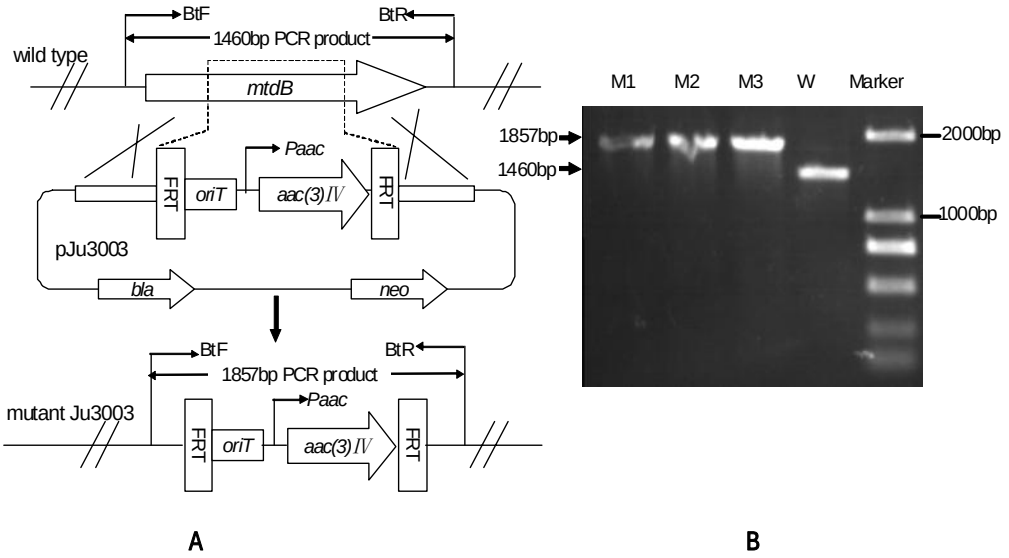


Fig. S1. Construction and gel electrophoresis analyses of mutant Δ*mtdB* (Ju3003). (A) Construction of mutant strain Δ*mtdB* and predicted PCR fragment size from wild-type and mutant; (B) Verification of the Δ*mtdB* mutant by PCR. DNA templates were from *M. thermotolerans* SCSIO 00652 (lane W) and Δ*mtdB* (lane M1, M2, M3). Marker, DNA marker DL2000.

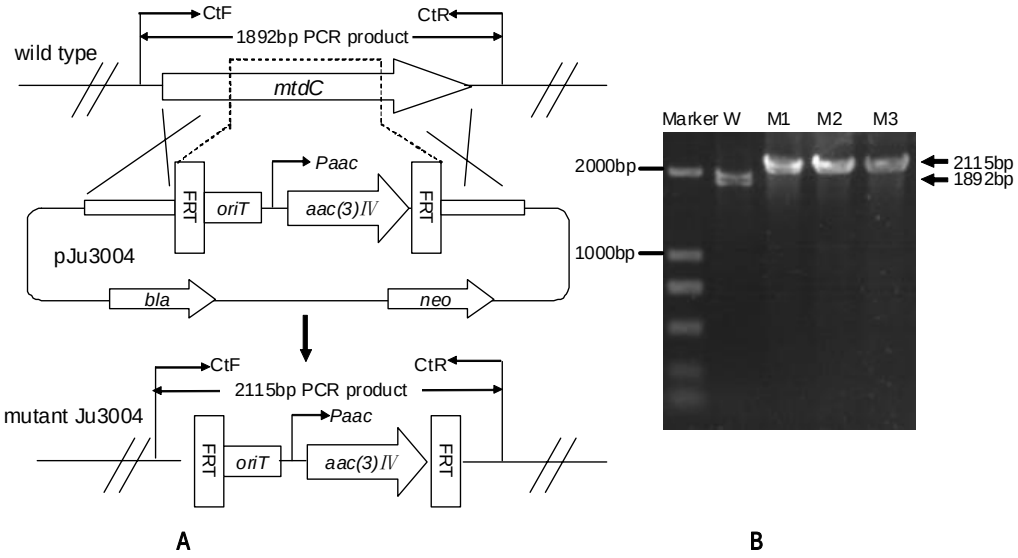


Fig. S2. Construction and gel electrophoresis analyses of mutant Δ*mtdC* (Ju3004). (A) Construction of mutant Δ*mtdC* and predicted PCR fragment size from wild-type and mutant; (B) Verification of the Δ*mtdC* mutant by PCR. DNA templates were from *M. thermotolerans* SCSIO 00652 (lane W) and Δ*mtdC* (lane M1, M2, M3). Marker, DNA marker DL2000.

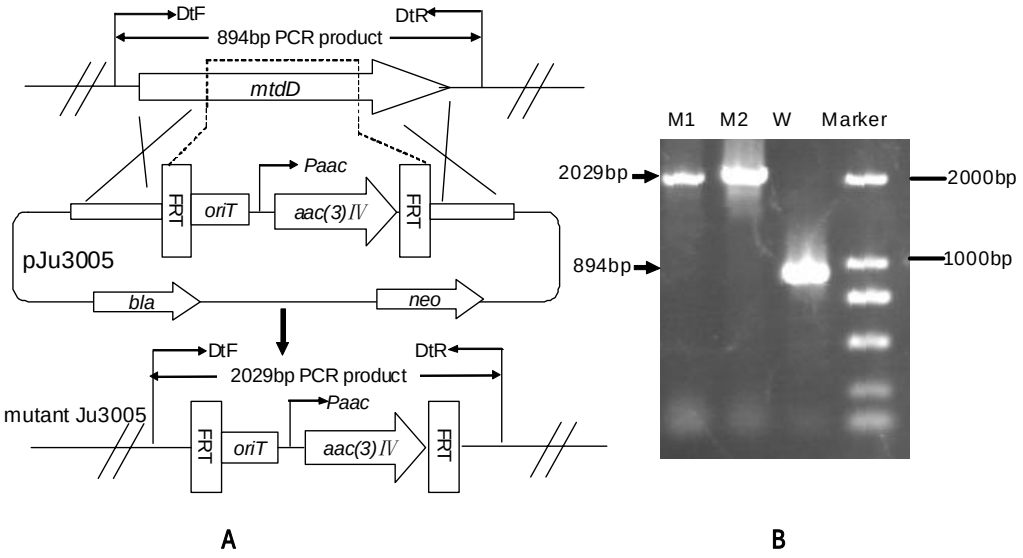


Fig. S3. Construction and gel electrophoresis analyses of mutant Δ*mtdD* (Ju3005). (A) Construction of mutant Δ*mtdD* and predicted PCR fragment size from wild-type and mutant; (B) Verification of the Δ*mtdD* mutant by PCR. DNA templates were from *M. thermotolerans* SCSIO 00652 (lane W) and Δ*mtdD* (lane M1, M2). Marker, DNA marker DL2000.

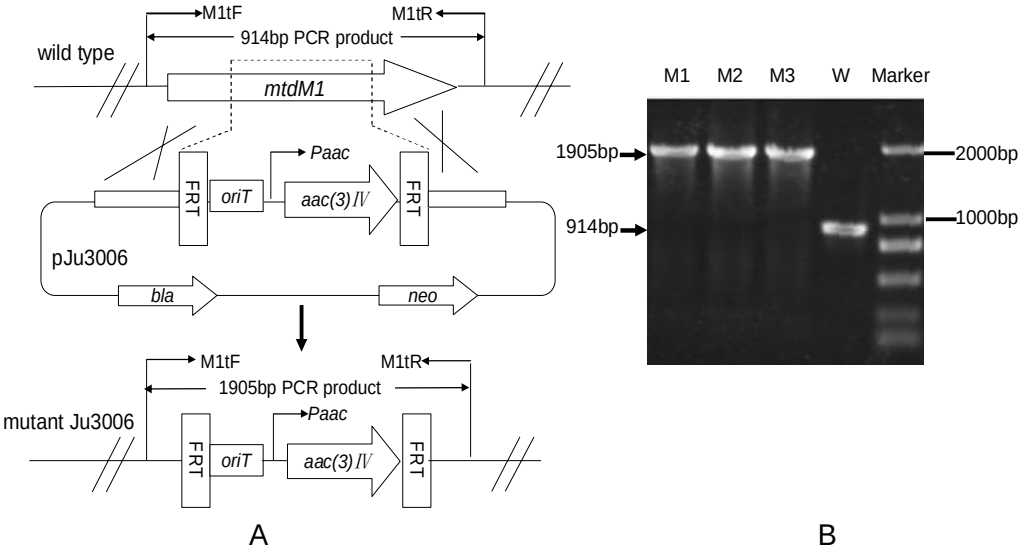


Fig. S4. Construction and gel electrophoresis analyses of mutant Δ*mtdM*$_1$ (Ju3006). (A) Construction of mutant Δ*mtdM*$_1$ and predicted PCR fragment size from wild-type and mutant; (B) Verification of the Δ*mtdM*$_1$ mutant by PCR. DNA templates were from *M. thermotolerans* SCSIO 00652 (lane W) and Δ*mtdM*$_1$ (lane M1, M2, M3). Marker, DNA marker DL2000.

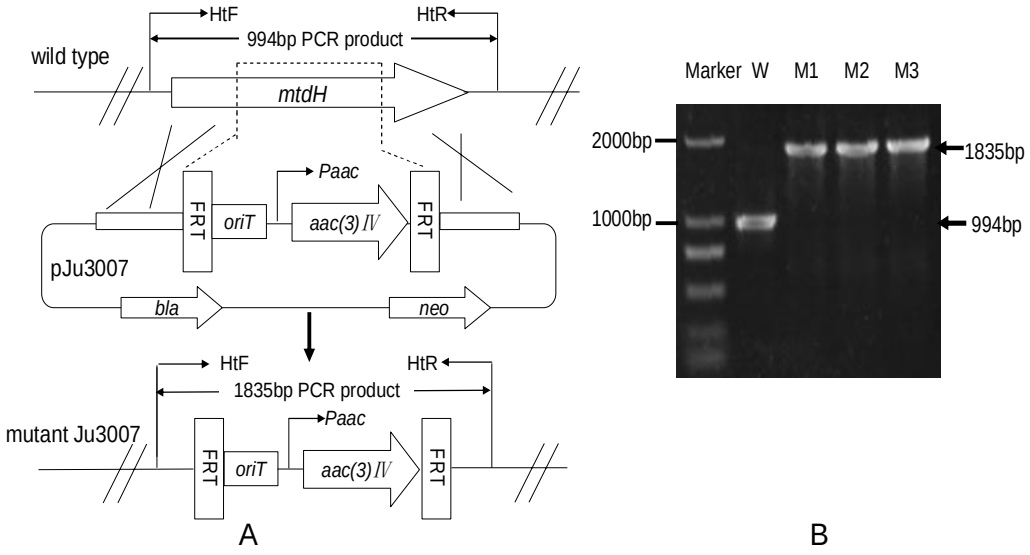


Fig. S5. Construction and gel electrophoresis analyses of mutant Δ*mtdH* (Ju3007). (A) Construction of mutant Δ*mtdH* and predicted PCR fragment size from wild-type and mutant; (B) Verification of the Δ*mtdH* mutant by PCR. DNA templates were from *M. thermotolerans* SCSIO 00652 (lane W) and Δ*mtdH* (lane M1, M2, M3). Marker, DNA marker DL2000.

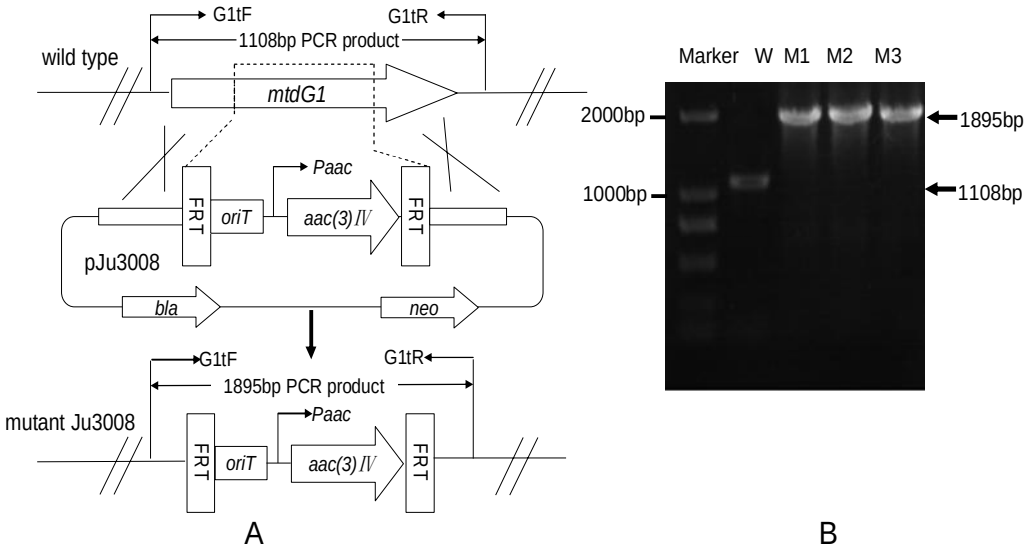


Fig. S6. Construction and gel electrophoresis analyses of mutant Δ*mtdG*$_1$ (Ju3008). (A) Construction of mutant Δ*mtdG*$_1$ and predicted PCR fragment size from wild-type and mutant; (B) Verification of the Δ*mtdG*$_1$ mutant by PCR. DNA templates were from *M. thermotolerans* SCSIO 00652 (lane W) and Δ*mtdG*$_1$ (lane M1, M2, M3). Marker, DNA marker DL2000.

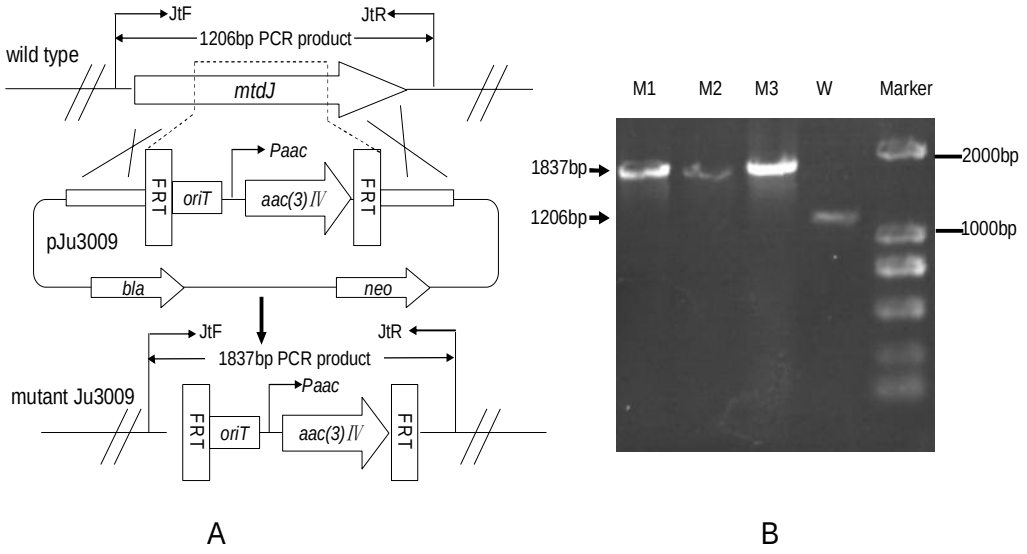


Fig. S7. Construction and gel electrophoresis analyses of mutant Δ*mtdJ* (Ju3009). (A) Construction of mutant Δ*mtdJ* and predicted PCR fragment size from wild-type and mutant; (B) Verification of the Δ*mtdJ* mutant by PCR. DNA templates were from *M. thermotolerans* SCSIO 00652 (lane W) and Δ*mtdJ* (lane M1, M2, M3). Marker, DNA marker DL2000.

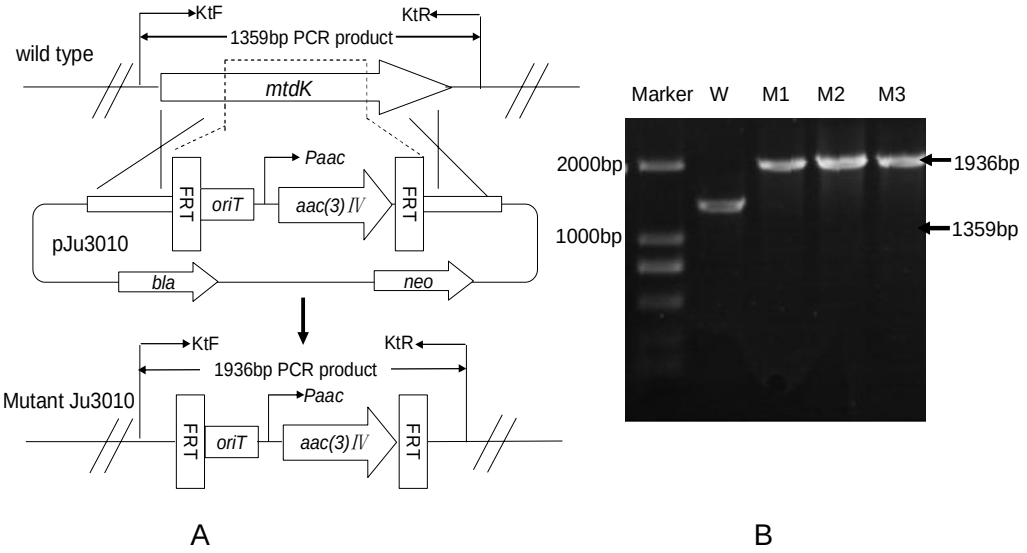


Fig. S8. Construction and gel electrophoresis analyses of mutant Δ*mtdK* (Ju3010). (A) Construction of mutant Δ*mtdK* and predicted PCR fragment size from wild-type and mutant; (B) Verification of the Δ*mtdK* mutant by PCR. DNA templates were from *M. thermotolerans* SCSIO 00652 (lane W) and Δ*mtdK* (lane M1, M2, M3). Marker, DNA marker DL2000.

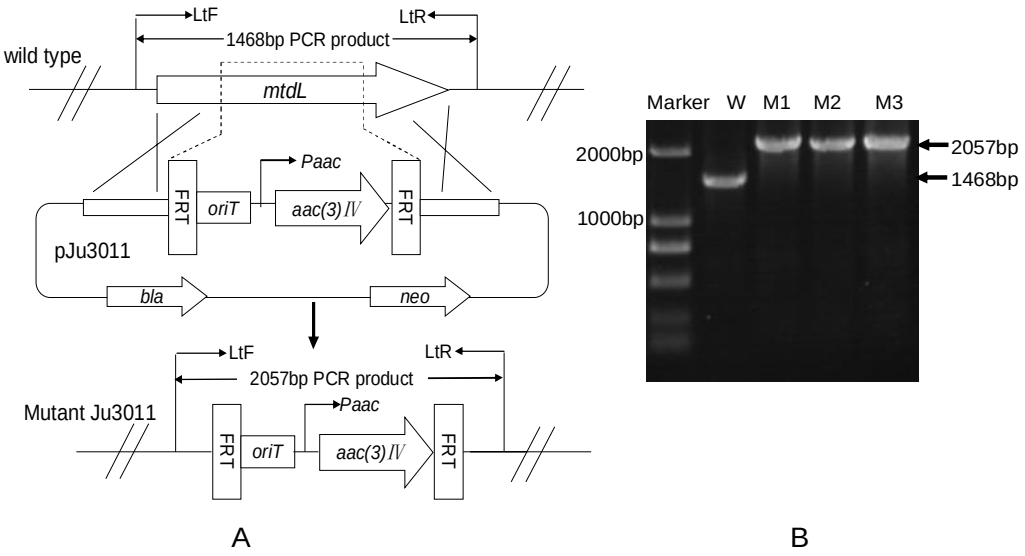


Fig. S9. Construction and gel electrophoresis analyses of mutant Δ*mtdL* (Ju3011). (A) Construction of mutant Δ*mtdL* and predicted PCR fragment size from wild-type and mutant; (B) Verification of the Δ*mtdL* mutant by PCR. DNA templates were from *M. thermotolerans* SCSIO 00652 (lane W) and Δ*mtdL*(lane M1, M2, M3). Marker, DNA marker DL2000.

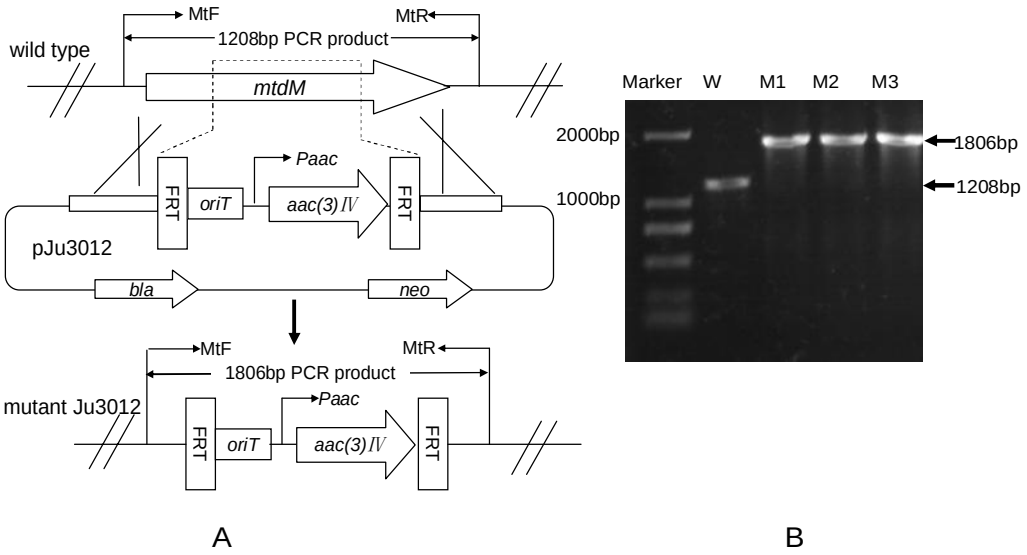


Fig. S10. Construction and gel electrophoresis analyses of mutant Δ*mtdM* (Ju3012). (A) Construction of mutant Δ*mtdM* and predicted PCR fragment size from wild-type and mutant; (B) Verification of the Δ*mtdM* mutant by PCR. DNA templates were from *M. thermotolerans* SCSIO 00652 (lane W) and Δ*mtdM* (lane M1, M2, M3). Marker, DNA marker DL2000.

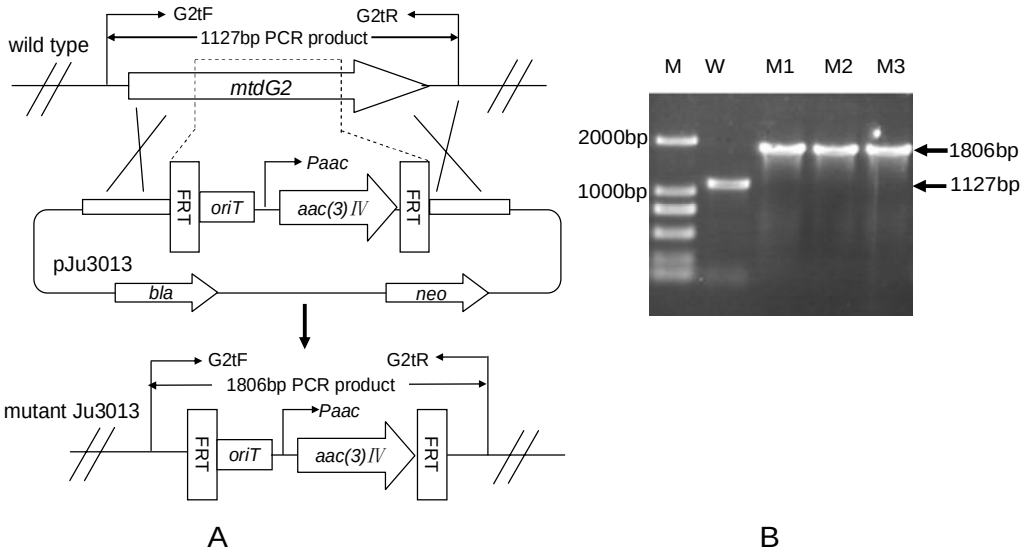


Fig. S11. Construction and gel electrophoresis analyses of mutant Δ*mtdG*$_2$(Ju3013). (A) Construction of mutant Δ*mtdG*$_2$ and predicted PCR fragment size from wild-type and mutant; (B) Verification of the Δ*mtdG*$_2$ mutant by PCR. DNA templates were from *M. thermotolerans* SCSIO 00652 (lane W) and Δ*mtdG*$_2$ (lane M1, M2, M3). M, DNA marker DL2000.

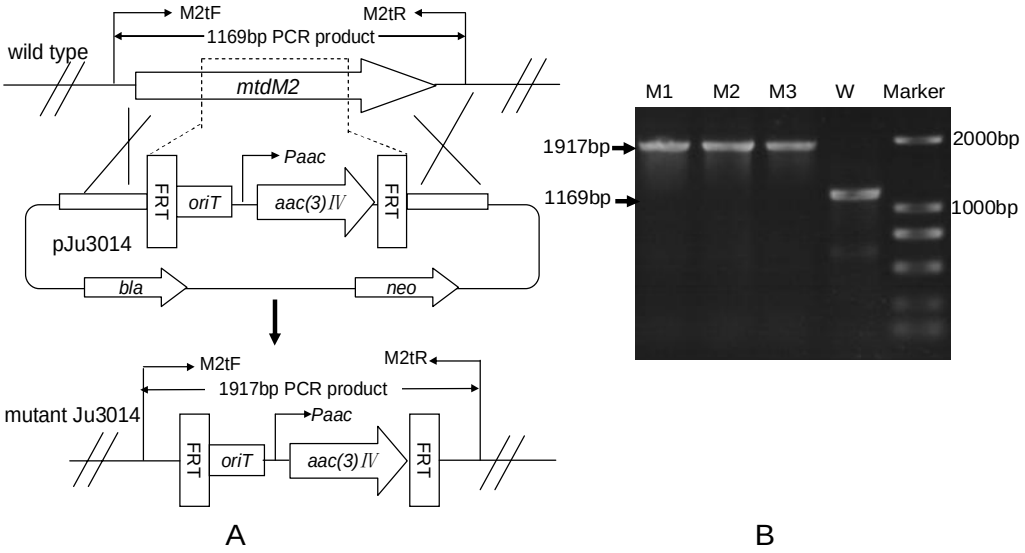


Fig. S12. Construction and gel electrophoresis analyses of mutant Δ*mtdM*$_2$ (Ju3014). (A) Construction of mutant Δ*mtdM*$_2$ and predicted PCR fragment size from wild-type and mutant; (B) Verification of the Δ*mtdM*$_2$ mutant by PCR. DNA templates were from *M. thermotolerans* SCSIO 00652 (lane W) and Δ*mtdM*$_2$ (lane M1, M2, M3). Marker, DNA marker DL2000.

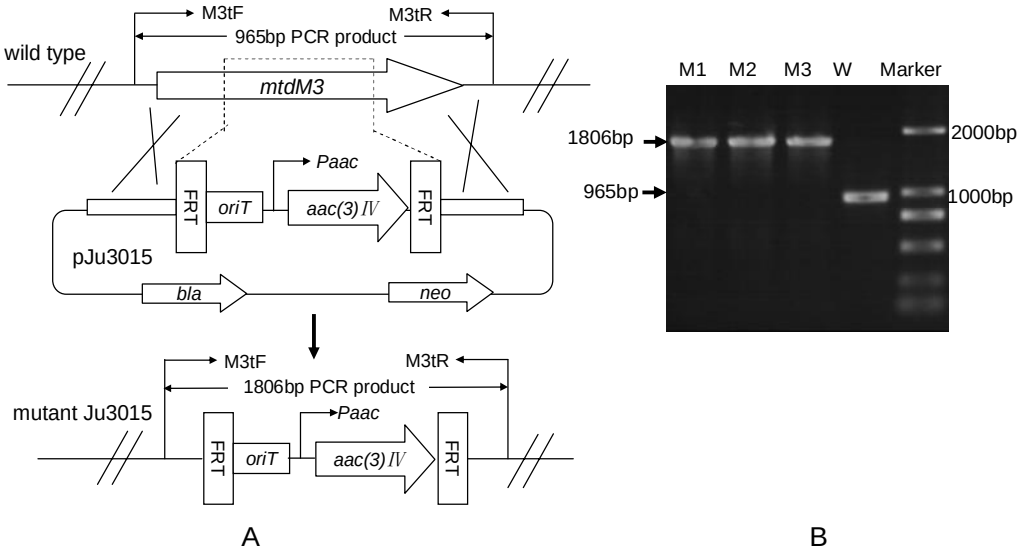


Fig. S13. Construction and gel electrophoresis analyses of mutant Δ*mtdM3* (Ju3015). (A) Construction of mutant Δ*mtdM*$_3$ and predicted PCR fragment size from wild-type and mutant; (B) Verification of the Δ*mtdM*$_3$ mutant by PCR. DNA templates were from *M. thermotolerans* SCSIO 00652 (lane W) and Δ*mtdM*$_3$ (lane M1, M2, M3). Marker, DNA marker DL2000.

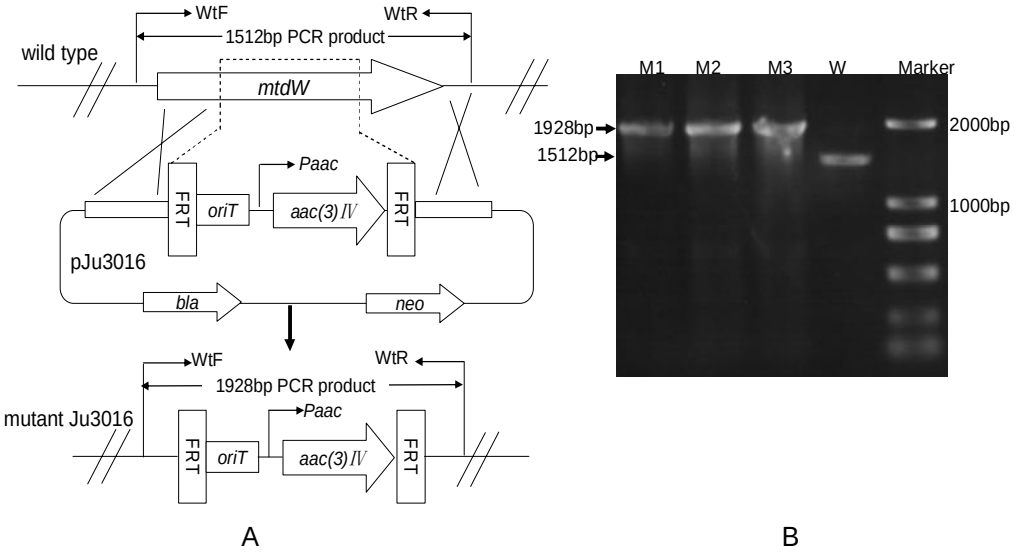


Fig. S14. Construction and gel electrophoresis analyses of mutant Δ*mtdW* (Ju3016). (A) Construction of mutant Δ*mtdW* and predicted PCR fragment size from wild-type and mutant; (B) Verification of the Δ*mtdW* mutant by PCR. DNA templates were from *M. thermotolerans* SCSIO 00652 (lane W) and Δ*mtdW* (lane M1, M2, M3). Marker, DNA marker DL2000.

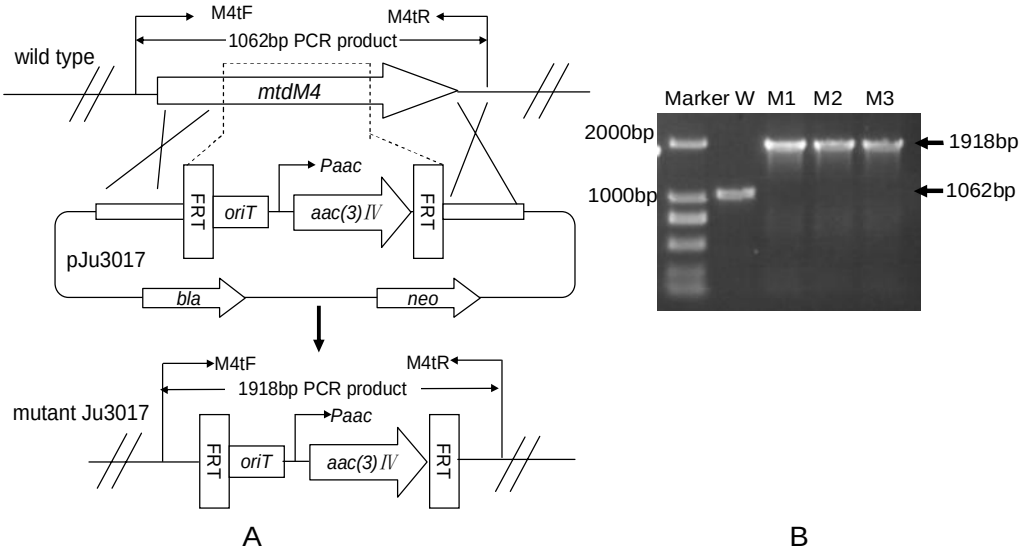


Fig. S15. Construction and gel electrophoresis analyses of mutant Δ*mtdM4* (Ju3017). (A) Construction of mutant Δ*mtdM*$_4$ and predicted PCR fragments size from wild-type and mutants; (B) Verification of the Δ*mtdM*$_4$ mutants by PCR. DNA templates were from *M. therm-otolerans* SCSIO 00652 (lane W) and Δ*mtdM*$_4$ (lane M1, M2, M3). Marker, DNA marker DL2000.

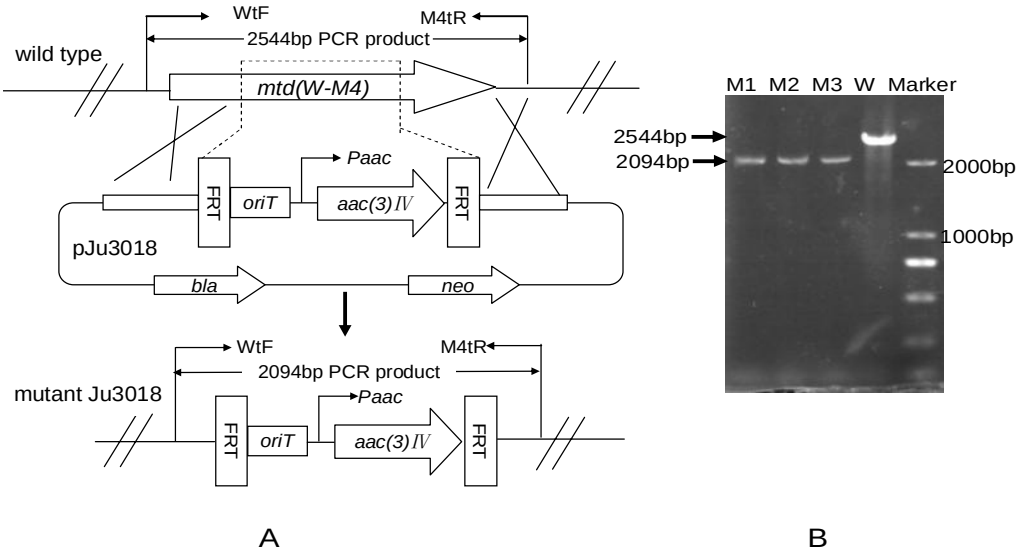


Fig. S16. Construction and gel electrophoresis analyses of mutant Δ*mtdWM*$_4$ (Ju3018). (A) Construction of mutant Δ*mtdWM*$_4$ and predicted PCR fragment size from wild-type and mutant; (B) Verification of the Δ*mtdWM*$_4$ mutant by PCR. DNA templates were from *M. thermotolerans* SCSIO 00652 (lane W) and Δ*mtdWM*$_4$ (lane M1, M2, M3). Marker, DNA marker DL2000.

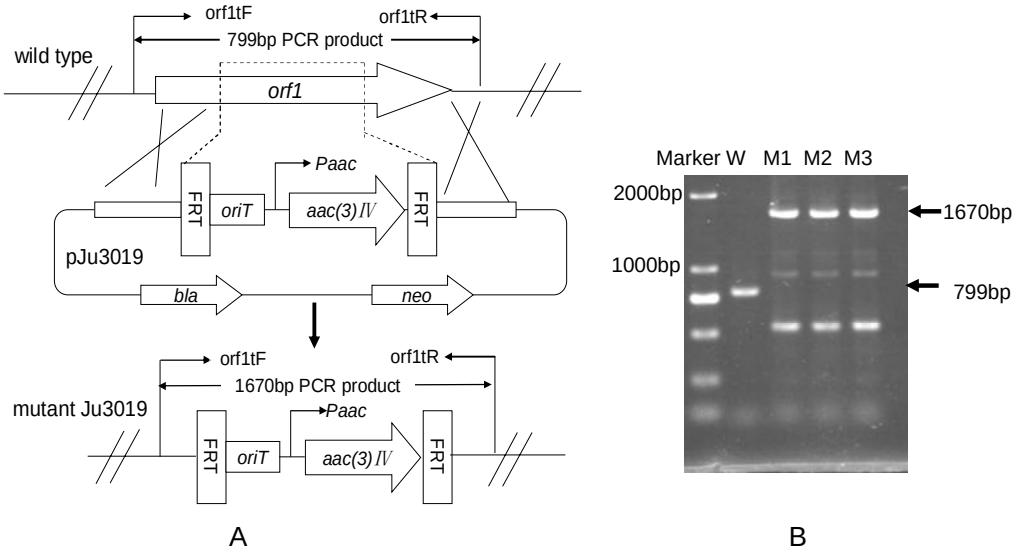


Fig. S17. Construction and gel electrophoresis analyses of mutant Δ*orf1* (Ju3019). (A) Construction of mutant Δ*orf1* and predicted PCR fragment size from wild-type and mutant; (B) Verification of the Δ*orf1* mutants by PCR. DNA templates were from *M. therm-otolerans* SCSIO 00652 (lane W) and Δ*orf1* (lane M1, M2, M3). Marker, DNA marker DL2000.

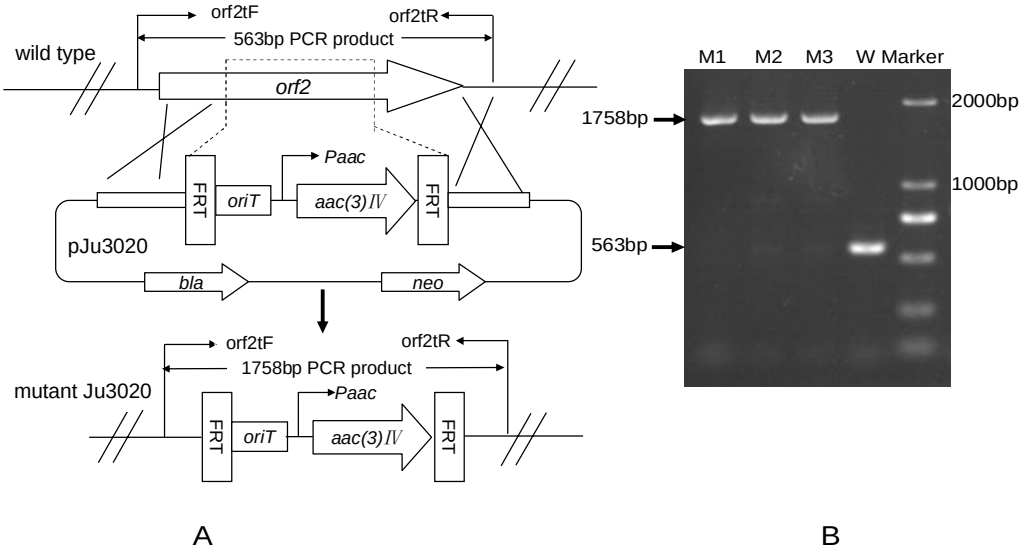


Fig. S18. Construction and gel electrophoresis analyses of mutant Δ*orf2* (Ju3020). (A) Construction of mutant Δ*orf2* and predicted PCR fragment size from wild-type and mutant; (B) Verification of the Δ*orf2* mutants by PCR. DNA templates were from *M. thermotolerans* SCSIO 00652 (lane W) and Δ*orf2* (lane M1, M2, M3). Marker, DNA marker DL2000.

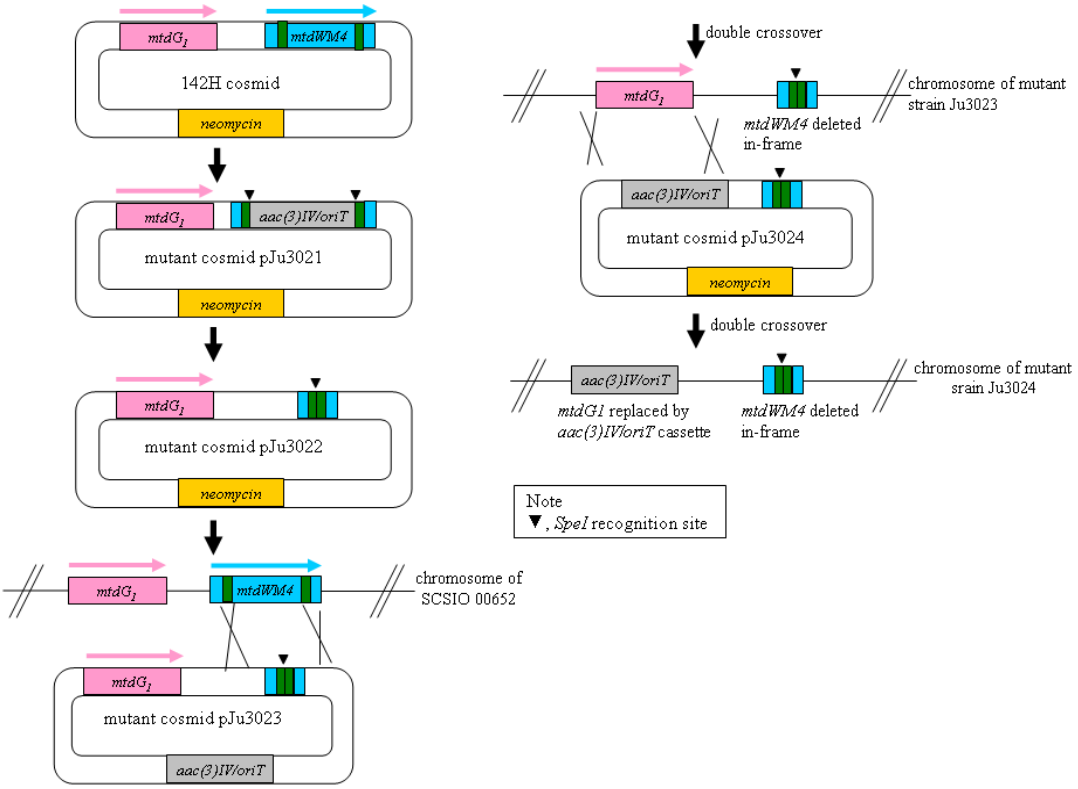


Fig. S19. Diagram showing construction strategy for mutant Δ*mtdG*$_1$*WM*$_4$.

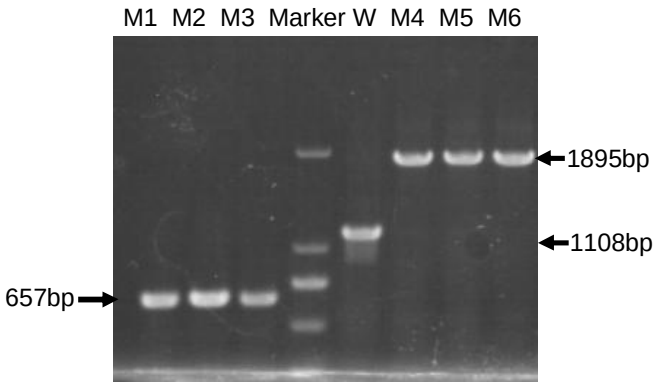


Fig. S20. PCR verification of the mutant Δ*mtdG*$_1$*WM*$_4$ (Ju3024). The whole construction process was shown in Fig. S19. On the one hand, the gene Δ*mtdWM*$_4$ was in-frame deleted, PCR using the primers WtF2 and M4tR was speculated to result in the amplification of a 657 bp fragment (lane M1, M2 and M3). On the other hand, the gene *mtdG*$_1$ was replaced by *aac(3)IV-oriT* fragment. Sizes of PCR products, using the primers G1tF and G1tR，was expected to be 1108 bp for the wild type strain *M. thermotolerans* SCSIO 00652(lane W) and 1895 bp for the mutants (lane M4, M5 and M6).

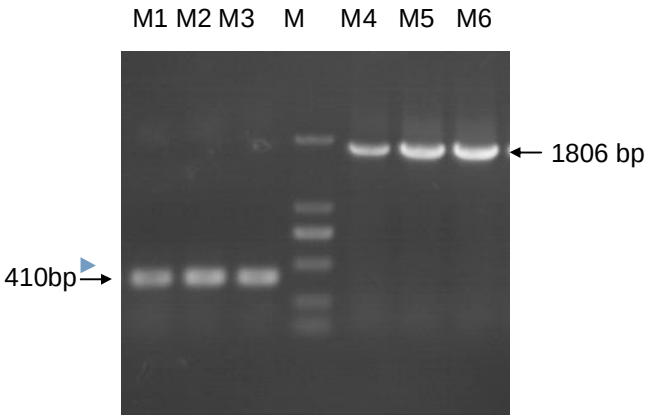


Fig. S21. PCR verification of the mutant Δ*mtdM₂M₃* (Ju3025). The whole construction process was the same as the mutants of Δ*mtdG₁WM₄* (Ju3024) shown in Fig. S19. On the one hand, the gene *mtdM₂* was in-frame deleted; PCR using the primers M2tF and M2tR was speculated to result in the amplification of a 410 bp fragment (lane M1, M2 and M3). On the other hand, the gene *mtdM3* was replaced by *aac(3)IV-oriT* fragment. Sizes of PCR products, using the primers M3tF and M3tR，was expected to be 1806 bp for the mutants (lane M4, M5 and M6).

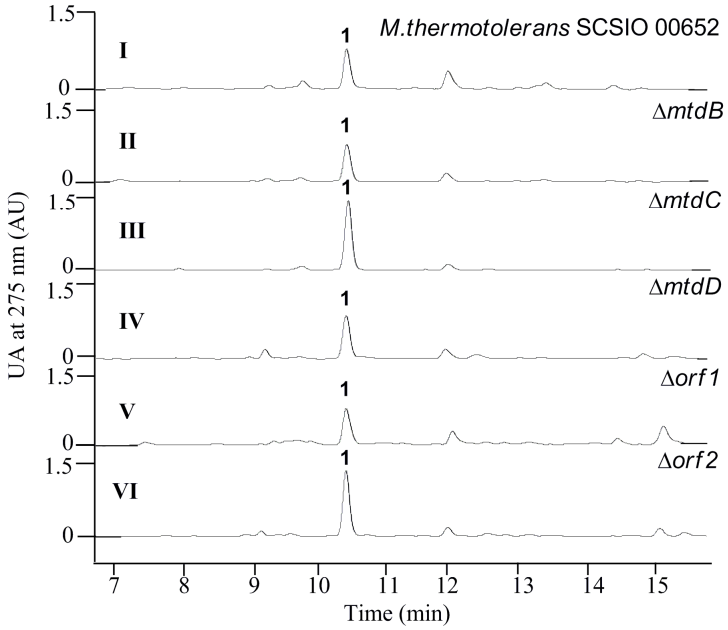


Fig. S22. Metabolite profiles of *M. thermotolerans* SCSIO 00652 mutant strains (II-VI), in comparison with the wild-type (I).

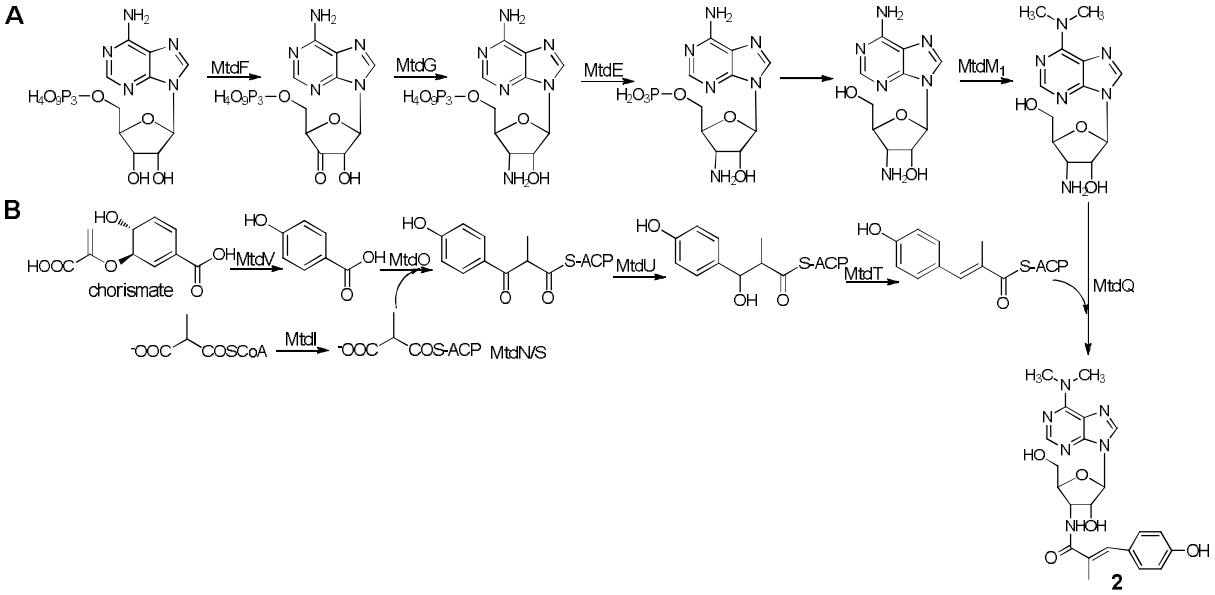


Fig. S23. The proposed biosynthetic pathway of *N*, *N*-dimethyl-3'-amino-3'-deoxyadenosine moiety and *p*-hydroxy-α-methylcinnamic acid moiety *en route* to **2**.

```
MtdL              --------------MSGRDIS-TAVVVTTISDG-GFLDRLAPALRDAGA----RLIVIP
Ata16             MPDHEPAIPADRADAADASDLSRTHIVLTTISDG-GFLDRFAPALRDSGA----RLTVIP
Hyg20             --------------MS----ETIDIVMTTIGSG-SFLDHFADALAEEGA----RLVVIP
UAM1              --MAGTVTVPSASVPSTPLLKDELDIVIPTIRN-LDFLEMWRPFFQP------YHLIIVQ
UAM3              --MAS--SDAAAAQAATPLLKDELDIVIPTIRN-LDFLEMWRPFFQP------YHLIIVQ
WP_030831156.1    --------------MS----ETIDIVMTTIGSG-SFLDHFADALAEEGA----RLVVIP
WP_029381493.1    --------------MS----ETIDIVMTTIGSG-TFLEHFTEALAEDGA----RLVVIP
WP_055557796.1    --------------MS----ETIDIVMTTIGSG-SFLDHYTDALAEDGA----RLVVIP
WP_018505382.1    --------------MSDSSGSTVDIVITTIGSG-EFLEQFTGTLTEDGA----RLVVIP
WP_004951028.1    --------------MS----ETIDIVITTIGSG-EFLEHYADTLAEGGA----RLVVIP
WP_007071477.1    --------------MSEP-LPRVDVVMTTIGDGEAFLSAYRTLLGAADAYRRVRIVVIP
                                           :*:.** .   **.     :         :: ::

MtdL              DRNTGPALFAACERHRRLGLDVVCPSVAEQQDLLERLAVPDLIPYHSDNRRNVGYLMAWM
Ata16             DRNTPPGLFAACERQRNRGLDVVCPTVEEQQELLDSLGAPELIPYASDNRRNVGYLLAWL
Hyg20             DRKTPSTFYEACERARRRGATIVAPDVEEQDRLLAKLGIPDLIPYDSDNRRNIGYLLSYL
UAM1              DGDPTKTIRVP------EGFDYELYNRNDINRILG--PKASCISFKDSACRCFGYMVSKK
UAM3              DGDPKKTIRVP------EGFDYELYNRDDINRILG--PRASCISFKDSACRCFGYMVSKK
WP_030831156.1    DRKTPSTFYDACERARRRGATIVAPDVEEQDRLLAKLGIPDLIPYDSDNRRNIGYLLSYL
WP_029381493.1    DRKTPRAFYDACERARGRGATVVFPDVEEQDRLLAKLGVPDLIPYDSDNRRNIGYLLSYL
WP_055557796.1    DRKTPSAFYDACERARARGATIVAPDVEEQDRLLAKLGIPDLIPCDSDNRRNIGYLLSYM
WP_018505382.1    DRKTPAALYQACDRARARGASIVAPDVEEQDRLLAKLGVPELIPYDSDNRRNIGYLLSYL
WP_004951028.1    DRRTPAAFHAACDRARARGAAILSPDVAEQDRLLAKLGVPELVPYDSDNRRNIGYLLSYL
WP_007071477.1    DRKTPAALFDAVERARRDGLHVVCPTVAEQDALLAGLGAPTLIPYDSDNRRNVGYLLSWQ
                  *  .   :  .       *         : : :*     .  :.  ..  * .**:::

                           D×D.....................................................R
MtdL              EGFDVIVSMDDDNLPTTDDFVERHQVVCQGPRTQPVTASSDGWFNNCALLEVEPTEVFPR
Ata16             QEAEVVISMDDDNLPRDPDFVRRHQVVRQGMRVQPVTTSATGWFNNCALLKTEPVDVFPR
Hyg20             NDSALAVSMDDDNLPVDRPFLDEHRIVLQGPARHRVVSAGNGWFNACDLLTVSPCRVFPR
UAM1              ---KYVFTIDDDCFVAKD--PSGKDINALEQHIKNLLSPSTPFFFNTLYDPYREGADFVR
UAM3              ---KYIYTIDDDCFVAKD--PSGKDINALEQHIKNLLNPSTPFFFNTLYDPYRDGADFVR
WP_030831156.1    NDSALAVSMDDDNLPADRPFLDEHRVVLQGPARHRVVSADNGWFNACDLLTVSPCRVFPR
WP_029381493.1    NGSALAVSMDDDNLPVDHPFLDEHRVVLRGPARHRTVSAANGWFNACDLLTVTPCRVFPR
WP_055557796.1    NDSALAVSMDDDNLPVASPFLDEHRIVLRGPTRHRVVSVGNGWFNACDLLTVSPCRVFPR
WP_018505382.1    NDSAFAVSMDDDNLPVDSPFFDEHRIVAAGPTEQRVVSNSDGWFNACDLLTVEPCRVFPR
WP_004951028.1    NGSACAVSMDDDNLPAVSPFLDEHRVVLEGPARHRTVSSPSGWFNCCDLLDVTPCRVHPR
WP_007071477.1    SDADFLISVDDDNFPIDGDFLTAHAVVAAGPRPARVVTAESGWWNPCGQLTVAPMPVYPR
                        ::*** :         : :               ::            . *
```

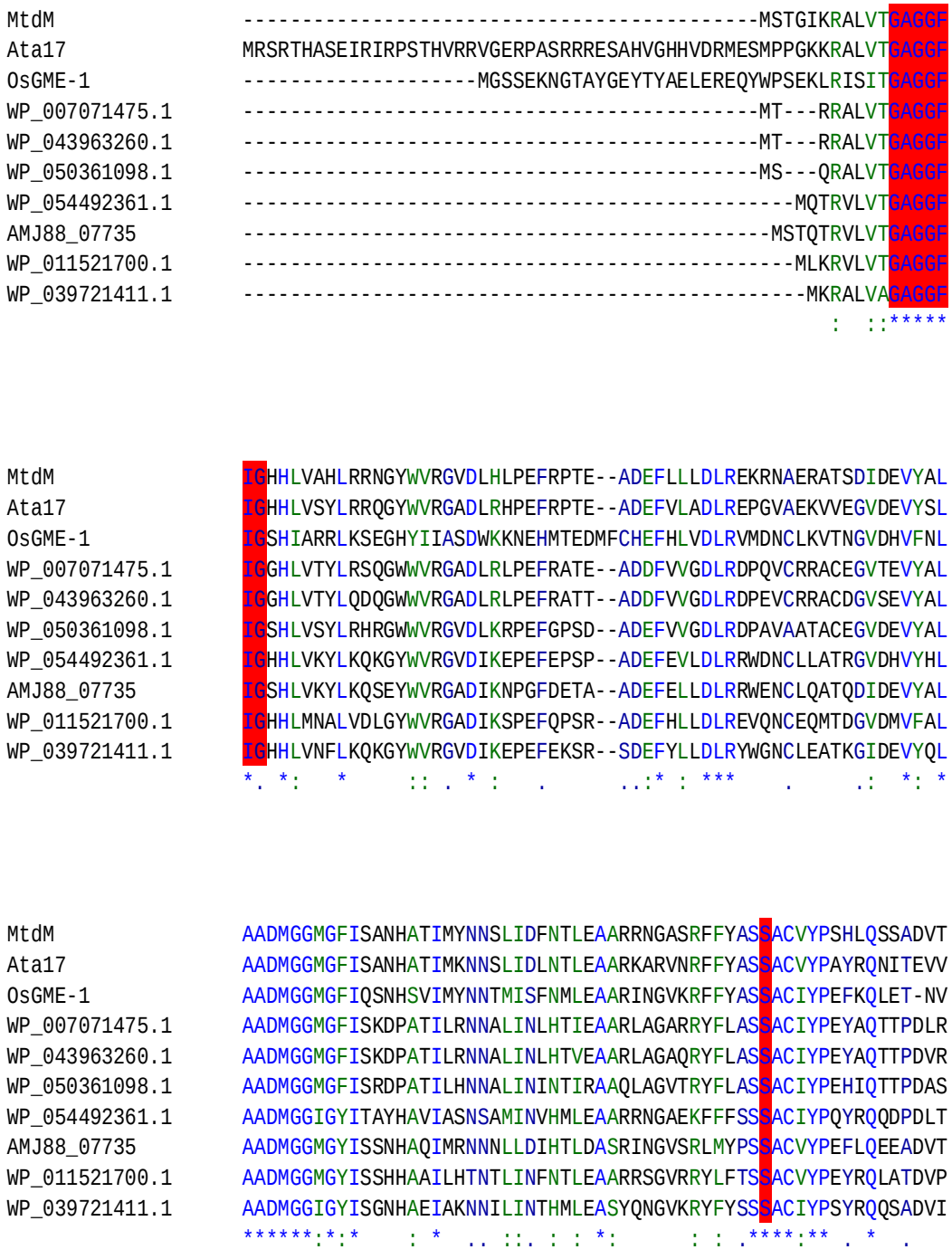
```
MtdL              GFPFHARPAHAQARTSVCERPADVRINAGLWLGDPDVDAITRLAVR-PNALAHSGGSVVL
Ata16             GFPLRHRATYDETALTTTRQPADVRVNAGLWLGDPDVDAITRVAVR-PEVTAHAGGNAVL
Hyg20             GFPYAPRT--AGTEVTSTEETADVRVNAGLWLDDPDVDAITRLAVR-PRVTAYGGEAAVL
UAM1              GYPFSLR-----------EGAKTAVSHGLWLNIPDYDAPTQMVKPRERNSRYVDAVMTV
UAM3              GYPFSLR-----------EGAPTAVSHGLWLNIPDYDAPTQLVKPLERNSRYVDAVMTI
WP_030831156.1    GFPYAPRT--AGTEVTSTEETADVRVNAGLWLDDPDVDAITRLAVR-PRVTAYGGEAAVL
WP_029381493.1    GFPYGPRQ--EHTEVTRTEEVADVRVNAGLWLDDPDVDAVTRLAVR-PQVTAYSGEAVVM
WP_055557796.1    GFPYAPRE--AGTEVTSTEETVDVRVNAGLWLDDPDVDAVTRLAVR-PRVTAYGGEAAVL
WP_018505382.1    GFPYGPRE--ASTELTETTRSVDVRVNAGLWLDDPDVDAVTRLAVR-PRVTGYERGTAVL
WP_004951028.1    GFPYGPRT--DPAAPTWTEETADVRVNAGLWLGDPDVDAVTRLAVR-PTVTAYRGPAAVL
WP_007071477.1    GFPYAHRS---PTPTSERTETVDVRINAGLWLGDPDVDAITRIAVR-PEVTAMPAPALVC
                  *:*   *            . . . :. ****. ** ** *::.             .

MtdL              AEGTWCPVNSQNTAVHRDAL-PAYYFLRMGQPVDGVPMERFGDIFSGYFVQVCAQHLGHA
Ata16             GRGTWCPVNSQNTALHRDAL-PAYYFLRMGQRVGGGVMERFGDIFSGYFLQACAQHLGHA
Hyg20             APDTWCSVNSQNTAVHHDAL-PAYYFLRMGQSIGGAPVERFGDIFSGYFVAACAKHLGHA
UAM1              PKGTLFPMCGMNLAFDRDLIGPAMYFGLMGD---GQPIGRYDDMWAGWCMKVICDHLSLG
UAM3              PKGTLFPMCGMNLAFDRDLIGPAMYFGLMGD---GQPIGRYDDMWAGWCTKVITDHLGLG
WP_030831156.1    APDTWCSVNSQNTAVHHDAL-PAYYFLRMGQSIGGAPVERFGDIFSGYFVAACAKHLGHA
WP_029381493.1    ARDTWCPVNSQNTAVHRDAL-PAYYFLRMGQPIGGAAVERFGDIFSGYFVAACAKHLGHA
WP_055557796.1    AGDTWCPVNSQNTAVHHDAL-PAYYFLRMGQSIGGATVERFGDIFSGYFVAACAKHLGHA
WP_018505382.1    AADTWCPVNSQNTAVHHDAL-AAYYFLRMGQPIYGAPVERFGDIFSGYFVAACAKHLGHA
WP_004951028.1    ARDTWCPVNSQNTAVHRDAL-PAYYFLRMGQPVGGAPLERFGDIFSGYFLAACTKHLGHS
WP_007071477.1    DTGTWAPVNSQNTAVHRDAI-PAYYFPRMGYRHHGQEIDRYADIFSGYFVQACAKRLGHA
                    .*  .: . * *..:* : .* **  **    *  : *: *:::*:   .  .:*. .

MtdL              VRFGDPVVEHPRNEHDLLDDLHKEVPAVRLLDDILDHLRDHPL--EGGDYLETYESLSYA
Ata16             VRFGDPLVDHPRNEHDLLDDLTKELPAVRLLDDLLDWLRECPL--EGGDYFTAYEALSHG
Hyg20             VRFGGPLVNHQRNDHDLLDDLAIELPAIRFMDELLDWLREFPI--EGSDYRESYESLSYG
UAM1              VKTGLPYIWHSKASNPFVN-LKKEYKGIFWQEDIIPFFQNATIPKECDTVQKCYLSLAEQ
UAM3              VKTGLPYIWHSKASNPFVN-LKKEYNGIFWQEELIPFFQSASLPKEADTVQKCYLELAKQ
WP_030831156.1    VRFGGPLVNHQRNDHDLLDDLAIELPAIRFMDELLEWLREFPI--EGSDYRESYESLSYG
WP_029381493.1    VRFGGPLVNHERNEHDLLHDLVIELPAIRFMDELLAWLQEFPL--EGGDYQEAYDSLSHG
WP_055557796.1    VRFGGPLVNHERNDHDLLNDLAIELPAIRFMDELLDWLREFSI--EGSDYRESYESLSHG
WP_018505382.1    VRFGGPLVNHERNAHDLLNDLAIELPAIRFFDEFLDWLVAHEL--SGSTYAETYESLSYG
WP_004951028.1    VRFGGPLVHHERNAHDLFADLTAELPAIRFMDELLDWLREFRP--DGSDYREAYASLAHG
WP_007071477.1    VRFGDPLARHTRNEHLLLRDLQQELTAIAILEDVLDWLHGCKL--DGDTYAEAYVSLSYQ
                  *: * *   * :  : :.  *  *  .:   ::.:  :       . .    *  *:
```

```
MtdL                LQEIAERVNGRA-WSPDARAFLHRSAHLMRSWTGALRTVAGT-----
Ata16               LQDFAEEASGPA-WTADARAFLHRSAHLMRTWLGVLRVGAGA-----
Hyg20               LQDFAEQARGRG-WTPDARAFLHRSAHLMRTWLTAIRRVHGG-----
UAM1                VREKLGKIDPYFVKLADAMVTWIEAWDELNPSTAAVE--NGKAK---
UAM3                VRAKLGKVDGYFNKLADSMVTWIEAWDQLNPPKGAVATANGTAKSK-
WP_030831156.1      LQDFAEQARGRG-WTPDARAFLHRSAHLMRTWLTAIRRVHGG-----
WP_029381493.1      LQDFAEQAGGRA-WTPEARAFLHRSAHHMRTWLSAVRRIDGR-----
WP_055557796.1      LQDFAEQARGRG-WTPDARAFLHRSAHLMRTWLTAIRRVNGS-----
WP_018505382.1      LQDFAEAAQGKG-YTPEARAFLHRSAHHMRTYLRAIRAIDGS-----
WP_004951028.1      LREFAEQARGPA-WTQDARAFLHRSAHLMLTWLSAVRRIDGG-----
WP_007071477.1      LQDAVEAMSGRV-WTHELRGFVHQMAHLMRQWVGVLQRCHGAGGTPA
                    ::              :      .  . :     .:    *
```

Fig. S24. Multiple sequence alignments of MtdL with its homologues (Ata16, Hyg20, UAM1, UAM3, WP_030831156.1, WP_029381493.1, WP_055557796.1, WP_018505382.1, WP_004951028.1, WP_007071477.1). The following proteins (with GenBank IDs) were used for amino acid alignment: Ata16 (CAD62189.1) from *Saccharothrix mutabilis* subsp. capreolus, Hyg20 (ABC42557.1) and WP_030831156.1 from *Streptomyces hygroscopicus*, UAM1 (Q8H8T0.1) and UAM3 (Q6Z4G3.1) from *Oryza sativa* Japonica Group, WP_029381493.1 from *Streptomyces leeuwenhoekii*, WP_055557796.1 from *Streptomyces* sp. NBRC 110028, WP_018505382.1 from *Frankia* sp. BCU110501, WP_004951028.1 from *Streptomyces mobaraensis*, WP_007071477.1 from *Micromonospora* sp. ATCC 39149. The DxD……R motif sequence was highlighted in red color.


MtdM ----------MSTGIKRALVTGAGGF
Ata17 MRSRTHASEIRIRPSTHVRRVGERPASRRRESAHVGHHVDRMESMPPGKKRALVTGAGGF
OsGME-1 ----------MGSSEKNGTAYGEYTYAELEREQYWPSEKLRISITGAGGF
WP_007071475.1 ----------MT---RRALVTGAGGF
WP_043963260.1 ----------MT---RRALVTGAGGF
WP_050361098.1 ----------MS---QRALVTGAGGF
WP_054492361.1 ----------MQTRVLVTGAGGF
AMJ88_07735 ----------MSTQTRVLVTGAGGF
WP_011521700.1 ----------MLKRVLVTGAGGF
WP_039721411.1 ----------MKRALVAGAGGF
IGHHLVAHLRRNGYWVRGVDLHLPEFRPTE--ADEFLLLDLREKRNAERATSDIDEVYAL
IGHHLVSYLRRQGYWVRGADLRHPEFRPTE--ADEFVLADLREPGVAEKVVEGVDEVYSL
IGSHIARRLKSEGHYIIASDWKKNEHMTEDMFCHEFHLVDLRVMDNCLKVTNGVDHVFNL
IGGHLVTYLRSQGWWVRGADLRLPEFRATE--ADDFVVGDLRDPQVCRRACEGVTEVYAL
IGGHLVTYLQDQGWWVRGADLRLPEFRATT--ADDFVVGDLRDPEVCRRACDGVSEVYAL
IGSHLVSYLRHRGWWVRGVDLKRPEFGPSD--ADEFVVGDLRDPAVAATACEGVDEVYAL
IGHHLVKYLKQKGYWVRGVDIKEPEFEPSP--ADEFEVLDLRRWDNCLLATRGVDHVYHL
IGSHLVKYLKQSEYWVRGADIKNPGFDETA--ADEFELLDLRRWENCLQATQDIDEVYAL
IGHHLMNALVDLGYWVRGADIKSPEFQPSR--ADEFHLLDLREVQNCEQMTDGVDMVFAL
IGHHLVNFLKQKGYWVRGVDIKEPEFEKSR--SDEFYLLDLRYWGNCLEATKGIDEVYQL
AADMGGMGFISANHATIMYNNSLIDFNTLEAARRNGASRFFYASSACVYPSHLQSSADVT
AADMGGMGFISANHATIMKNNSLIDLNTLEAARKARVNRFFYASSACVYPAYRQNITEVV
AADMGGMGFIQSNHSVIMYNNTMISFNMLEAARINGVKRFFYASSACIYPEFKQLET-NV
AADMGGMGFISKDPATILRNNALINLHTIEAARLAGARRYFLASSACIYPEYAQTTPDLR
AADMGGMGFISKDPATILRNNALINLHTVEAARLAGAQRYFLASSACIYPEYAQTTPDVR
AADMGGMGFISRDPATILHNNALININTIRAAQLAGVTRYFLASSACIYPEHIQTTPDAS
AADMGGIGYITAYHAVIASNSAMINVHMLEAARRNGAEKFFFSSSACIYPQYRQQDPDLT
AADMGGMGYISSNHAQIMRNNNLLDIHTLDASRINGVSRLMYPSSACVYPEFLQEEADVT
AADMGGMGYISSHHAAILHTNTLINFNTLEAARRSGVRRYLFTSSACVYPEYRQLATDVP
AADMGGIGYISGNHAEIAKNNILINTHMLEASYQNGVKRYFYSSSACIYPSYRQQSADVI

```
MtdM              GLREEVSHPADPEDGYGWEKLHIEHACAYYREEFGLETRVARLHNVYGPYSTYAGGREKA
Ata17             GLREEDAYPAAPEDGYGWEKLNTEHLCSYYREEYGLPVRVARLHNVYGPYCTYDGGREKS
OsGME-1           SLKESDAWPAEPQDAYGLEKLATEELCKHYTKDFGIECRVGRFHNIYGPFGTWKGGREKA
WP_007071475.1    PLREDDAFPAGPQDSYGWEKLMAERLCVYYAEQYGLAVRIARYHNVYGPYGTYDGGREKA
WP_043963260.1    PLREDDAFPAGPQDSYGWEKLMAERLCVYYAEQYGLDVRIARYHNVYGPYGTWDGGREKA
WP_050361098.1    ALRETDAFPADPQDSYGWEKLMAELLCRYYHEQYGMEVRVARYHNIYGPNGTYDGGREKA
WP_054492361.1    PLREEDAYPADPEEGYGWEKLFMEKLCQYYYEDYGFKTYVARFHNIFGPLGTWEGGKEKA
AMJ88_07735       PLKEEDAYPAMPQDAYGWQKLISERTCMHFQEEFDIETRVVRFHNIFGPFGTWDGGREKA
WP_011521700.1    ALREEDAYPAAPQDAYGWEKLITERLCTHYREDYGMEMRIIRFHNIFGPLGTWEGGREKA
WP_039721411.1    PLKEEDAMPADPEEGYGWEKLFAEKLCQYYQEDKGMETRVARFHNVYGPLGTYKGGREKA
                   *:*  : ** *::.** :**  *  * :: :: .:   : * **::**  *: **:**:

MtdM              PAALARKAALAAPGG--EMEIWGDGRQTRSFCYVDDCVEGIRRLTASDFPGPVNIGTEHL
Ata17             PAALARKAALAEPGG--RMEIWGDGMQTRSYCYVDDCVEGIHRLTRSDFPGPVNLGTERL
OsGME-1           PAAFCRKAQTSTDR----FEMWGDGLQTRSFTFIDECVEGVLRLTKSDFREPVNIGSDEM
WP_007071475.1    PAALCRKVAEAPPGG--EVEIWGDGRQTRSFCYVDDCLEGTYRLMRSDHGEPVNIGSDRL
WP_043963260.1    PAALCRKVAEAPPGG--EVEIWGDGRQTRSFCYVDDCVEGTYRLMRSDHAEPVNIGSDRL
WP_050361098.1    PAAMCRKVALASEDG--SIEVWGDGSQTRSFCYVDDCVEGTYLLMRSDHREPLNIGSDRL
WP_054492361.1    PAALCRKIAITKLTGNPEVEIWGDGEQTRSFCYVDDLVEGVYRLMQSDYHQPLNIGQDRM
AMJ88_07735       PAALCRKVAVAKLTGNPEVEIWGDGEQTRSFCYIDDCLVGLHKLMRSDYHQPLNLGQDRM
WP_011521700.1    PAAMCRKVAIAKLTGNHEIEIWGDGKQTRSFCYIDDCVTGIHKLMVSDFAYPLNLGQDRM
WP_039721411.1    PAAICRKIALAEDSS--EIEIWGDGKQTRSFLYIQDCVEGIYLITQSDYPKPLNLGSEEL
                  ***:.**   :       .*:**** ****: :::: : *   :  **.  *:*:* :.:

MtdM              IAIDDLARMLLSIAGKEDVRLVH-RPGPQGVRGRNSDNTLLREKLRWEPATPLWEGMSAM
Ata17             IAINDLARMLLEIAGKPGVTLEH-RPGPQGVRGRNSDNALLRAELGWEPSTPLETGMAAT
OsGME-1           VSMNEMAEIILSFEDR-ELPIHH-IPGPEGVRGRNSDNTLIKEKLGWAPTMKLKDGLRFT
WP_007071475.1    VTIDELAALVMAAAGRDDLRLRH-VSGPQGVRGRNSDNTRVRQVLGWAPGIPLEQGLAVT
WP_043963260.1    VTIDELAALVMTAAGRDDLRLRH-VAGPQGVRGRNSDNTLVRQVLGWAPRIPLEEGLAVT
WP_050361098.1    VSVDELAHLVFAAAGRDGLGIEH-IEGPQGVRGRNSDNTLLRKVLGWEPTVPLEKGLLET
WP_054492361.1    VTINELADIIANIAG-IEIVKKH-VPGPQGVRGRNSDNTRLREVLGWEPQISLEEGLART
AMJ88_07735       VTINQLVEMIAEIAG-IQIKRKH-IPGPQGVRGRNSDNTRLKEVLNWEPTISLEEGLART
WP_011521700.1    VSINELADLVADIAG-IRVNKRH-VSGPMGVRGRNSDNTLLRQVLGWTPVISLEDGLRRT
WP_039721411.1    VTIDQLVEMTAKVANKN-IRIRHNLSKPQGVRGRNSDNSKLYKITGWMPKFSLLEGLKLT
                  ::::::. :         :   *    * *********  :     * *   *  *:
```

```
MtdM                YHWIERDIARRGRDVAVSSIVPRAEEPHVRP--------------
Ata17               YHWIRSDIERRAGTVQAASEIVRVGDPGA----------------
OsGME-1             YFWIKEQIEKEKTQGVDIAGYGSSKVVSTQAPVQLGSLRAADGKE
WP_007071475.1      YRWIAEQVAARRDATAQAAPA------------------------
WP_043963260.1      YRWIAGEVAAARGAAAHATPA------------------------
WP_050361098.1      YQWISRELAARSVA-------------------------------
WP_054492361.1      YAWIEEQVRQKLAREMDISSASV----------------------
AMJ88_07735         YAWIEKQVQAKLDREHE----------------------------
WP_011521700.1      YRWIEAQVAAKLSEKCSSSFTSKVAATTP----------------
WP_039721411.1      YPWIAERVAQERNMQGCQ---------------------------
                    * **   :
```

Fig. S25. Multiple sequence alignments of MtdM with its homologues (Ata17, OsGME-1, WP_007071475.1, WP_043963260.1, WP_050361098.1, WP_054492361.1, AMJ88_07735, WP_011521700.1, WP_039721411.1). The following proteins (with GenBank IDs) were used for amino acid alignment: Ata17 (CAD62190.1) from *Saccharothrix mutabilis* subsp. capreolus, OsGME-1 (BAD66930.1) from *Oryza sativa* Japonica Group, WP_007071475.1 from *Micromonospora* sp. ATCC 39149, WP_043963260.1 from *Micromonospora carbonacea*, WP_050361098.1 from *Streptomyces*, WP_054492361.1 from *Ardenticatena maritime*, AMJ88_07735 from *Anaerolineae bacterium* SM23_63, WP_011521700.1from *Candidatus Koribacter versatilis*, WP_039721411.1 from *Methylacidiphilum kamchatkense*. The GxxGxxG……S……YxxxK motif sequence was highlighted in red color.

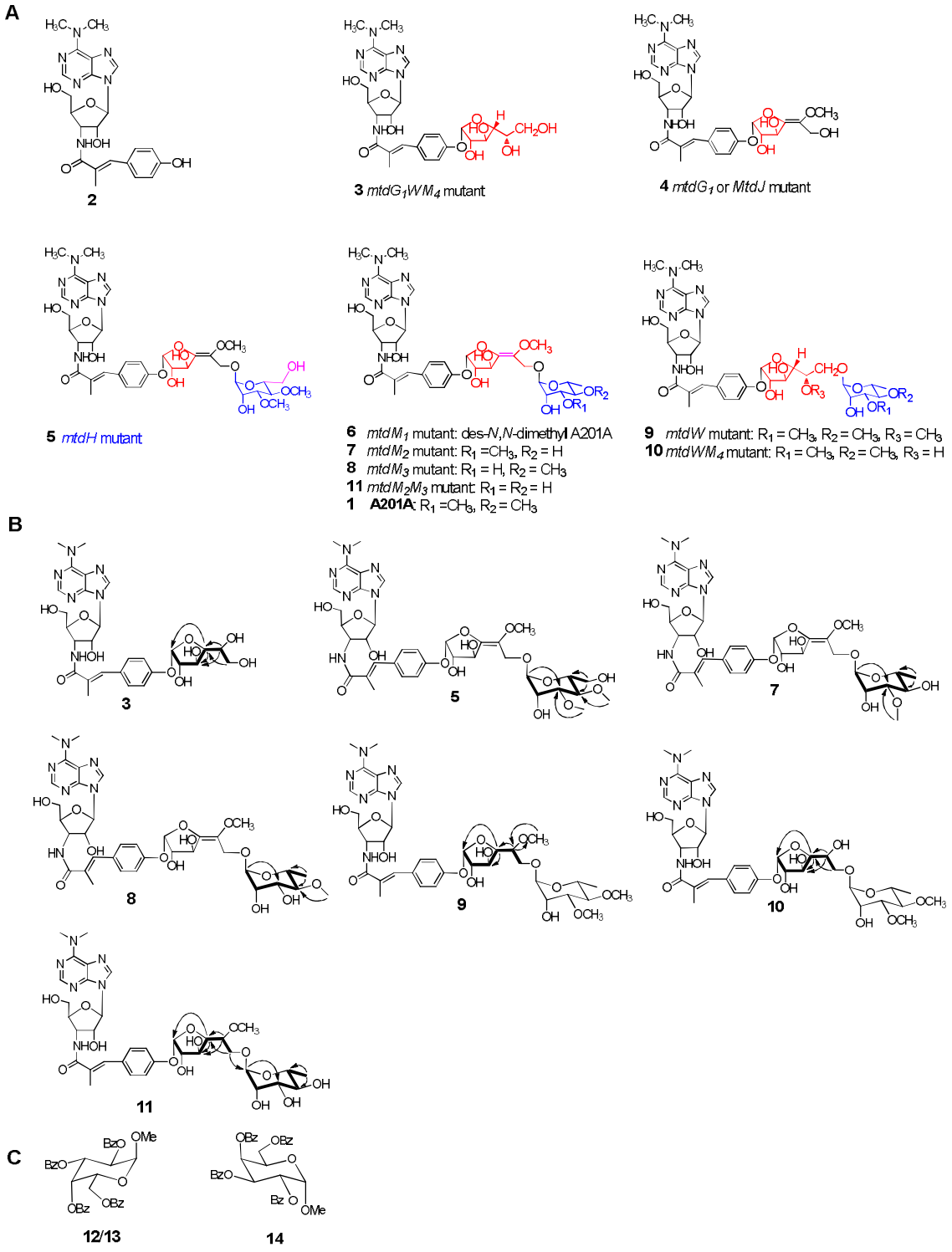


Fig. S26. (A) Compounds **1**–**11** as identified during metabolite analyses carried out during this study; (B) Key COSY and HMBC correlations of **3**, **5**, and **7**–**11**; (C) Chemical structures of compounds **12**–**14**.

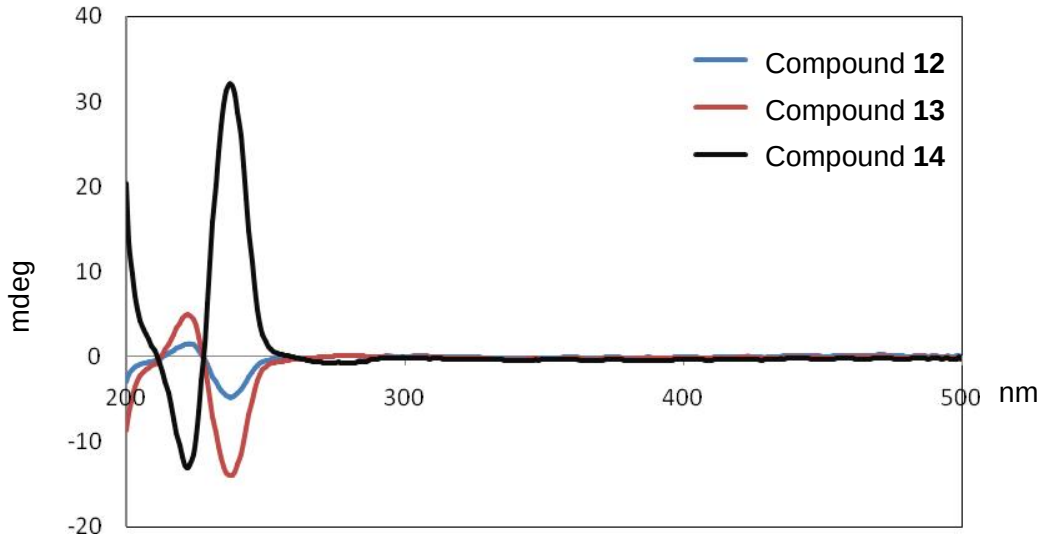
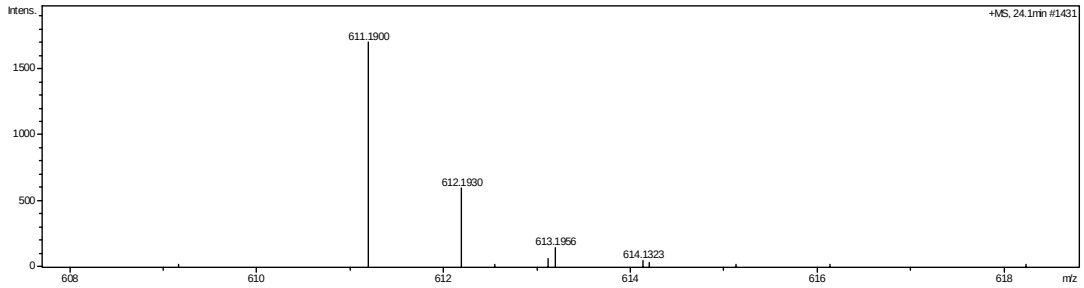
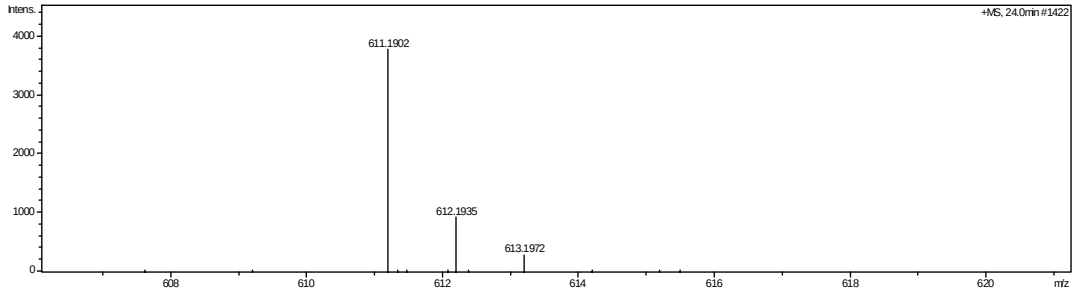


Fig. S27. The CD spectra of compounds **12**, **13**, and **14** showed that **12** and **13** possessed the same L-configuration. CD spectra were obtained at 25 °C using a 1 mm quartz cuvette in MeCN. The concentrations of **12**, **13**, and **14** were 0.000013 and 0.000049, and 0.000115 M, respectively.

A. HR-ESI-MS spectrum for compound **12**.



B. HR-ESI-MS spectrum for compound **13**.



C. HR-ESI-MS spectrum for compound **14**.

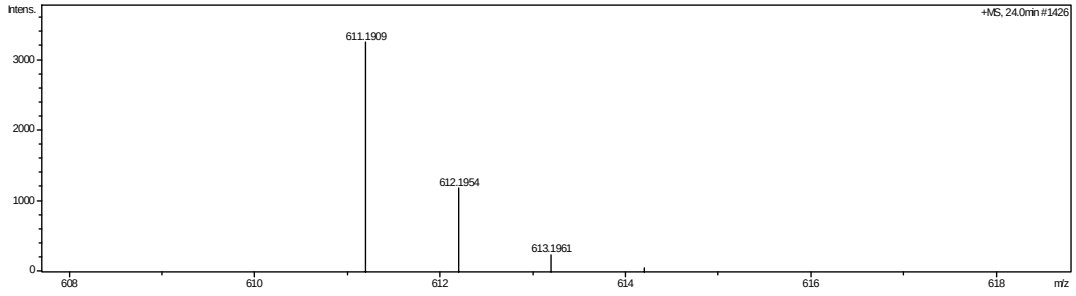
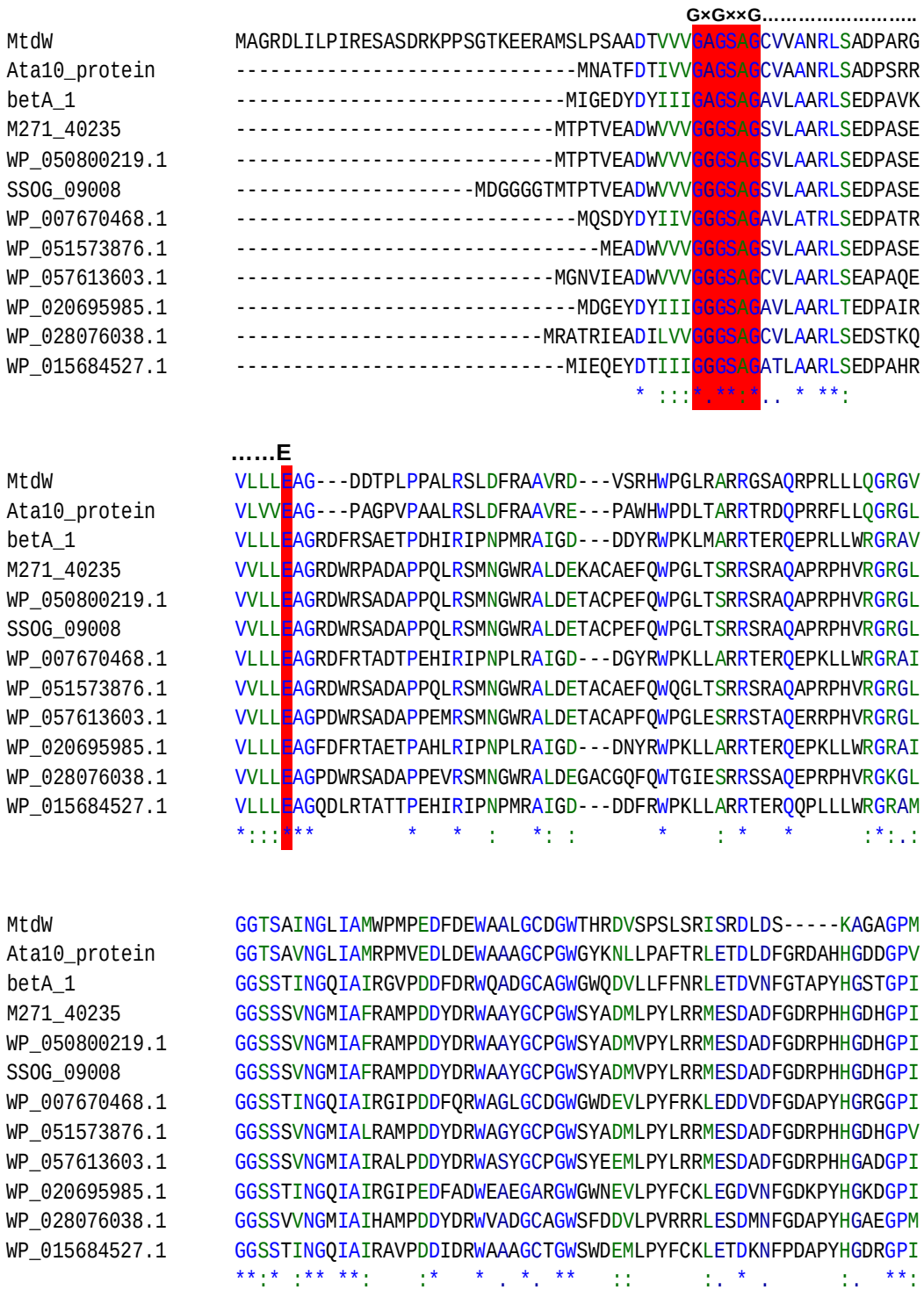


Fig. S28. The HR-ESI-MS spectra for compounds **12**–**14**.


G×G××G.....
MtdW MAGRDLILPIRESASDRKPPSGTKEERAMSLPSAADTVVVGAGSAGCVVANRLSADPARG
Ata10_protein ----------MNATFDTIVVGAGSAGCVAANRLSADPSRR
betA_1 ----------MIGEDYDYIIIGAGSAGAVLAARLSEDPAVK
M271_40235 ----------MTPTVEADWVVVGGGSAGSVLAARLSEDPASE
WP_050800219.1 ----------MTPTVEADWVVVGGGSAGSVLAARLSEDPASE
SSOG_09008 ----------MDGGGGTMTPTVEADWVVVGGGSAGSVLAARLSEDPASE
WP_007670468.1 ----------MQSDYDYIIVGGGSAGAVLATRLSEDPATR
WP_051573876.1 ----------MEADWVVVGGGSAGSVLAARLSEDPASE
WP_057613603.1 ----------MGNVIEADWVVVGGGSAGCVLAARLSEAPAQE
WP_020695985.1 ----------MDGEYDYIIIGGGSAGAVLAARLTEDPAIR
WP_028076038.1 ----------MRATRIEADILVVGGGSAGCVLAARLSEDSTKQ
WP_015684527.1 ----------MIEQEYDTIIIGGGSAGATLAARLSEDPAHR
.....E
MtdW VLLLEAG---DDTPLPPALRSLDFRAAVRD---VSRHWPGLRARRGSAQRPRLLLQGRGV
Ata10_protein VLVVEAG---PAGPVPAALRSLDFRAAVRE---PAWHWPDLTARRTRDQPRRFLLQGRGL
betA_1 VLLLEAGRDFRSAETPDHIRIPNPMRAIGD---DDYRWPKLMARRTERQEPRLLWRGRAV
M271_40235 VVLLEAGRDWRPADAPPQLRSMNGWRALDEKACAEFQWPGLTSRRSRAQAPRPHVRGRGL
WP_050800219.1 VVLLEAGRDWRSADAPPQLRSMNGWRALDETACPEFQWPGLTSRRSRAQAPRPHVRGRGL
SSOG_09008 VVLLEAGRDWRSADAPPQLRSMNGWRALDETACPEFQWPGLTSRRSRAQAPRPHVRGRGL
WP_007670468.1 VLLLEAGRDFRTADTPEHIRIPNPLRAIGD---DGYRWPKLLARRTERQEPKLLWRGRAI
WP_051573876.1 VVLLEAGRDWRSADAPPQLRSMNGWRALDETACAEFQWQGLTSRRSRAQAPRPHVRGRGL
WP_057613603.1 VVLLEAGPDWRSADAPPEMRSMNGWRALDETACAPFQWPGLESRRSTAQERRPHVRGRGL
WP_020695985.1 VLLLEAGFDFRTAETPAHLRIPNPLRAIGD---DNYRWPKLLARRTERQEPKLLWRGRAI
WP_028076038.1 VVLLEAGPDWRSADAPPEVRSMNGWRALDEGACGQFQWTGIESRRSSAQEPRPHVRGKGL
WP_015684527.1 VLLLEAGQDLRTATTPEHIRIPNPMRAIGD---DDFRWPKLLARRTERQQPLLLWRGRAM
GGTSAINGLIAMWPMPEDFDEWAALGCDGWTHRDVSPSLSRISRDLDS-----KAGAGPM
Ata10_protein GGTSAVNGLIAMRPMVEDLDEWAAAGCPGWGYKNLLPAFTRLETDLDFGRDAHHGDDGPV
betA_1 GGSSTINGQIAIRGVPDDFDRWQADGCAGWGWQDVLLFFNRLETDVNFGTAPYHGSTGPI
M271_40235 GGSSSVNGMIAFRAMPDDYDRWAAYGCPGWSYADMLPYLRRMESDADFGDRPHHGDHGPI
WP_050800219.1 GGSSSVNGMIAFRAMPDDYDRWAAYGCPGWSYADMVPYLRRMESDADFGDRPHHGDHGPI
SSOG_09008 GGSSSVNGMIAFRAMPDDYDRWAAYGCPGWSYADMVPYLRRMESDADFGDRPHHGDHGPI
WP_007670468.1 GGSSTINGQIAIRGIPDDFQRWAGLGCDGWGWDEVLPYFRKLEDDVDFGDAPYHGRGGPI
WP_051573876.1 GGSSSVNGMIALRAMPDDYDRWAGYGCPGWSYADMLPYLRRMESDADFGDRPHHGDHGPV
WP_057613603.1 GGSSSVNGMIAIRALPDDYDRWASYGCPGWSYEEMLPYLRRMESDADFGDRPHHGADGPI
WP_020695985.1 GGSSTINGQIAIRGIPEDFADWEAEGARGWGWNEVLPYFCKLEGDVNFGDKPYHGKDGPI
WP_028076038.1 GGSSVVNGMIAIHAMPDDYDRWVADGCAGWSFDDVLPVRRRLESDMNFGDAPYHGAEGPM
WP_015684527.1 GGSSTINGQIAIRAVPDDIDRWAAAGCTGWSWDEMLPYFCKLETDKNFPDAPYHGDRGPI

```
MtdW                PVRRPPRPTWGSLDLALAEWAGDRGFPAAEDHNAPGSTGLSPYAFSARDGGRVSAADAFL
Ata10_protein       PVRRTRPAAWGALDLALASWAGDRGLPRVEDHNAPDTTGLAPYAFNAWSDTRVSAADAFL
betA_1              PVYRAPVADWGSVDVALKDASIGLGYGWCEDHNAPEGTGASPYAINSRAGARVSTNDGYL
M271_40235          PVQRLGREDWGPADHALAEAALGRGHAWCDDHNGPAGTGVSPYAISSRDGARVTTNDGYL
WP_050800219.1      PVQRLGREDWGPADHALAEAALGRGHAWCDDHNGPTGTGVSPYAISSRDGARVTTNDGYL
SSOG_09008          PVQRLGREDWGPADHALAEAALGRGHAWCDDHNGPTGTGVSPYAISSRDGARVTTNDGYL
WP_007670468.1      PVYRAPVEDWGYVDRALKASAESLGYPWCEDHNAPDSTGASPYAINSRDGYRISTNDGYL
WP_051573876.1      PVQRLGREDWGPADHALAEAALGHGHAWCDDHNGPTGTGVSPYAISSRDGARVTTNDGYL
WP_057613603.1      PVQRLGREAWGPADHALAEAALGIGHPWCDDHNGPTGTGVSPYGINSRDGARVTANDGYL
WP_020695985.1      PVYRAPIDDWGPADRALMKAAVGLGYGWCEDHNAPEGTGASPYAINSRAGLRISTNDGYL
WP_028076038.1      PITRLDRAAWGPVDEALAAGAGAAGHGWCEDHNAPTGTGISPYGISSRDGARVTANDGYL
WP_015684527.1      PVYRAPIPDWGNVDRALRSSALSLGYGWCEDHNAPDGTGVSPYAINSEGGLRISTNDGYL
                    *: *     **  * **   :   *     * *.*  ** :** :.:    *::  *.:*

MtdW                APILDRPNLTVLARTVAERLIVRGGR--VTGVCCRTPDGEVVIEA-GEVVVTAGAVGSPA
Ata10_protein       APVLDRPNLTVLTGTVCRRLVVRGGR--VTGVECDGPTG--VVTA-AEVVVAAGVLGSPA
betA_1              EPARGRANLTIVGHALVDRLQVEGNRPHASGVHVTVDGRTYAPRANRSVILSAGAIHSPA
M271_40235          EPARQRPNLRVFGGATVDTVLVQGGR--AVGVRVRRGNDWTEVRA-QHVVLCAGAVHSPA
WP_050800219.1      EPARQRPNLRVFGGATVDTVLIEGGR--AVGVRLRRGNDWTEVRA-QHVVLCAGAVHSPA
SSOG_09008          EPARQRPNLRVFGGATVDTVLIEGGR--AVGVRLRRGNDWTEVRA-QHVVLCAGAVHSPA
WP_007670468.1      ESARDRGNLTIVGNALVDTLTIEGNRPHISGVSVWIDGERRSVRATREVILSAGAIHSPA
WP_051573876.1      EPARQRPNLRVFGCATVDTVLVEGGR--AVGVRLRRGNDWTEVRA-QHVVLCAGAVHSPA
WP_057613603.1      EPARQRPNLRIVGDATVDQVIVEGGR--AVGVRVRVDGTWTQVRA-ARVVLCAGAVNSPS
WP_020695985.1      EPARHRSNLTIVGQALVDTLQFEGNRPHVSGVAVRVGGETRTVRARREVILSAGAIHSPA
WP_028076038.1      EPARDRSNLRVIGGATVDQVLLEGGR--AAGVRARIAGEWVEARA-ENVILCAGAIGSPA
WP_015684527.1      EPARGRANLDIVGHALVDTITFEGNRLHASGVLVRVNGQSYTPRATREVMLCAGAIHSPA
                     .   : .* :   :    : .  .*    **               *:: **.  **:

MtdW                LLLRSGIGPAHRLERVGVAPVADLPGVGQGLQDHPALTLPIRLAAPGAA---PPRPTNCC
Ata10_protein       LLLRSGLGPADHLTSVGVPVRADLPGVGRNLQDHAALTLPVRLTGDAPA---PSRPTNCC
betA_1              VLQRSGIGPAGVLKGLGIPVVADLP-VGENLLDHPMMPLFLTLKDHARVGTLMHRHTNCC
M271_40235          ILLRSGIGPDGPVAAL--------P-VGEGMQEHPLALFWLYARPGNHP-DVDARQTNCC
WP_050800219.1      ILLRSGIGPDGPVAAL--------P-VGEGMQEHPLALFWLYARPGKQP-DIDARHANCC
SSOG_09008          ILLRSGIGPDGPVAAL--------P-VGEGMQEHPLALFWLYARPGKQP-DIDARHANCC
WP_007670468.1      ILQRSGIGPAALLQDLGVPVAVDRP-VGANLLDHPILGILLHLRSEAQVGTRMHRHTNCC
WP_051573876.1      ILLRSGIGPDGPVAAL--------P-VGEGMQEHPLALFWLYARPGKHP-DVDARQTNCC
WP_057613603.1      ILLRSGIGPGGRVASL--------P-VGEGMQEHPLALFWLHQRPEARP-GLDERQTNCC
WP_020695985.1      ILQRSGIGPRTVLDRLGIETVADRP-VGEHLLDHPILSLMLHLREDARVSTLMHRHTNCC
WP_028076038.1      ILLRSGIGPEGSAARL--------P-VGEGMQEHPLALFWLFHRPEARP-SVDDRQANCC
WP_015684527.1      ILQRSGIGPAALLERLGIPIRADLP-VGENLLDHPIVNALLHLRDGSQVNTLMHRHTNCC
                    :* ***:**      :        * **  : :*.     :            * :***
```

```
MtdW                LRFASGESDGEANELMLNALN----------------------EVGPGVGA-----VVL
Ata10_protein       VRLDAGLPGSRPNELMVNALN----------------------EVEPGLGA-----VVL
betA_1              LRYSSGLGGAGENDMIVIAGNLVS-L-----------------GANGDTGRGR---LVV
M271_40235          LRYSSGLEGAGENDMMIASINQTLALPDPGDSHLVAQGTGGTWGGAGGGHTAGGPGLLCL
WP_050800219.1      LRYSSGLAGAGENDMMITSINQTLALPNPGDSHLVAEGTGGTWGGAGGGHAAGGPGLLCL
SSOG_09008          LRYSSGLAGAGENDMMITSINQTLALPNPGDSHLVAEGTGGTWGGAGGGHAAGGPGLLCL
WP_007670468.1      LRYSSGLPGAGENDMIMIAGNLR--S-----------------EAEGDLARGR---IAV
WP_051573876.1      LRYSSGLEGAGENDMMIASINQTLALPDPGDSHLVAEGTGGTWGGAGGGHTAGGPGLLCL
WP_057613603.1      VRYSSDLADANENDMMIVSINQTLALPDMNDSHLVADGARGTWGGAGGGHASGGPGLLCL
WP_020695985.1      VRYSSGLAGAGVNDMIMIAGNLR--A-----------------EADGGTALAR---LAV
WP_028076038.1      VRYSSGLEGAGENDMMIVAVNQTLALPPDVTSHLVADQLAGTWGGAGGGQTAGGPGLLCL
WP_015684527.1      LRYSSGLAGAGDNDMIMIAGNLARSQ-----------------QSIAEVTLGR---IAV
                    :*  :.   .  *:::: : *                                 : :

MtdW                ALFRPEARGRVEVTGAGLDDDLVVDLGLLGTERDLRRMRAGVARLADLARHPAFERVGSP
Ata10_protein       ALFRPESTGRVELAGP--DRGLLVDLDFLSTDADLARMRAGAALLAEIAGHPALTAVGQP
betA_1              SVYQAFSQGHVRITTPDPSVDPAVEERMLSDERDLVRMRDGVLRLRALARHEAIRAITTR
M271_40235          WANQQFSRGVLRLASADPDVHPVMDQDLLSDRADLVRMRDGVRHCVELLRSGPFDTAFEH
WP_050800219.1      WANQQFSRGVLRLASADPDAHPVIDQDLLSDPSDLVRMRDGVKHCLELLRSGPFDTAFEH
SSOG_09008          WANQQFSRGVLRLASADPDAHPVIDQDLLSDPSDLVRMRDGVKHCLELLRSGPFDTAFEH
WP_007670468.1      SVYQAFSQGTVRIASSDPHLDPVVEERMLSDERDLLRMRDGVRRLCAVGRQDGVHSIATR
WP_051573876.1      WANQQFSRGVLRLASADPDVHPVMDQDLLSHPADLVRMRDGVKHCVELLRSGPFDTAFEH
WP_057613603.1      WANQEFSRGTLRLASPDPDVHPVIEQNLLNDPGDLTRMRDGVHRCLELIRSGAFDTAFEH
WP_020695985.1      SAFQAFSEGTVRIASRDPAIDPQVDERMLSSESDYLRMRDGVMRLQEIARHPDVQAIATR
WP_028076038.1      WLNQQFGRGSLRLASPDPDVHPLIDQDLLNVDADLVRARDGIKRSLELLRGGSFDREFAH
WP_015684527.1      SVYQAFSQGHVCIVTTDPTIDPIVEERMLSDSRDLVRLRDGVRRLRDICLQPAVTDIAHR
                       :  . * : :          :   :*.   *  * . *      :     .

MtdW                VGADLLR---GLARRPEELDSWMAARCHEAWHLVGSCRMGDPADPATVVDPRGRVKGVAG
Ata10_protein       IGAAALR---ARLGDPAALDAWLRARCHEAWHLVGTCRMGSPADPGAVVGPDCRVHGVAG
betA_1              VDYGSTGRSIEEDLSPAELDDWLFSECSDAQHASGTCRMGAADDPRSVVDPDCRVIGCTG
M271_40235          IAIDIAGRGLDALSDDATVDRWLMETIGDTGHICGTCRMGAPDDPRTVVDPSGRVLGVDG
WP_050800219.1      IAIDIAGRGLDTLSDDATVDRWLMETIGDTGHICGTCRMGAPDDPRAVVDPSGRVLGVDG
SSOG_09008          IAIDIAGRGLDTLSDDATVDRWLMETIGDTGHICGTCRMGAPDDPRAVVDPSGRVLGVDG
WP_007670468.1      VEYGISGRSIEDELTGIDLDDWMMAECSDAQHASGTCRMGPVSDPRSVVDPACRVIGCTG
WP_051573876.1      IAIDIAGRGLDTLSDDATVDRWLMETIGDTGHICGTCRMGAPDDPRAVVDPSGRVLGVDG
WP_057613603.1      IAVDLAGTGTDALTDDTAIDRWLLETIGDTGHICGTCRMGAPDDPRAVVDPAGRVLGVEG
WP_020695985.1      ADYGMSGRSIDEPFAPKELEDWMFAECSDAQHASGTCRMGAADDPRSVVDPECRVIGCTG
WP_028076038.1      VAIDLTGRGIDELSDDRAIDAWLLATIGDTGHICGTCRMGSPEDPRTVVDPEGRVLGVDG
WP_015684527.1      VDYGASGRSMDEALGDAELDDWLFAECSDAQHASGTCRMGAPDDARSVVDHECRVIGCTG
                                  :  ::     :: *  .:****    . :**.    * *  *
```

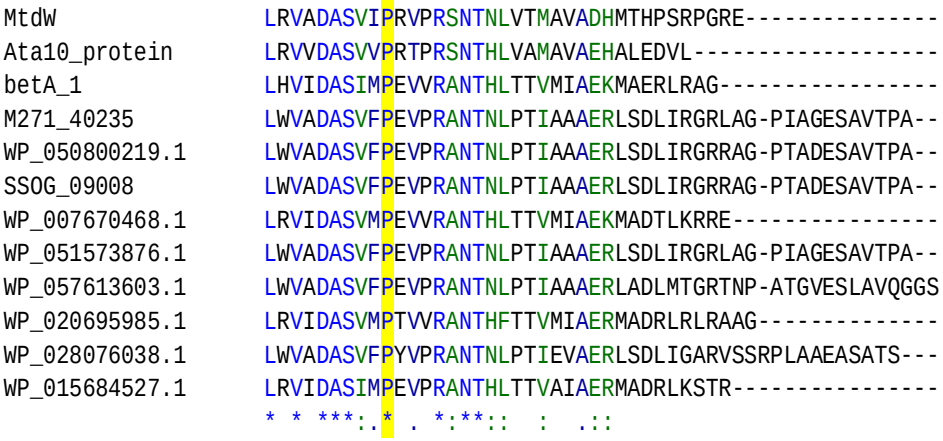


Fig. S29. Multiple sequence alignments for MtdW with its homologues (Ata10_protein, betA_1, M271_40235, WP_050800219.1, SSOG_09008, WP_007670468.1, WP_051573876.1, WP_057613603.1, WP_020695985.1, WP_028076038.1, WP_015684527.1). The following proteins (with GenBank IDs) were used for amino acid alignment: Ata10_protein (CAD62204.1) from *Saccharothrix mutabilis* subsp. capreolus, betA_1 (CEJ10070.1) from *bacterium* YEK0313, M271_40235 (AGP59428.1) from *Streptomyces rapamycinicus* NRRL 5491, WP_050800219.1 from *Streptomyces himastatinicus*, SSOG_09008 (EFL29294.1) from *Streptomyces himastatinicus* ATCC 53653, WP_007670468.1 from *alpha proteobacterium* BAL199, WP_051573876.1 from *Streptomyces* sp. PRh5, WP_057613603.1 from *Streptomyces* sp. Root369, WP_020695985.1 from *Reyranella massiliensis*, WP_028076038.1 from *Solirubrobacterales bacterium* URHD0059, WP_015684527.1 from *Bradyrhizobium* sp. S23321. The binding motif GxGxxG(x)$_{18}$E in the N terminus for ribose moiety of FAD was highlighted in red color. The conserved H(x)$_{37}$P motif in the C terminus presumbly involved in substrate oxidation was highlighted in yellow color.

Fig. S30. ^{1}H NMR (500 MHz) spectrum of compound **1** in CD_3OD

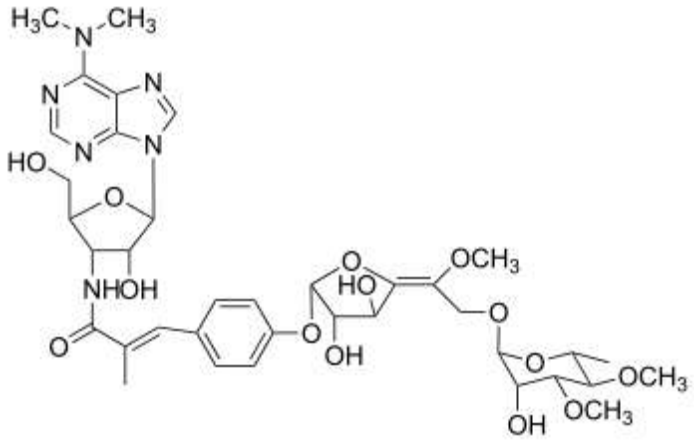


Fig. S31. ^{13}C NMR (125 MHz) spectrum of compound **1** in CD_3OD

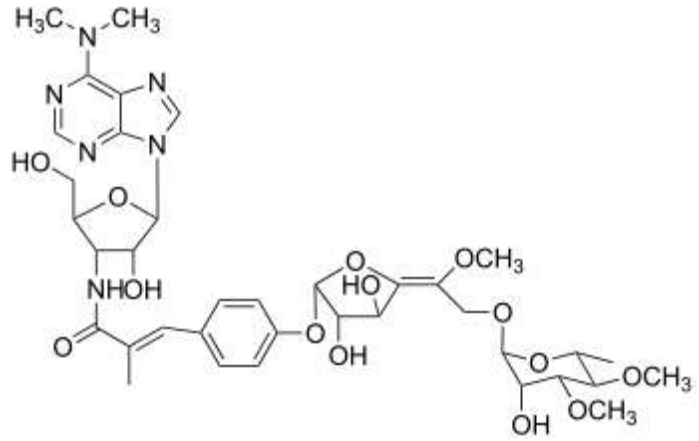


Fig. S32. DEPT 135 NMR (125 MHz) spectrum of compound **1** in CD_3OD

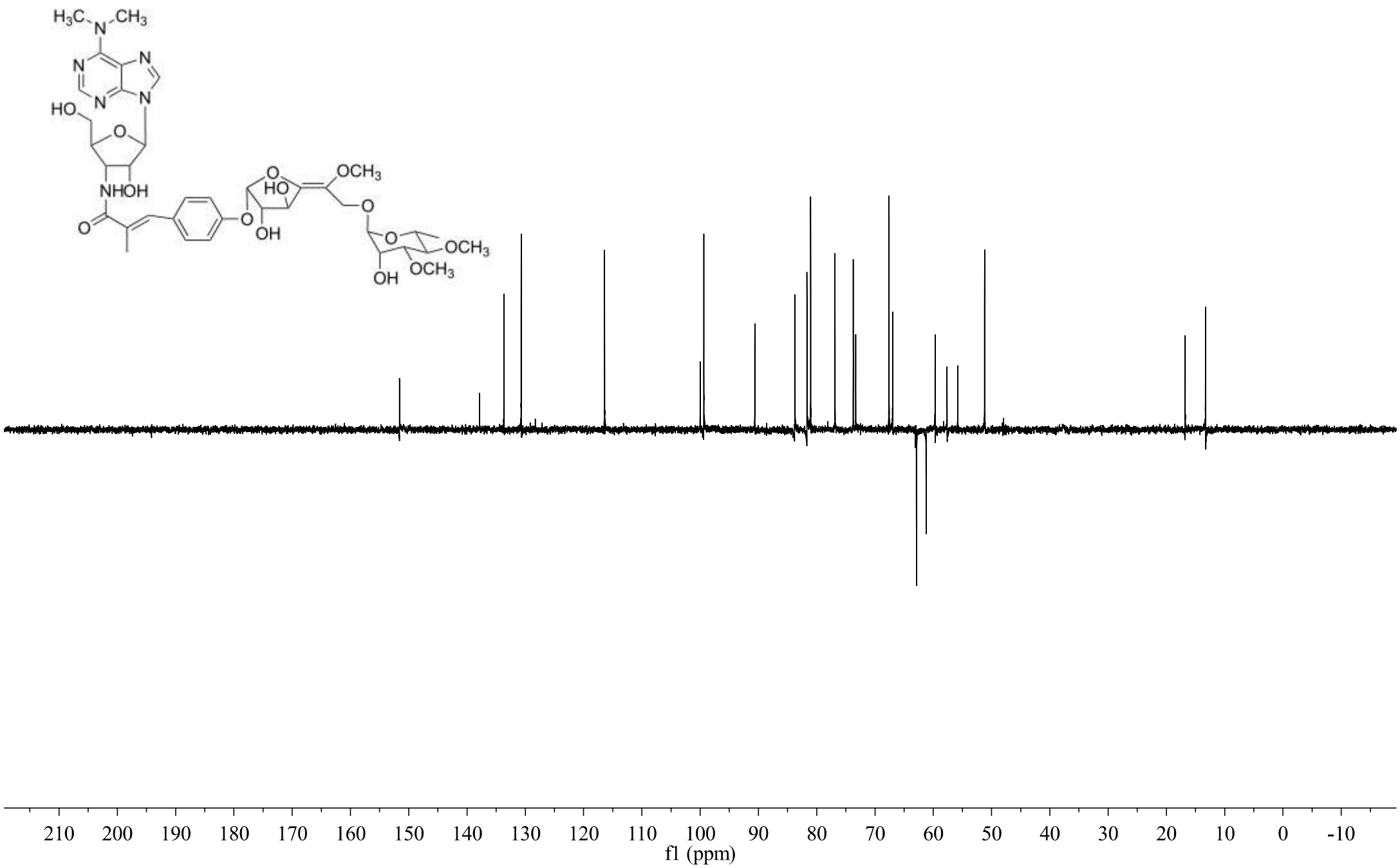


Fig. S33. HSQC spectrum of compound **1** in CD_3OD

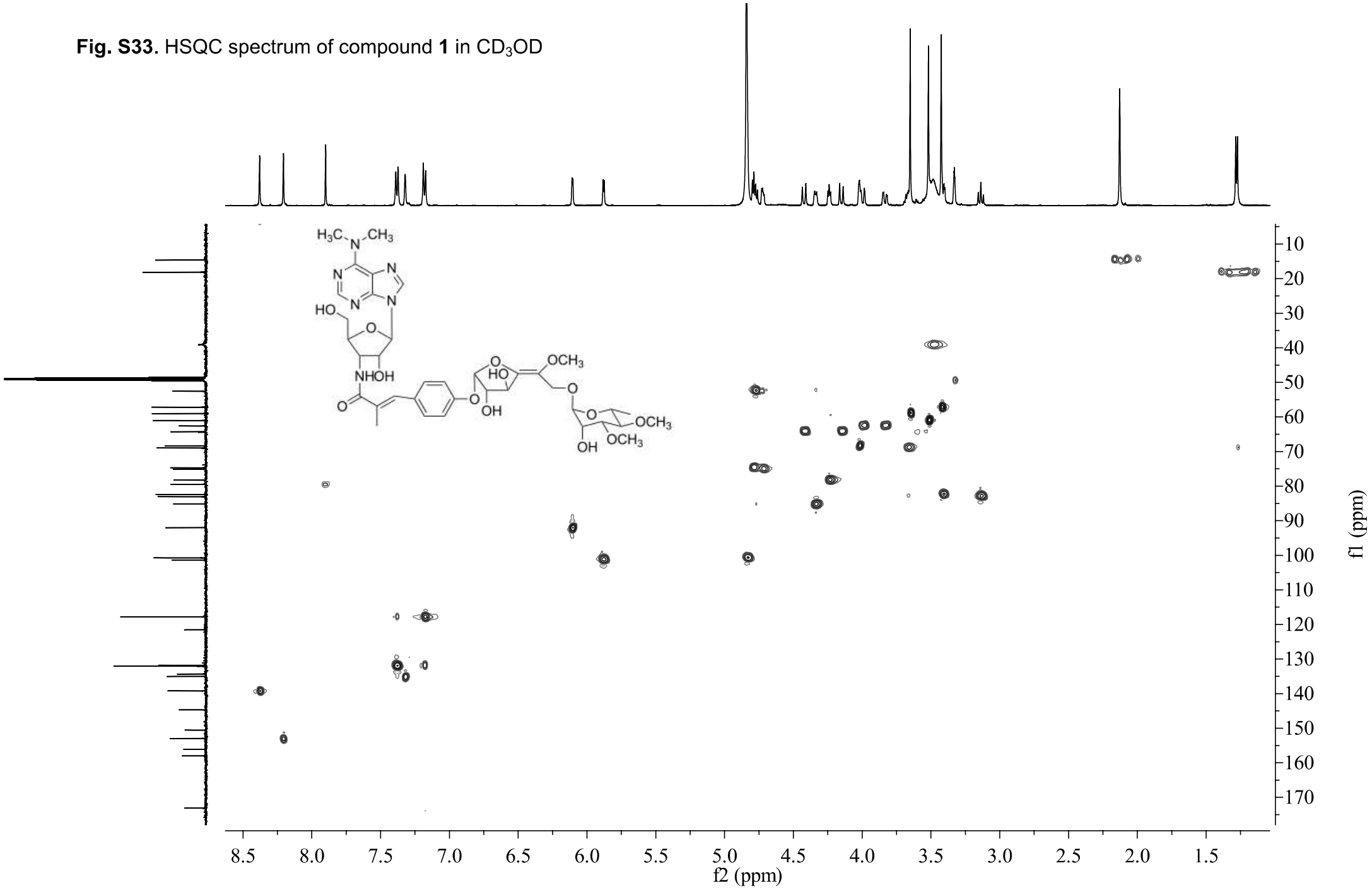


Fig. S34. ^{1}H-^{1}H COSY spectrum of compound **1** in CD_3OD

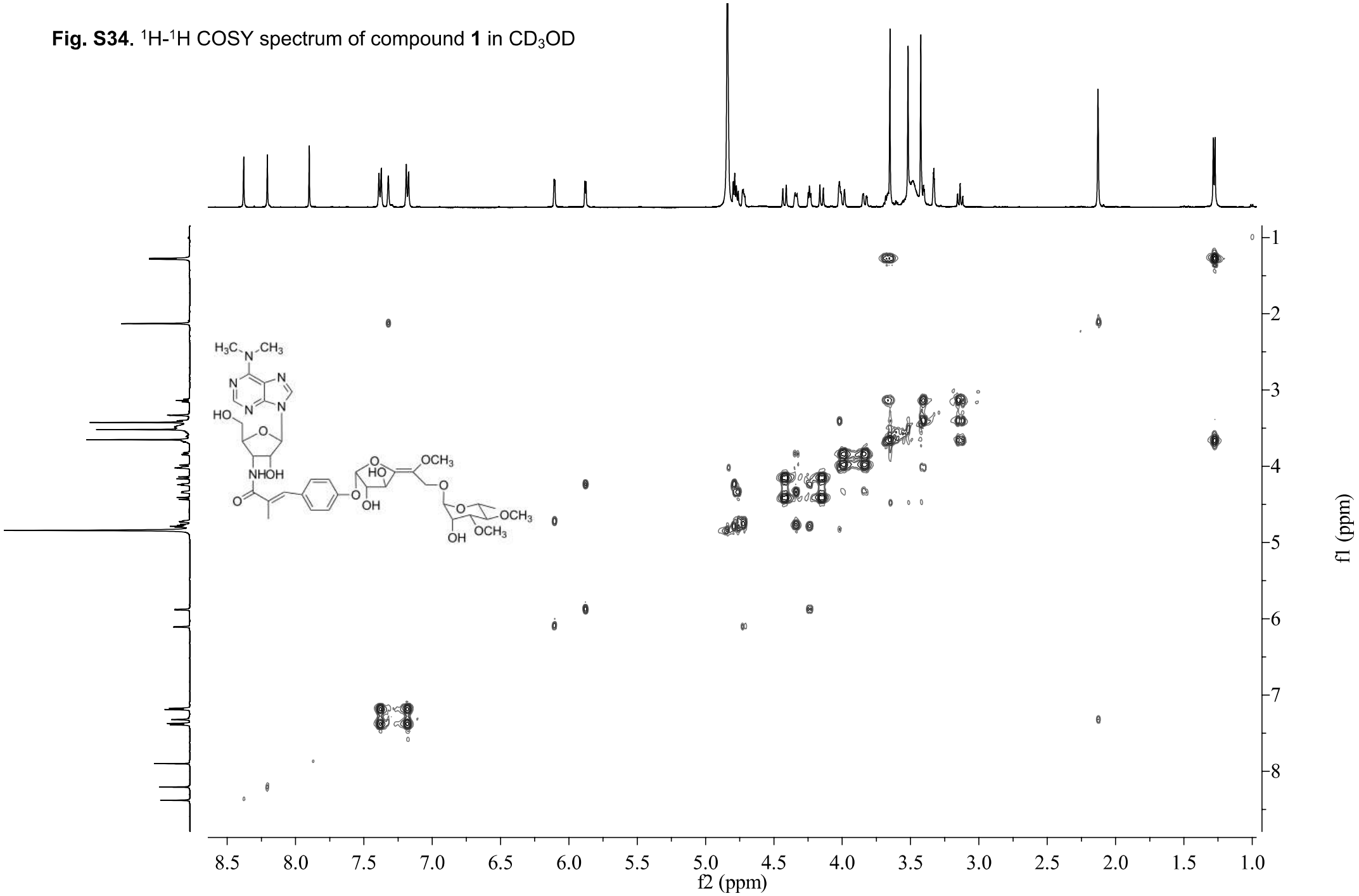


Fig. S35. HMBC spectrum of compound **1** in CD_3OD

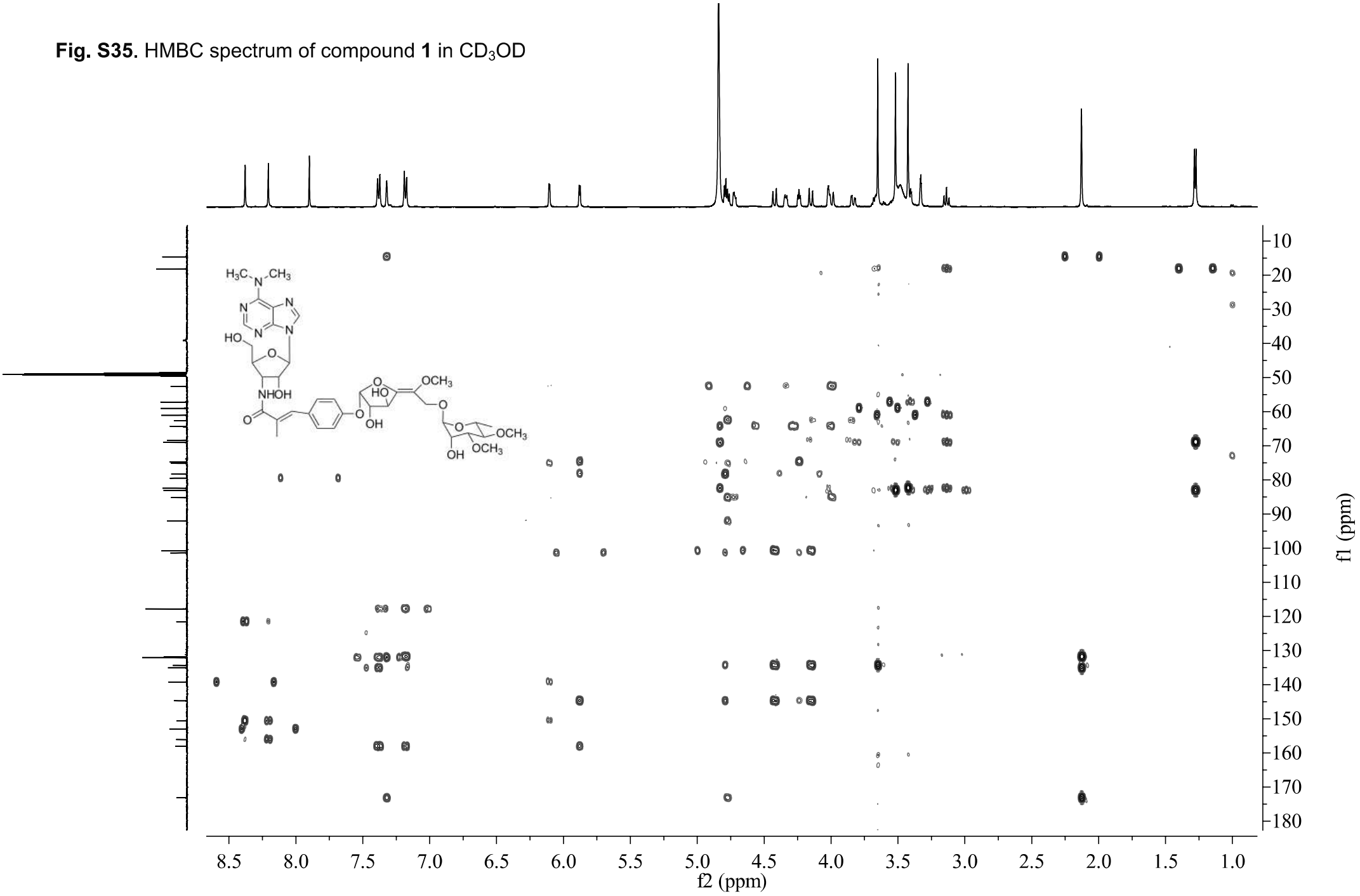


Fig. S36. NOESY spectrum of compound **1** in CD_3OD

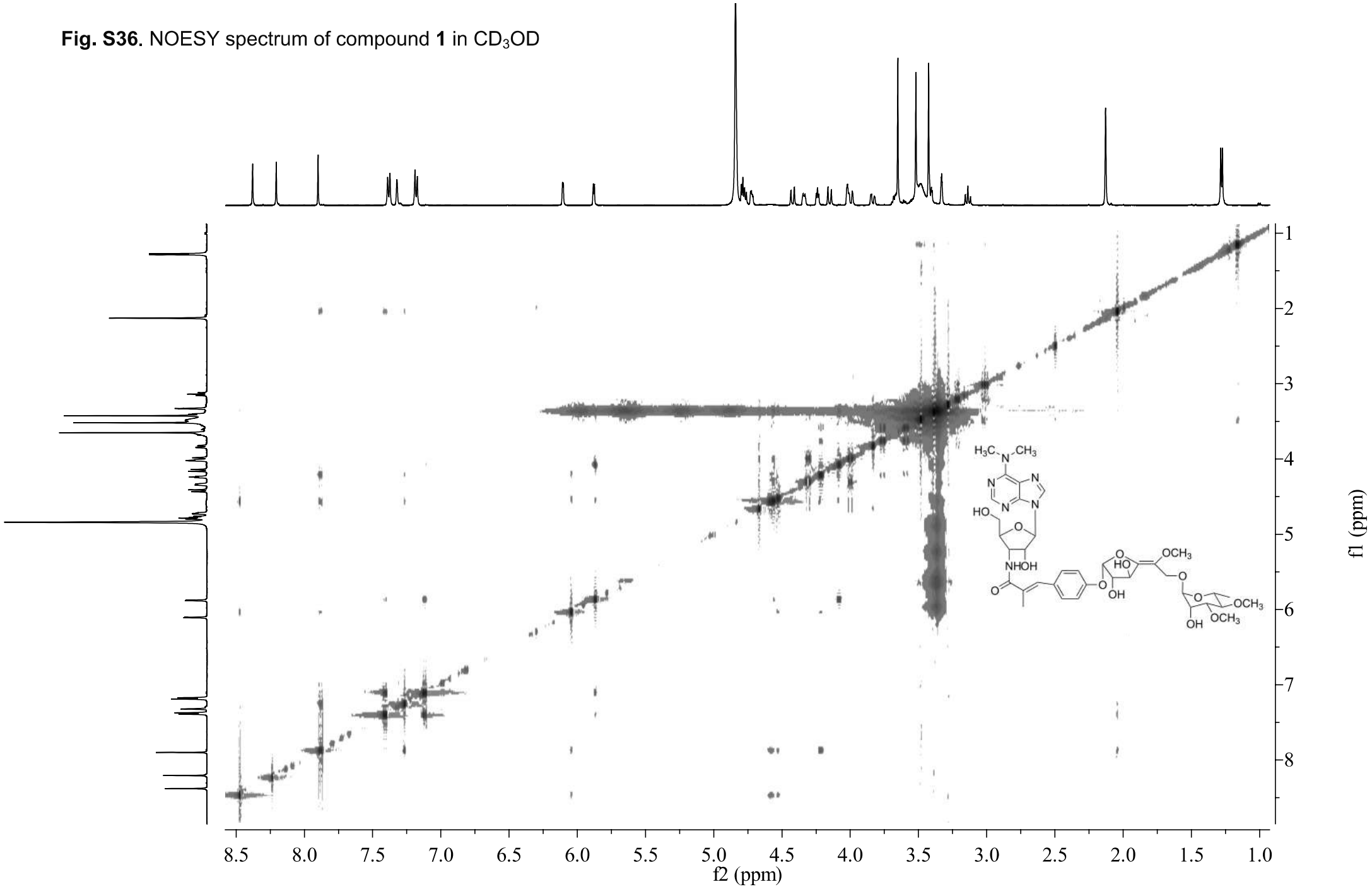


Fig. S37. ^{1}H NMR (500 MHz) spectrum of compound **2** in CD_3OD

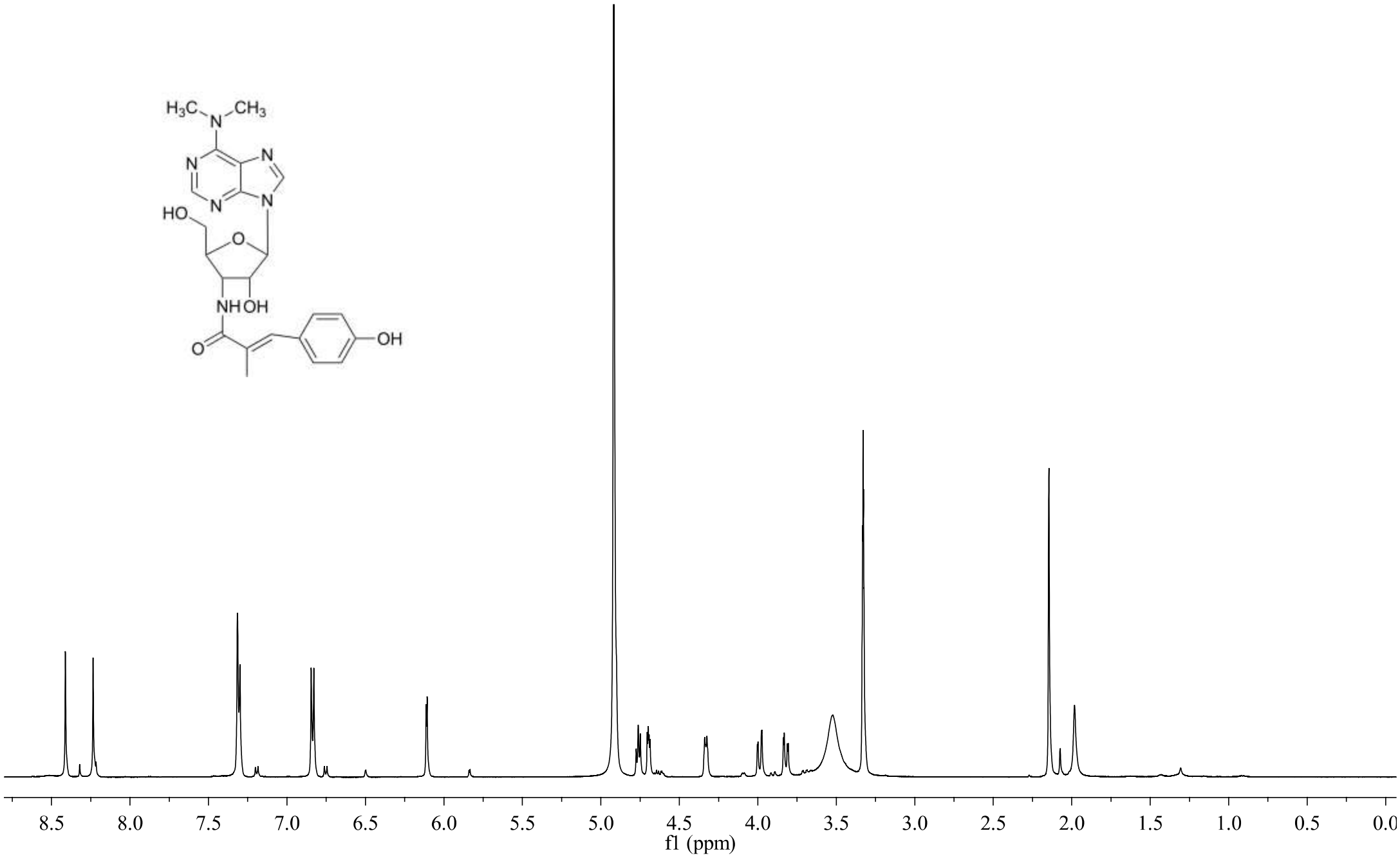


Fig. S38. ^{13}C NMR (125 MHz) spectrum of compound **2** in CD_3OD

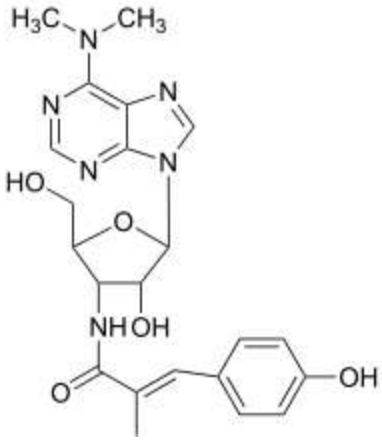


Fig. S39. ^{1}H NMR (500 MHz) spectrum of compound **3** in CD_3OD

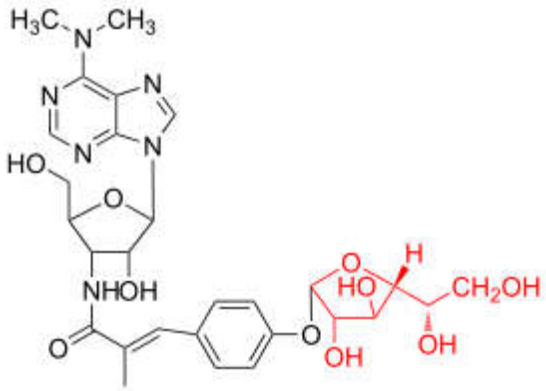


Fig. S40. HSQC spectrum of compound **3** in CD_3OD

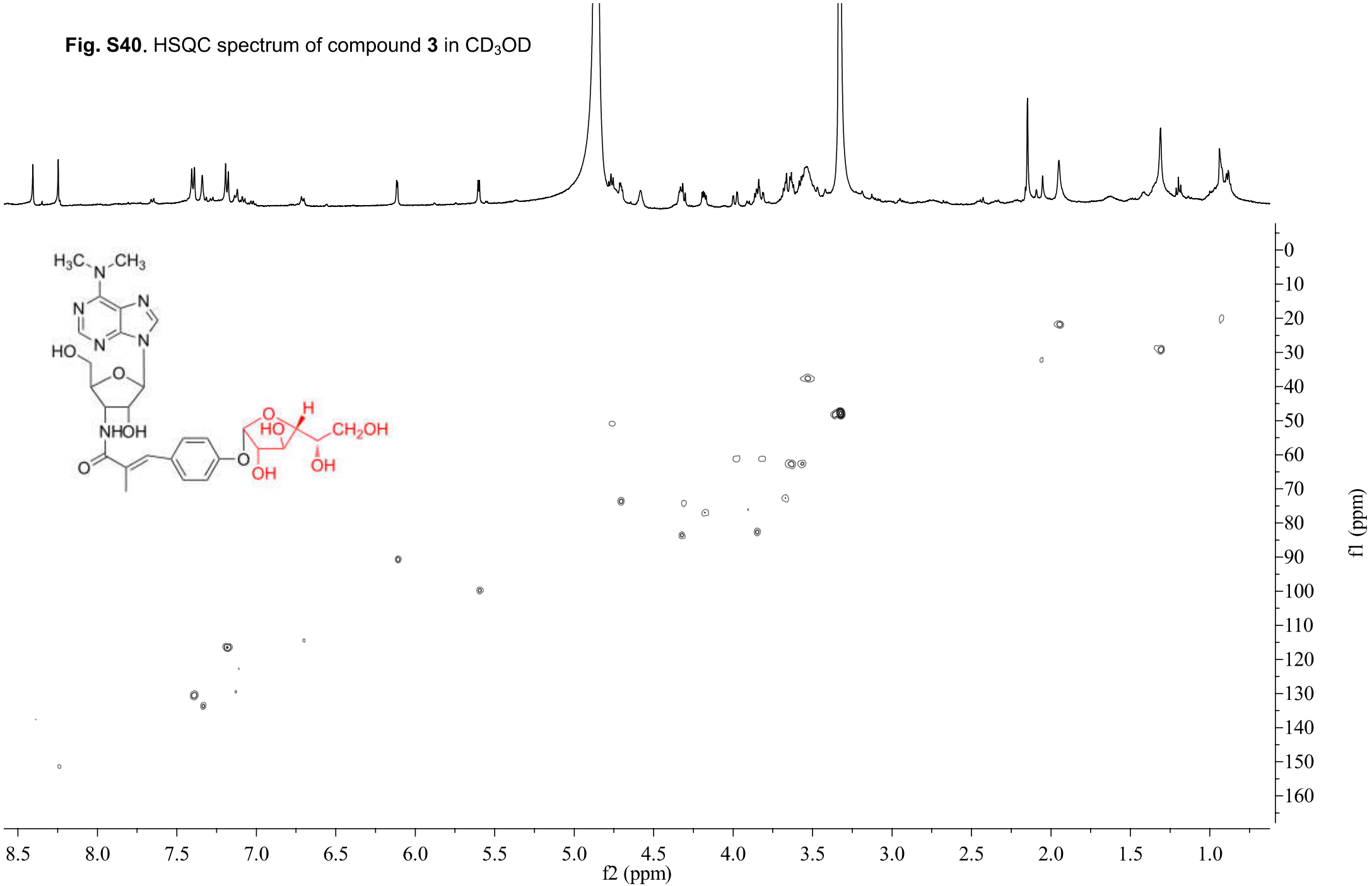


Fig. S41. ^{1}H-^{1}H COSY spectrum of compound **3** in CD_3OD

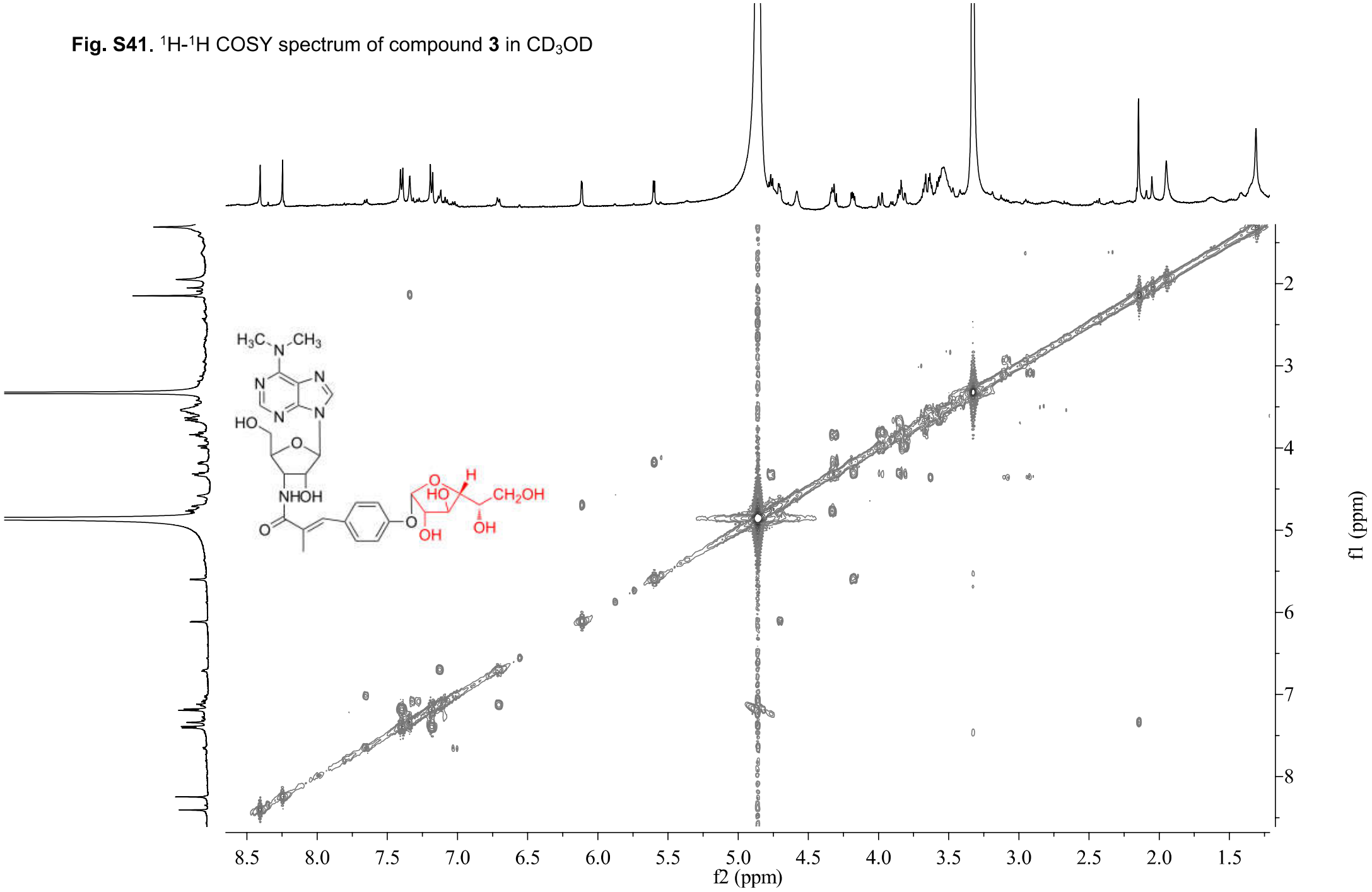


Fig. S42. HMBC spectrum of compound **3** in CD_3OD

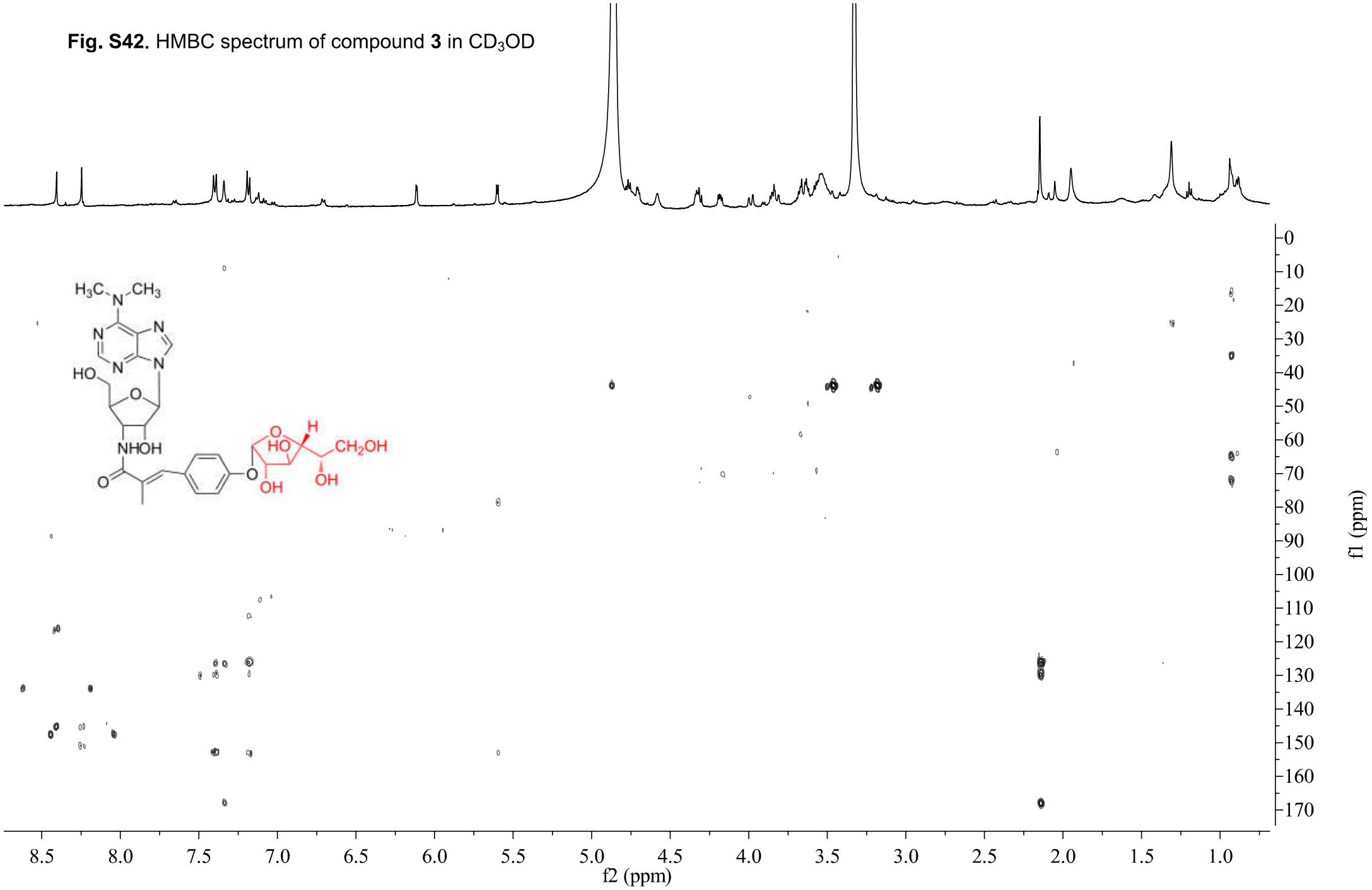


Fig. S43. NOESY spectrum of compound **3** in CD_3OD

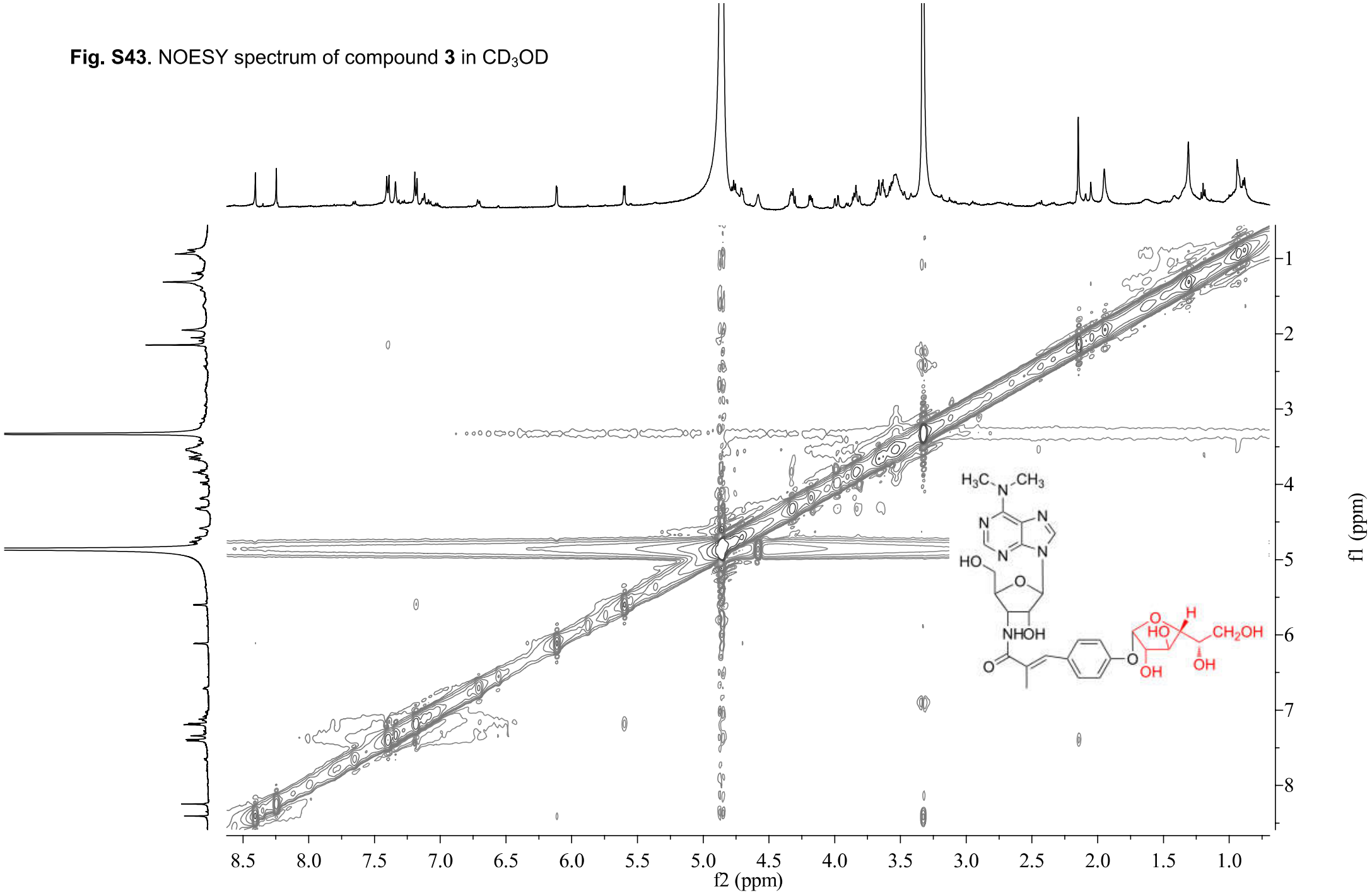


Fig. S44. ^{1}H NMR (500 MHz) spectrum of compound **4** in CD_3OD and $CDCl_3$

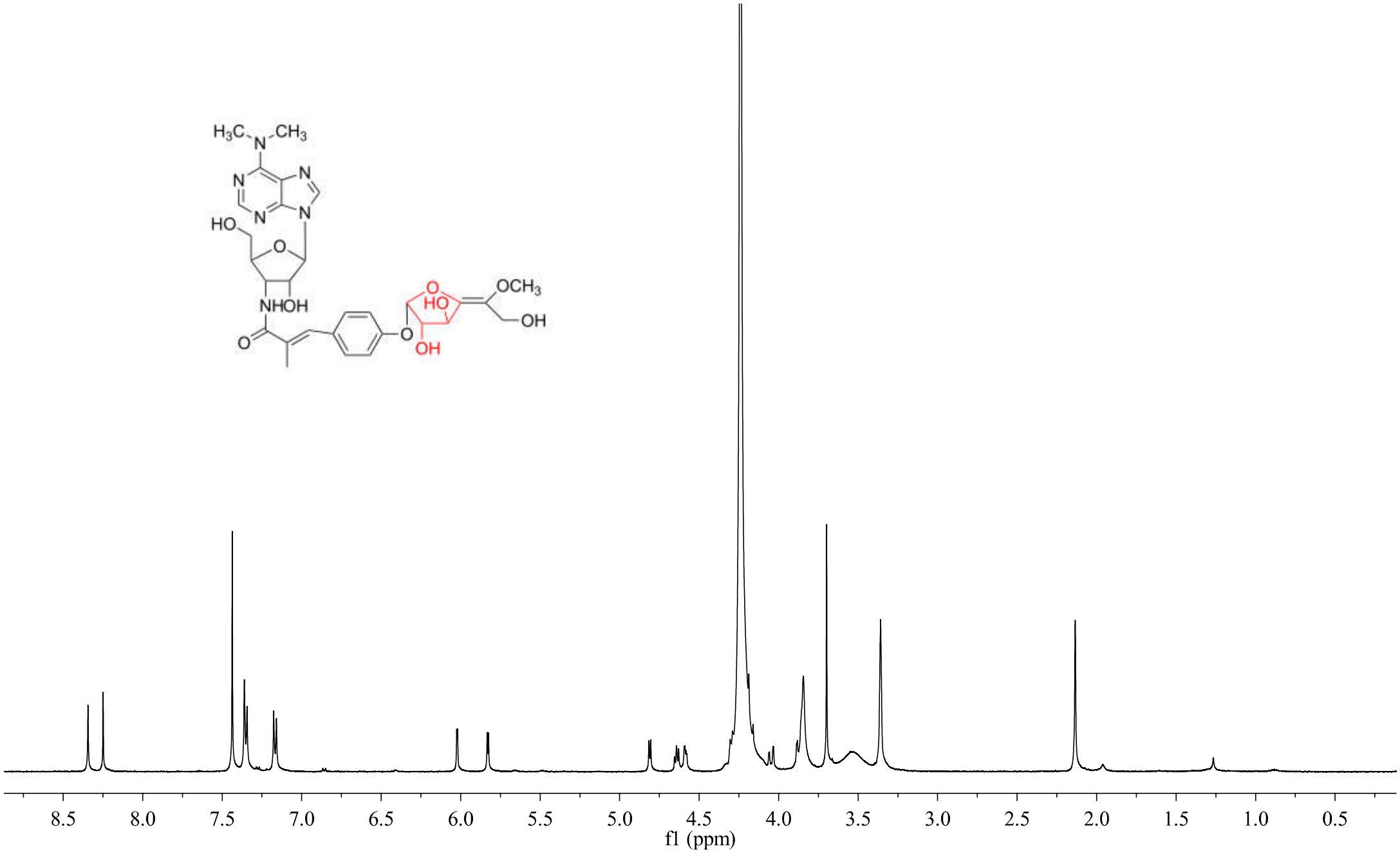


Fig. S45. ^{13}C NMR (125 MHz) spectrum of compound **4** in CD_3OD and $CDCl_3$

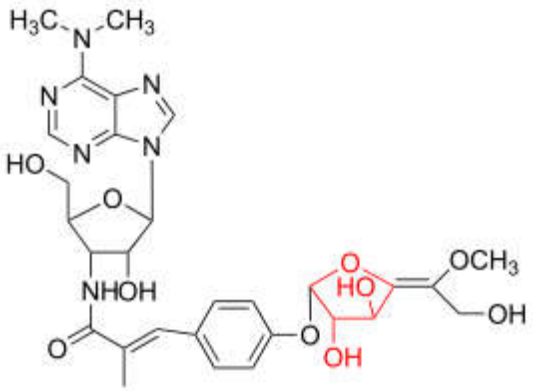


Fig. S46. ^{1}H NMR (500 MHz) spectrum of compound **5** in CD_3OD

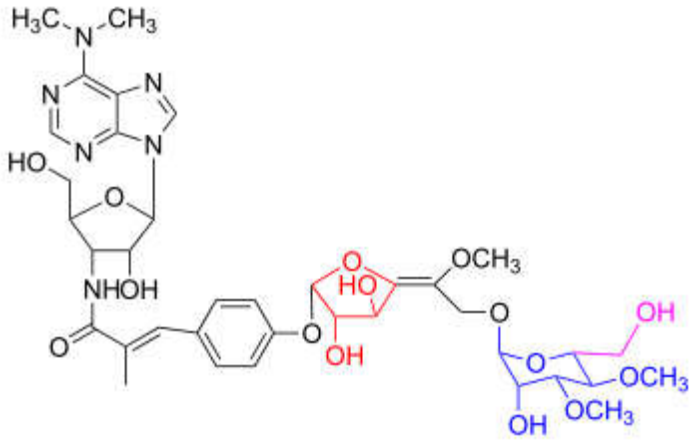


Fig. S47. ^{13}C NMR (125 MHz) spectrum of compound **5** in CD_3OD

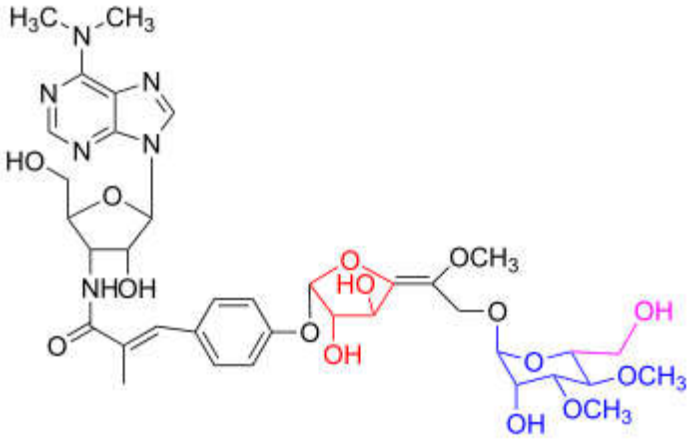


Fig. S48. HSQC spectrum of compound **5** in CD_3OD

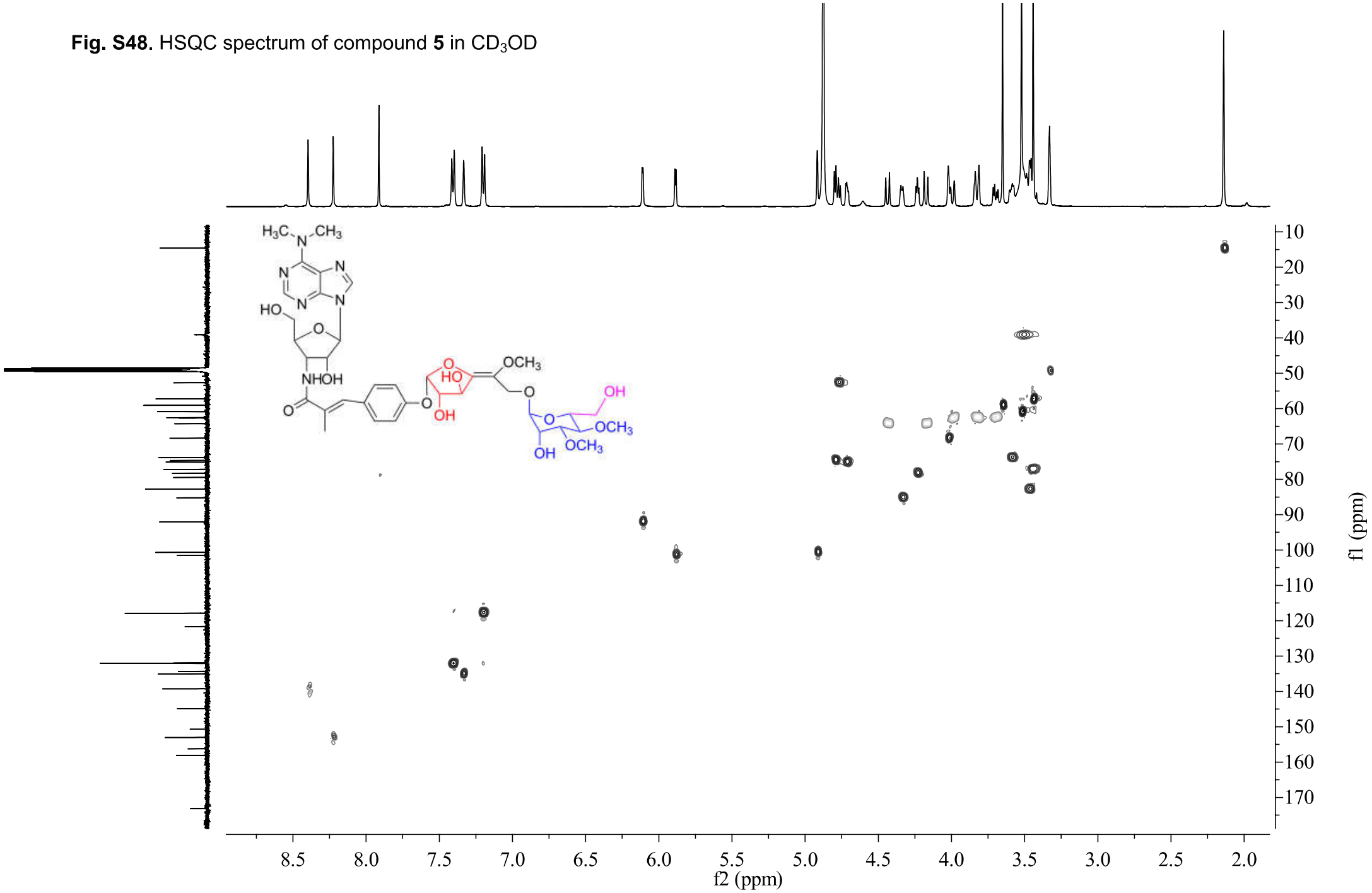


Fig. S49. ^{1}H-^{1}H COSY spectrum of compound **5** in CD_3OD

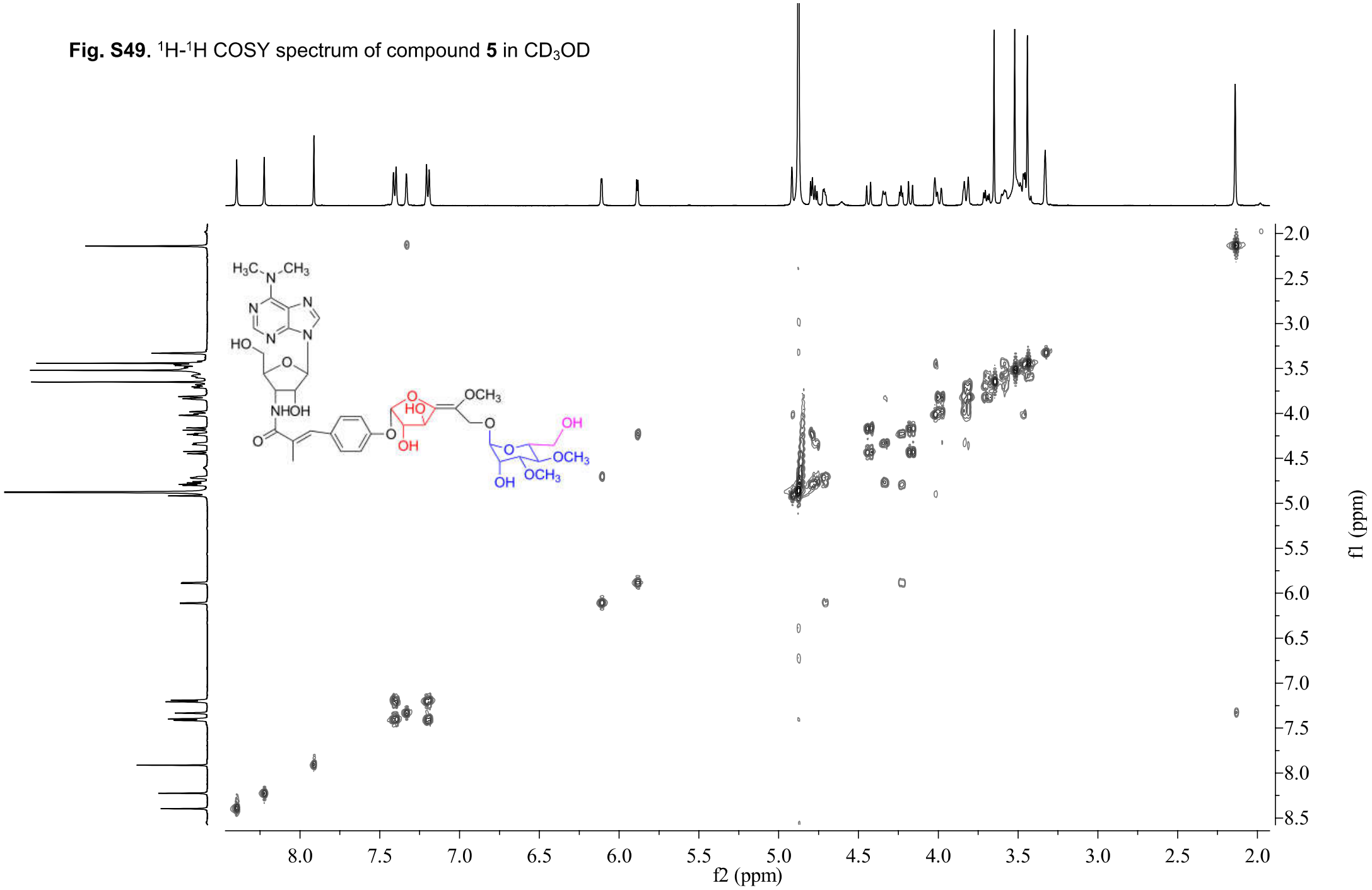


Fig. S50. HMBC spectrum of compound **5** in CD_3OD

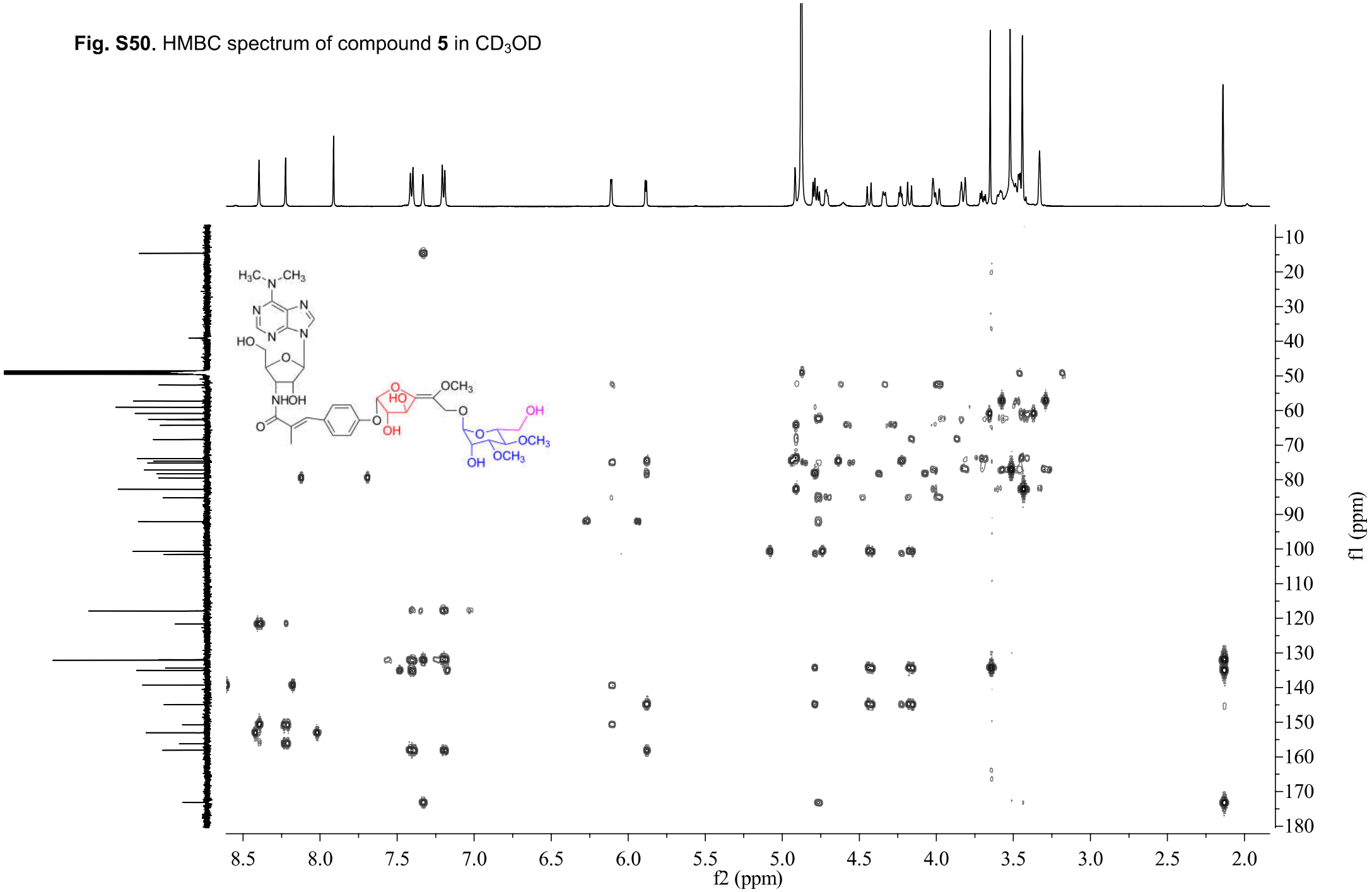


Fig. S51. ^{1}H NMR (500 MHz) spectrum of compound **6** in CD_3OD

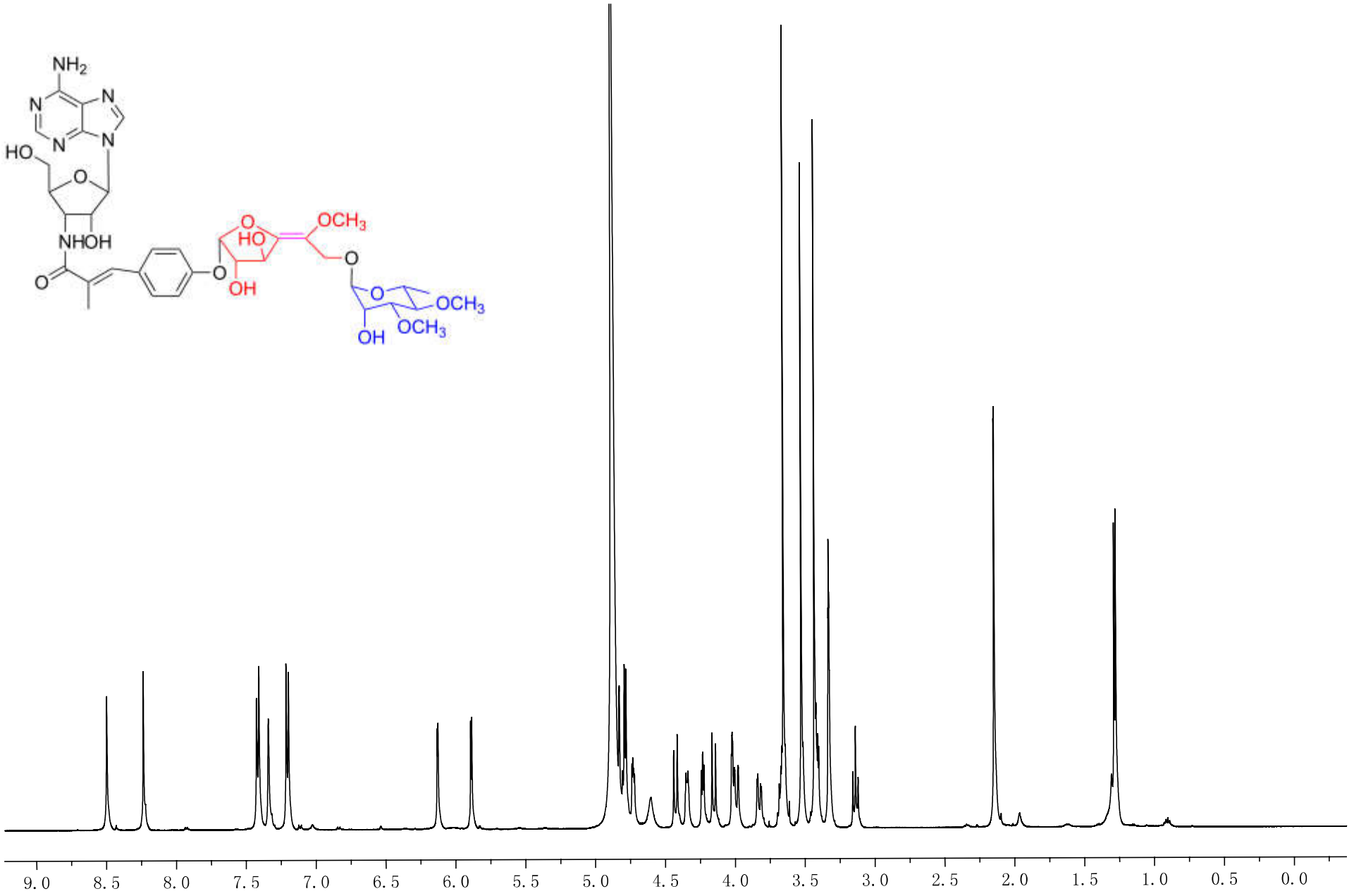


Fig. S52. ^{13}C NMR (125 MHz) spectrum of compound **6** in CD_3OD

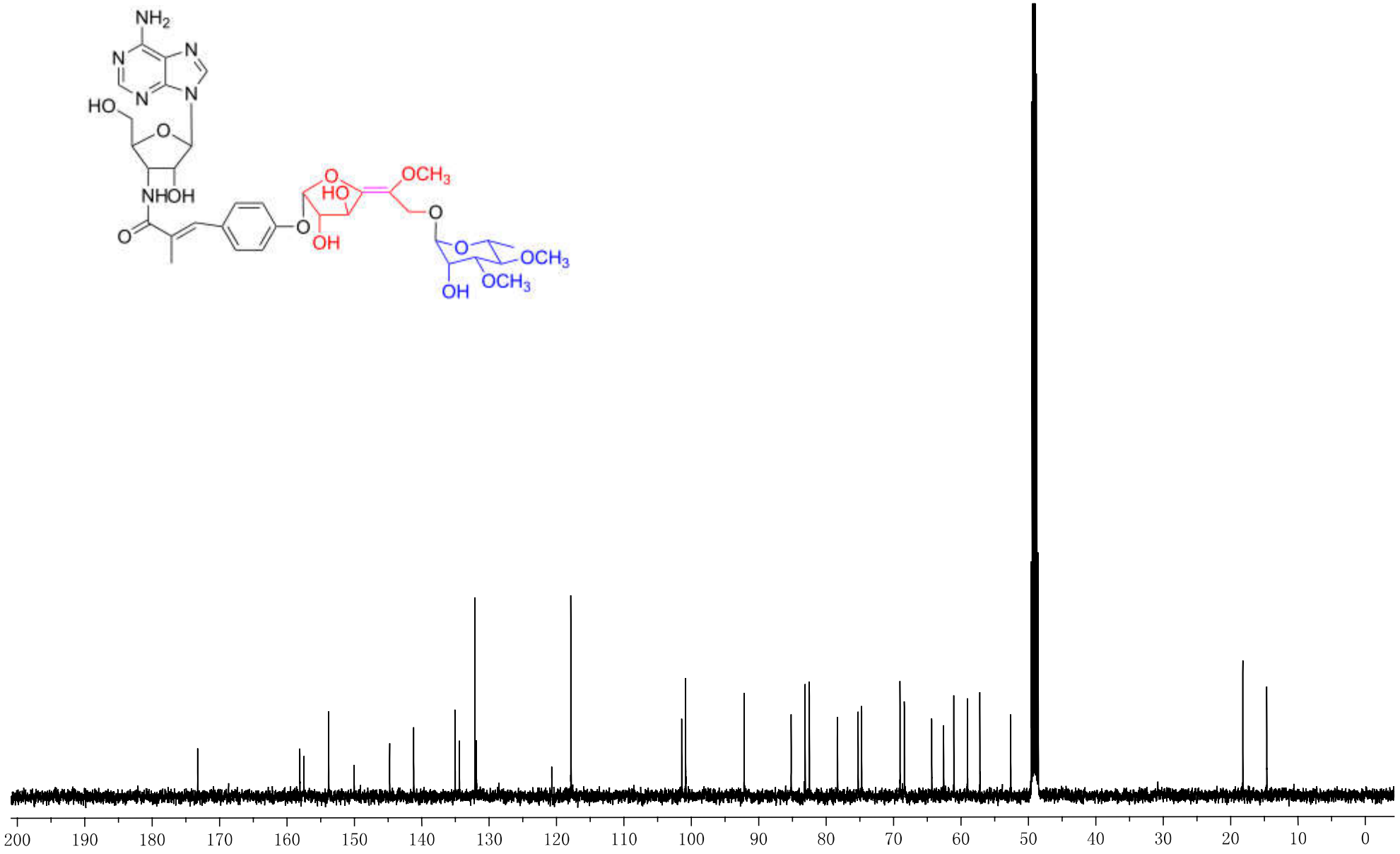


Fig. S53. HSQC spectrum of compound **6** in CD_3OD

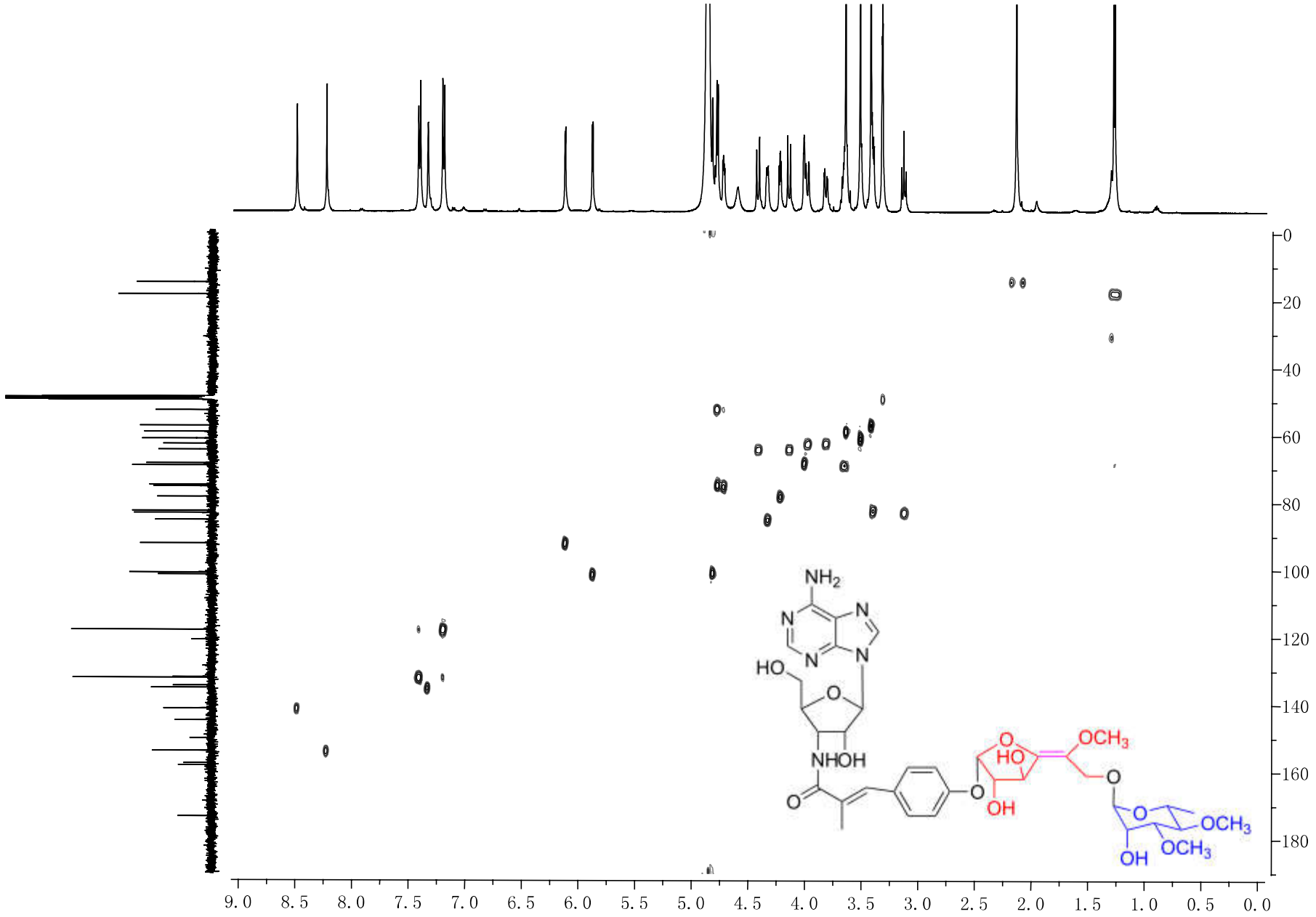


Fig. S54. ^{1}H-^{1}H COSY spectrum of compound **6** in CD_3OD

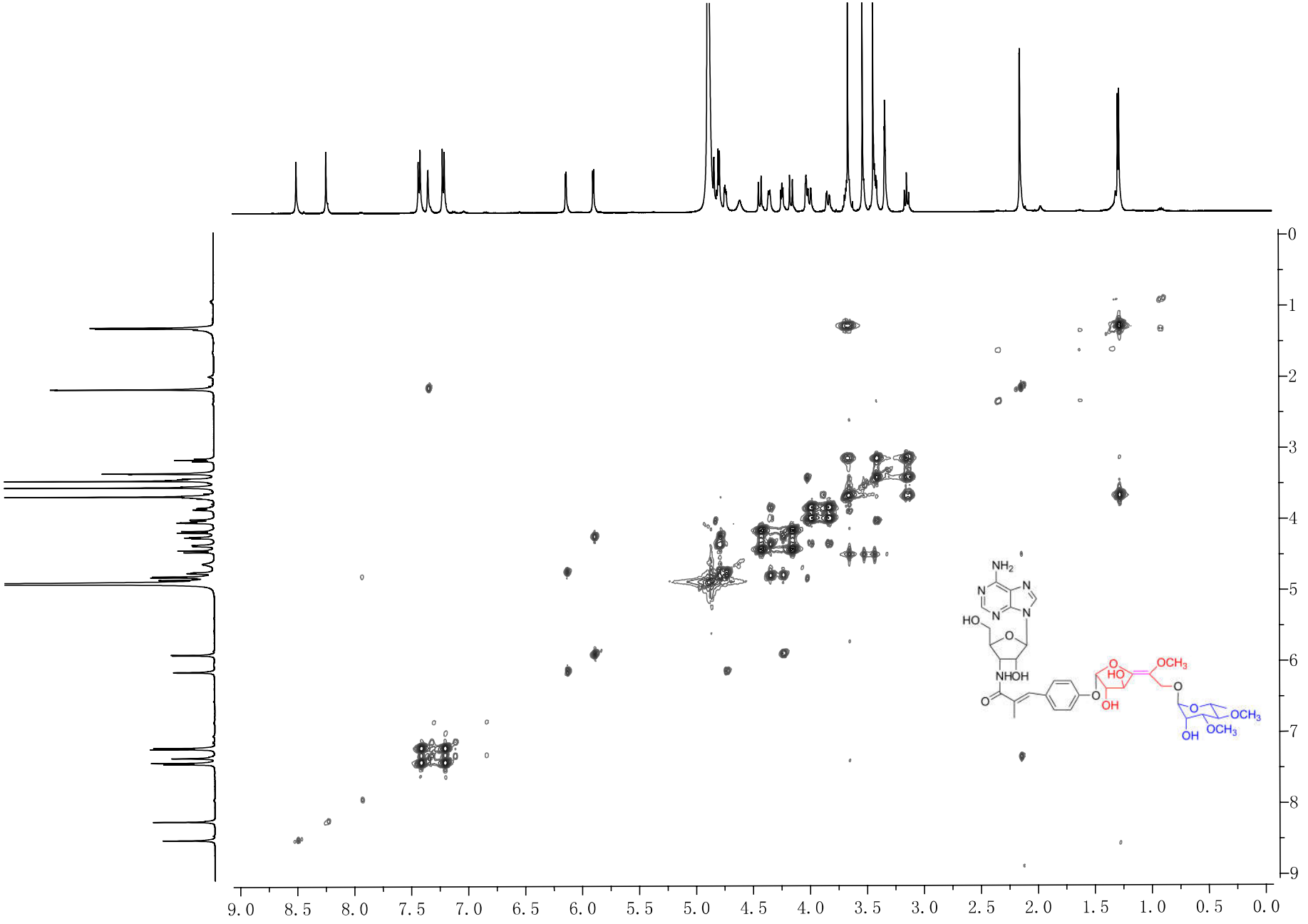


Fig. S55. HMBC spectrum of compound **6** in CD_3OD

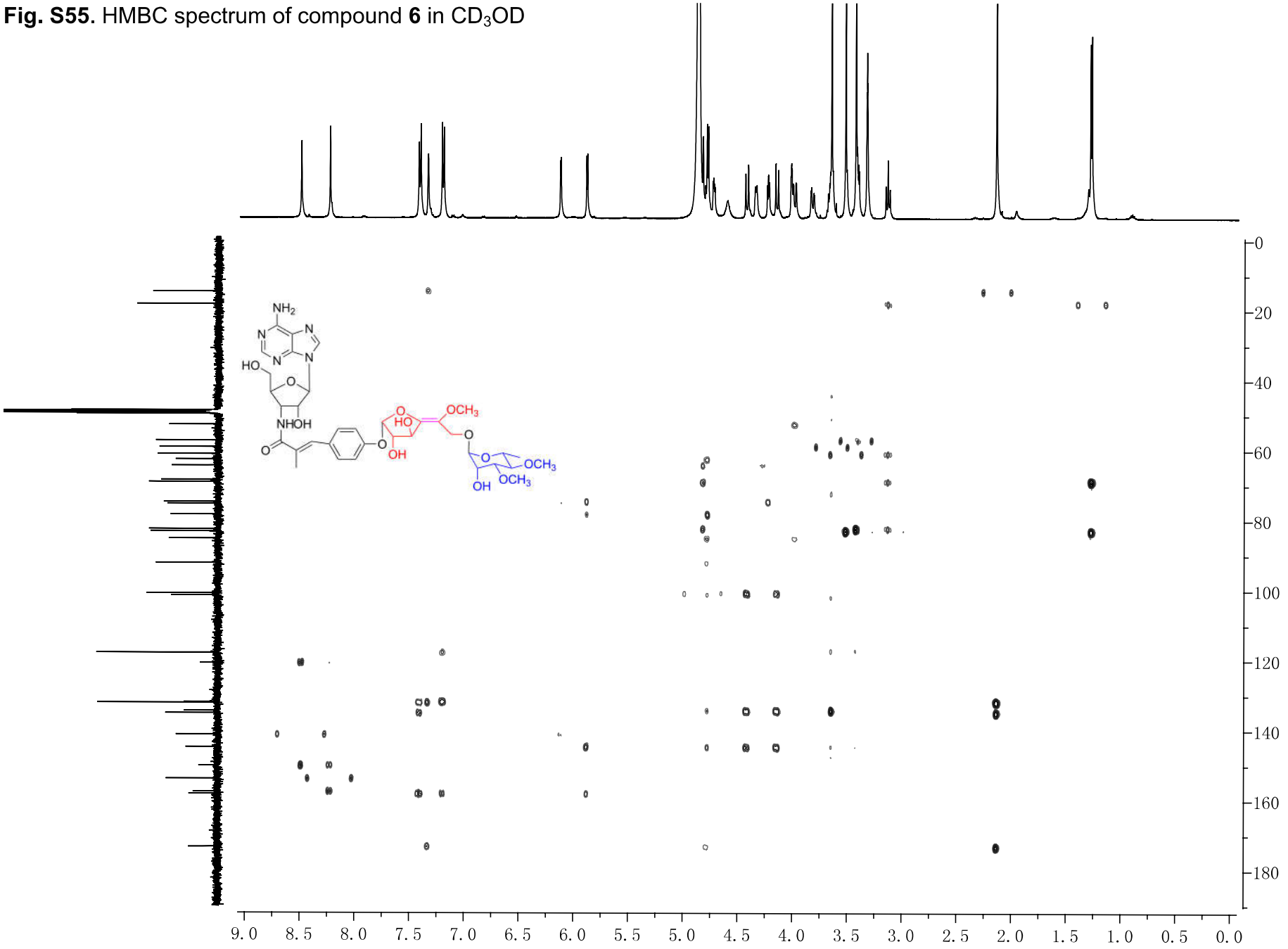


Fig. S56. ^{1}H NMR (500 MHz) spectrum of compound **7** in CD_3OD

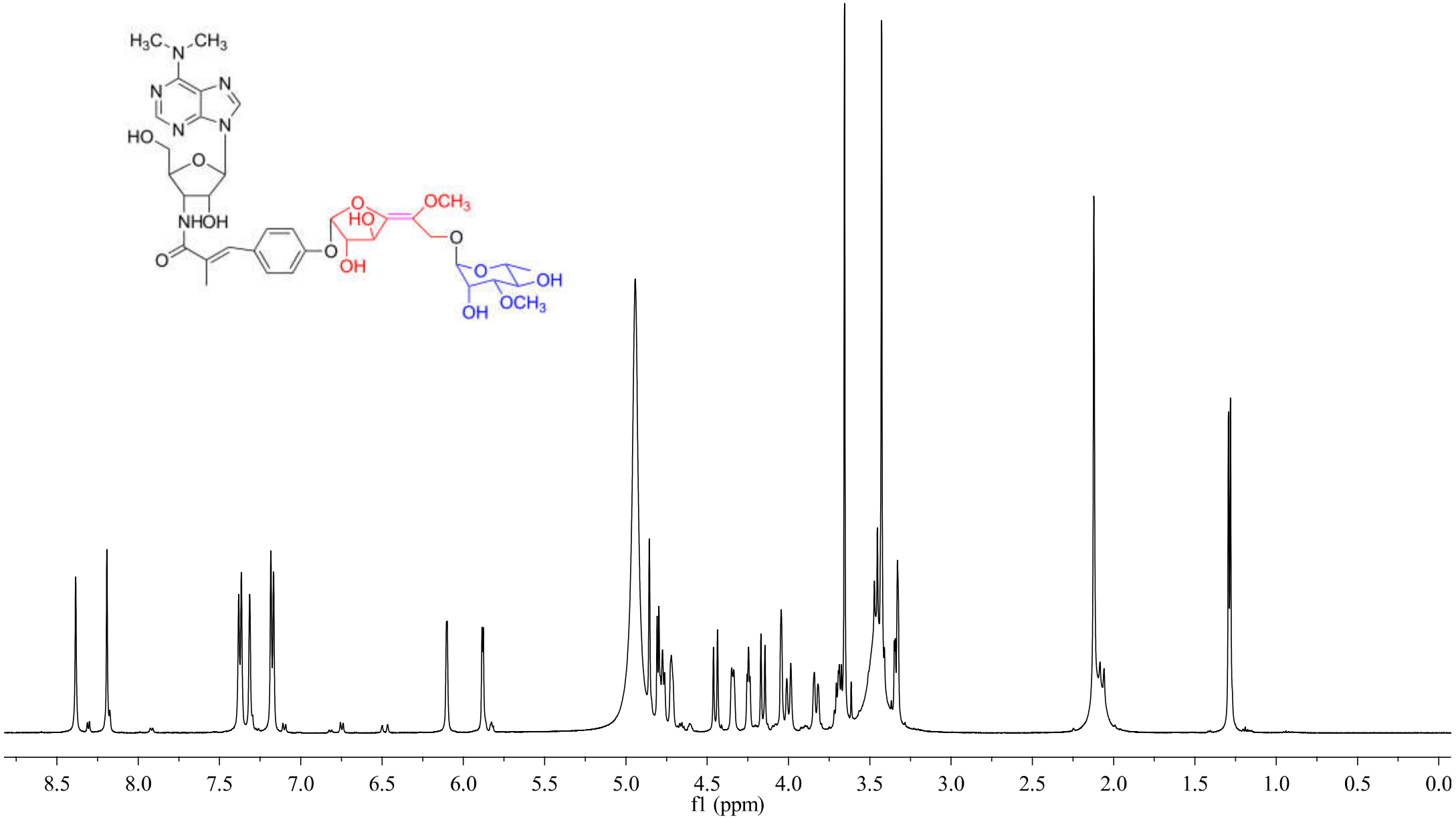


Fig. S57. ^{13}C NMR (125 MHz) spectrum of compound **7** in CD_3OD

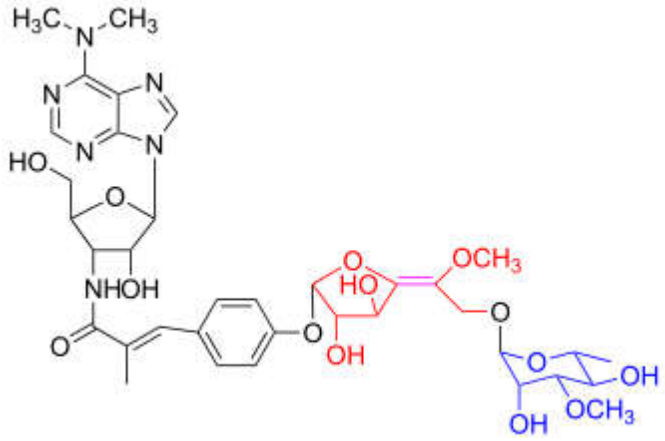


Fig. S58. HSQC spectrum of compound **7** in CD_3OD

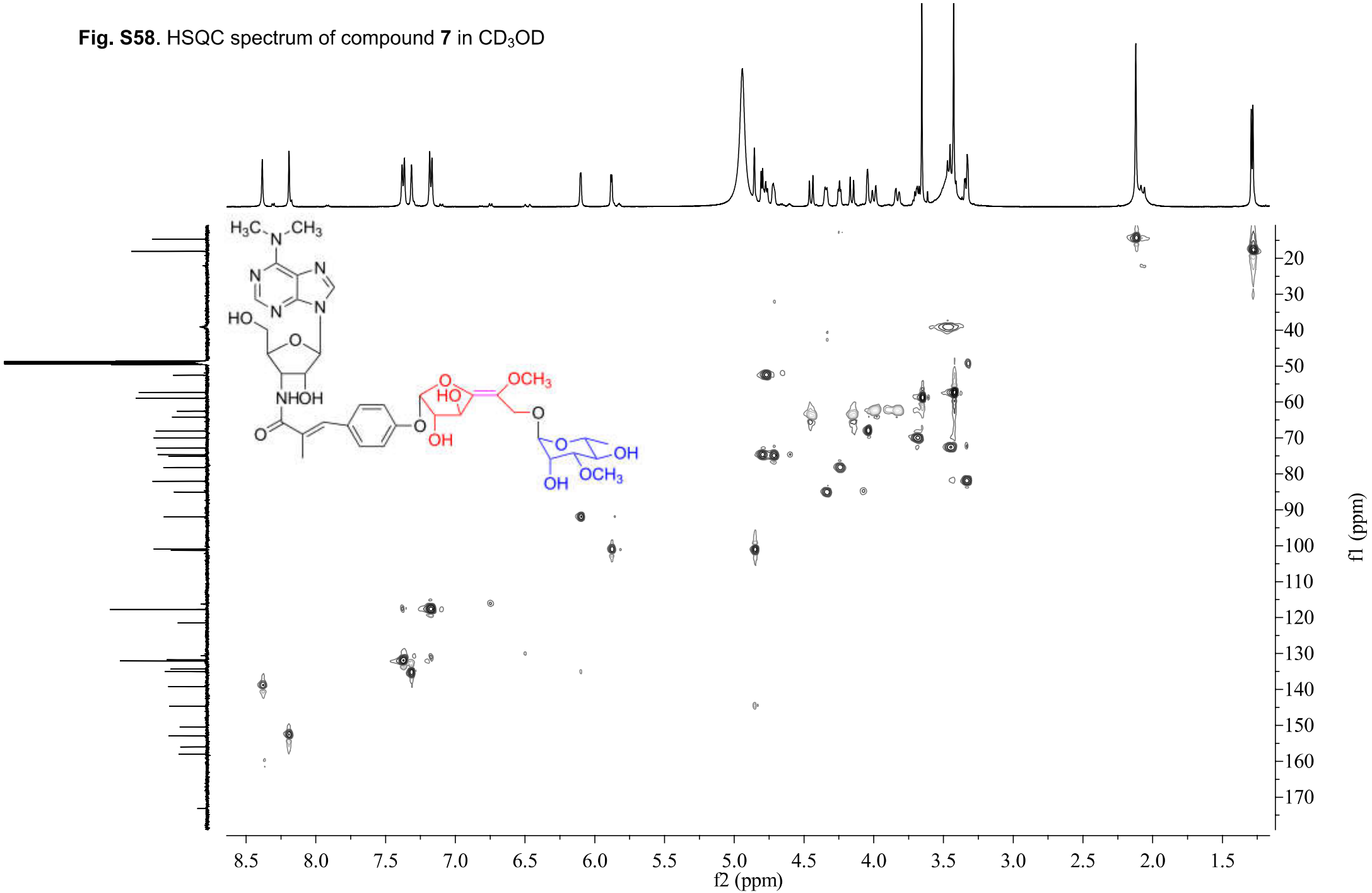


Fig. S59. ^{1}H-^{1}H COSY spectrum of compound **7** in CD_3OD

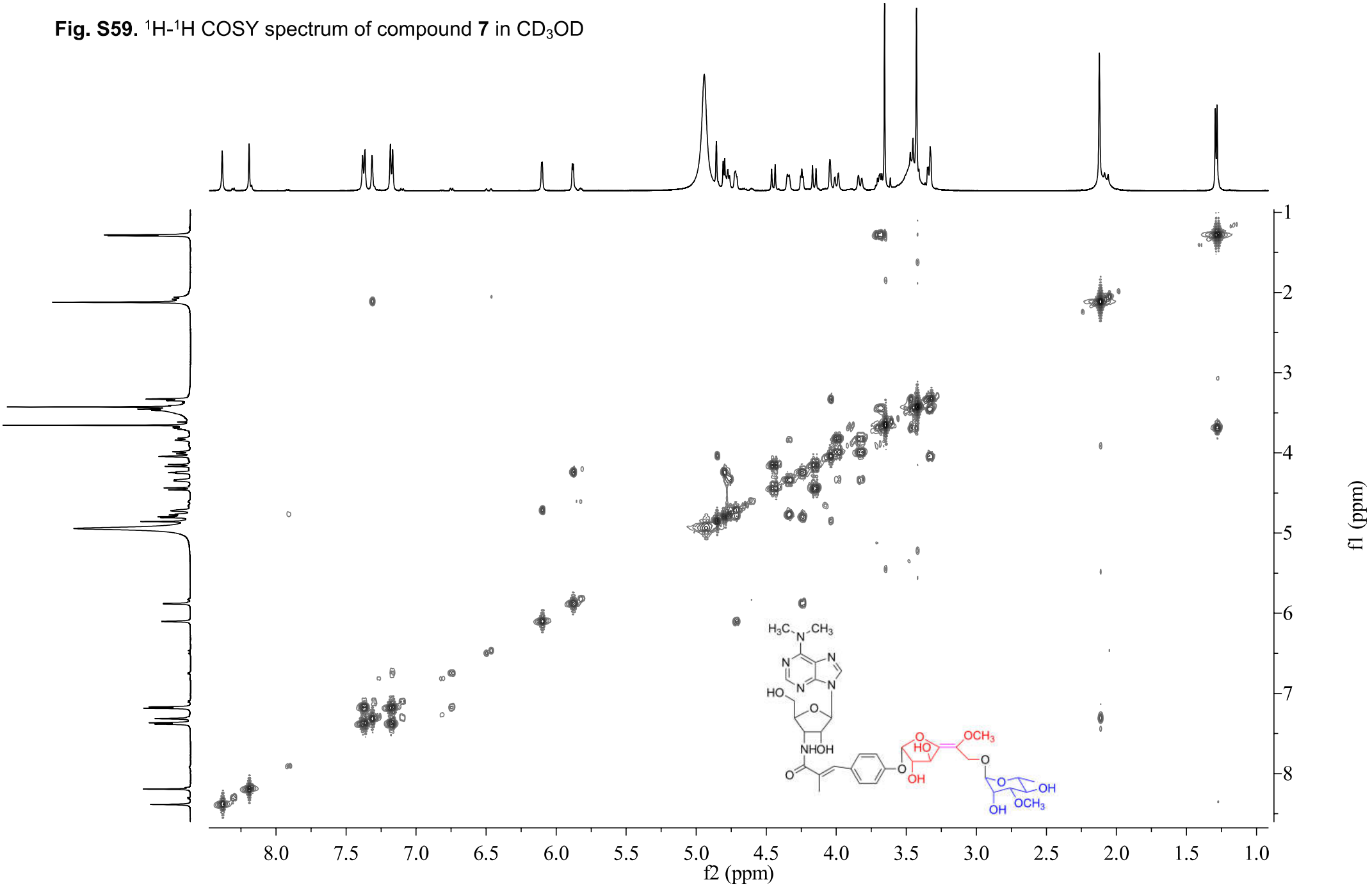


Fig. S60. HMBC spectrum of compound **7** in CD_3OD

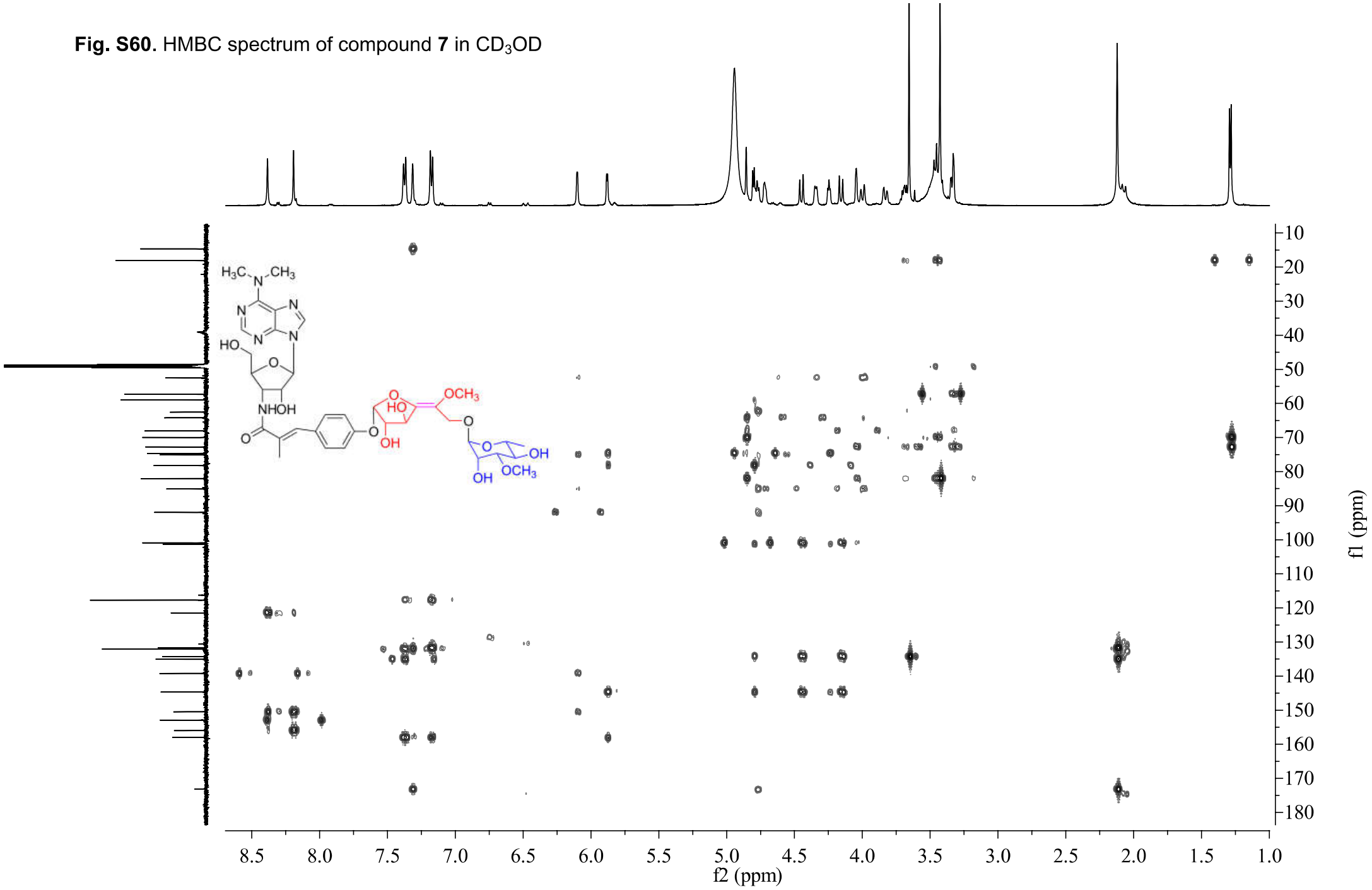


Fig. S61. ^{1}H NMR (500 MHz) spectrum of compound **8** in CD_3OD

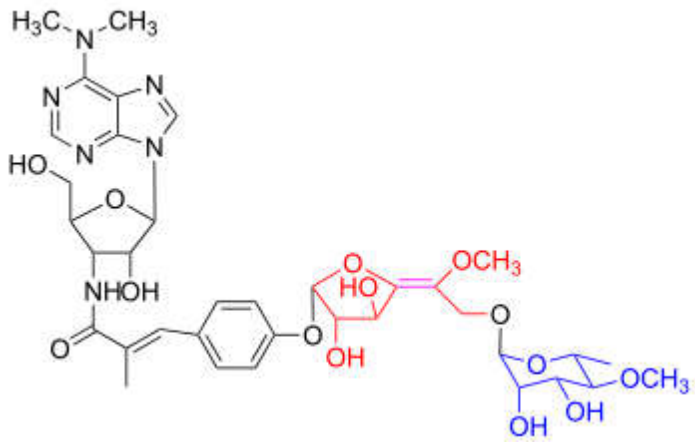


Fig. S62. ^{13}C NMR (125 MHz) spectrum of compound **8** in CD_3OD

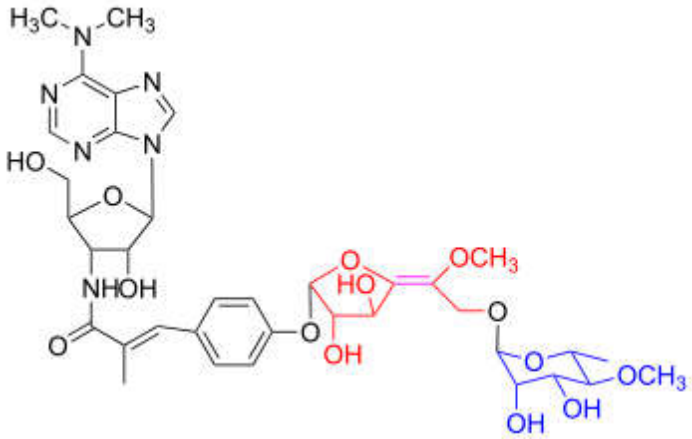


Fig. S63. HSQC spectrum of compound **8** in CD_3OD

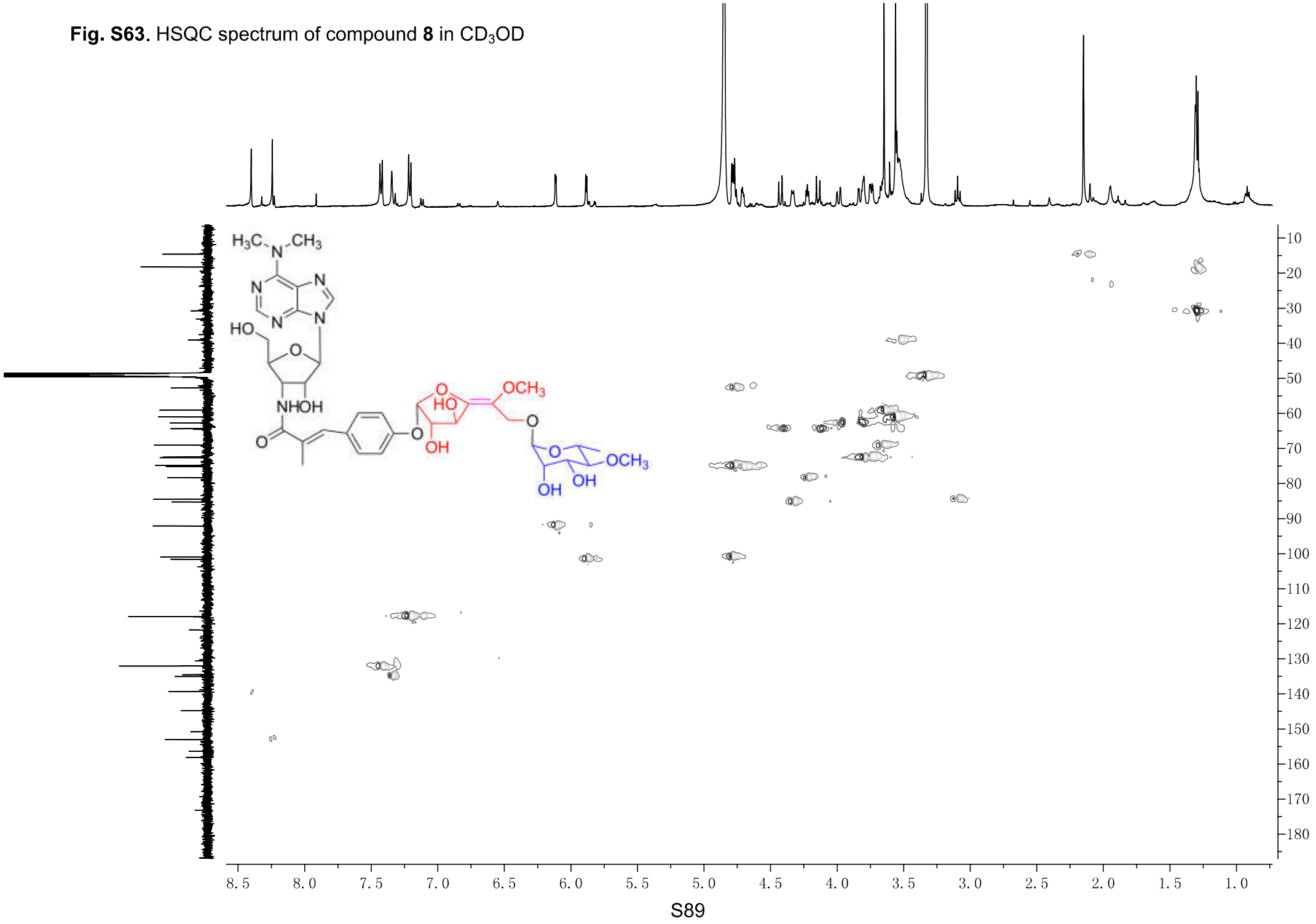


Fig. S64. ^{1}H-^{1}H COSY spectrum of compound **8** in CD_3OD

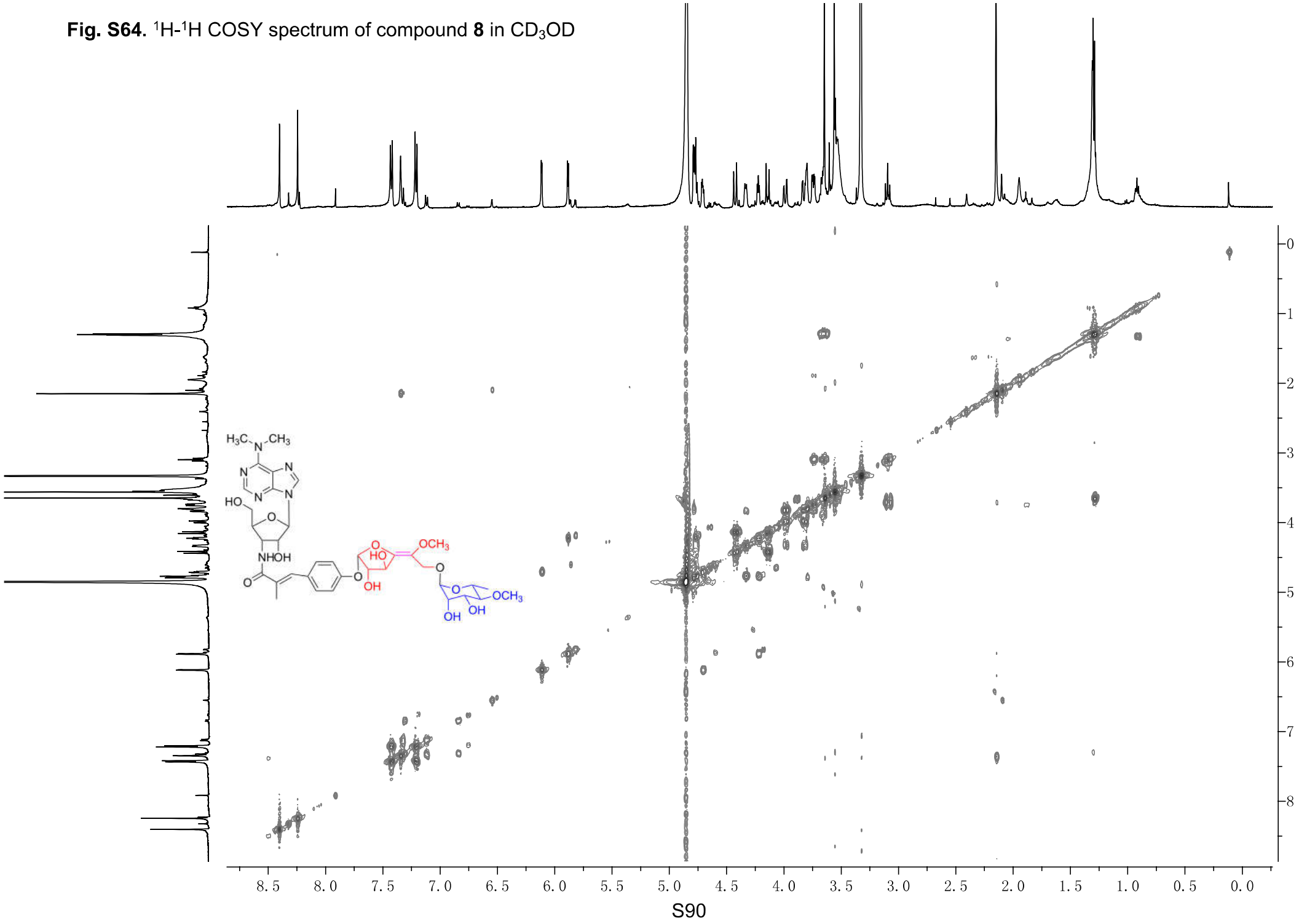


Fig. S65. HMBC spectrum of compound **8** in CD_3OD

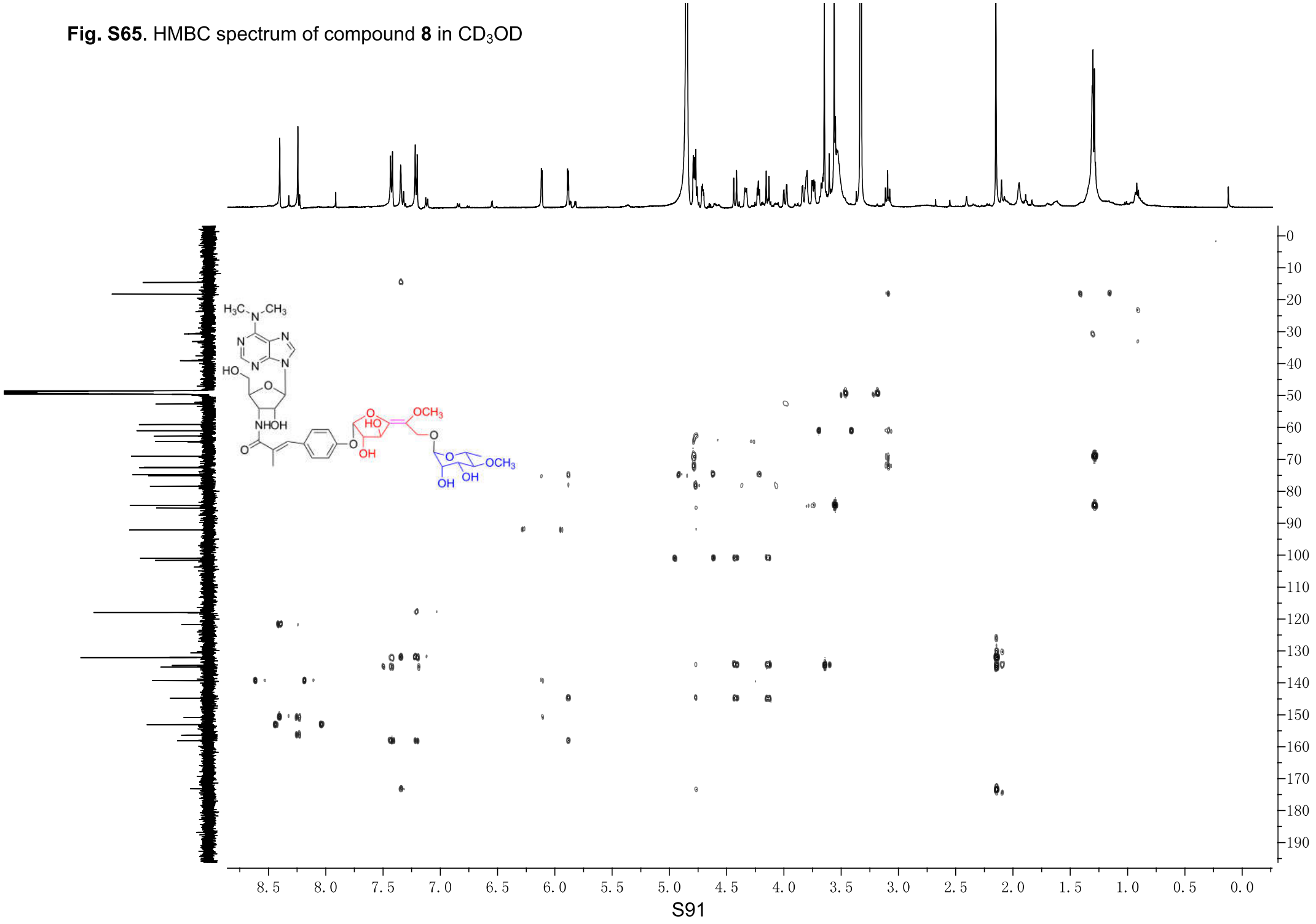


Fig. S66. ^{1}H NMR (500 MHz) spectrum of compound **9** in CD_3OD

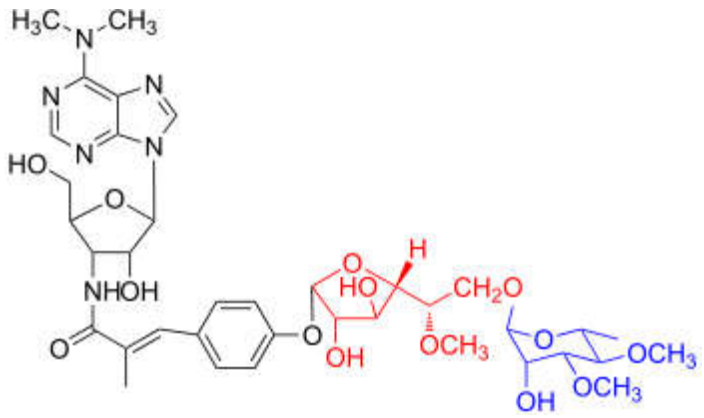


Fig. S67. ^{13}C NMR (125 MHz) spectrum of compound **9** in CD_3OD

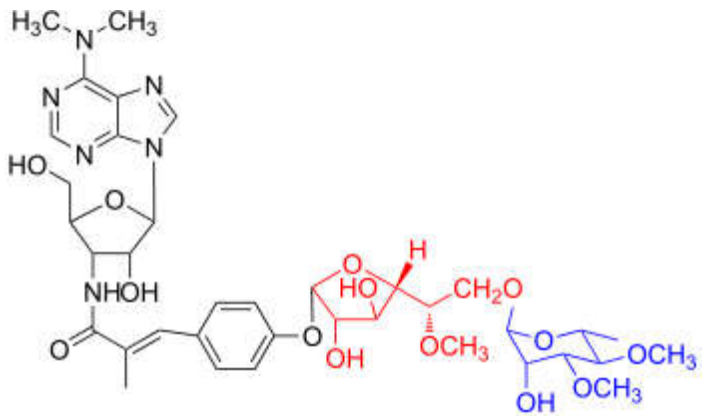


Fig. S68. HSQC spectrum of compound **9** in CD_3OD

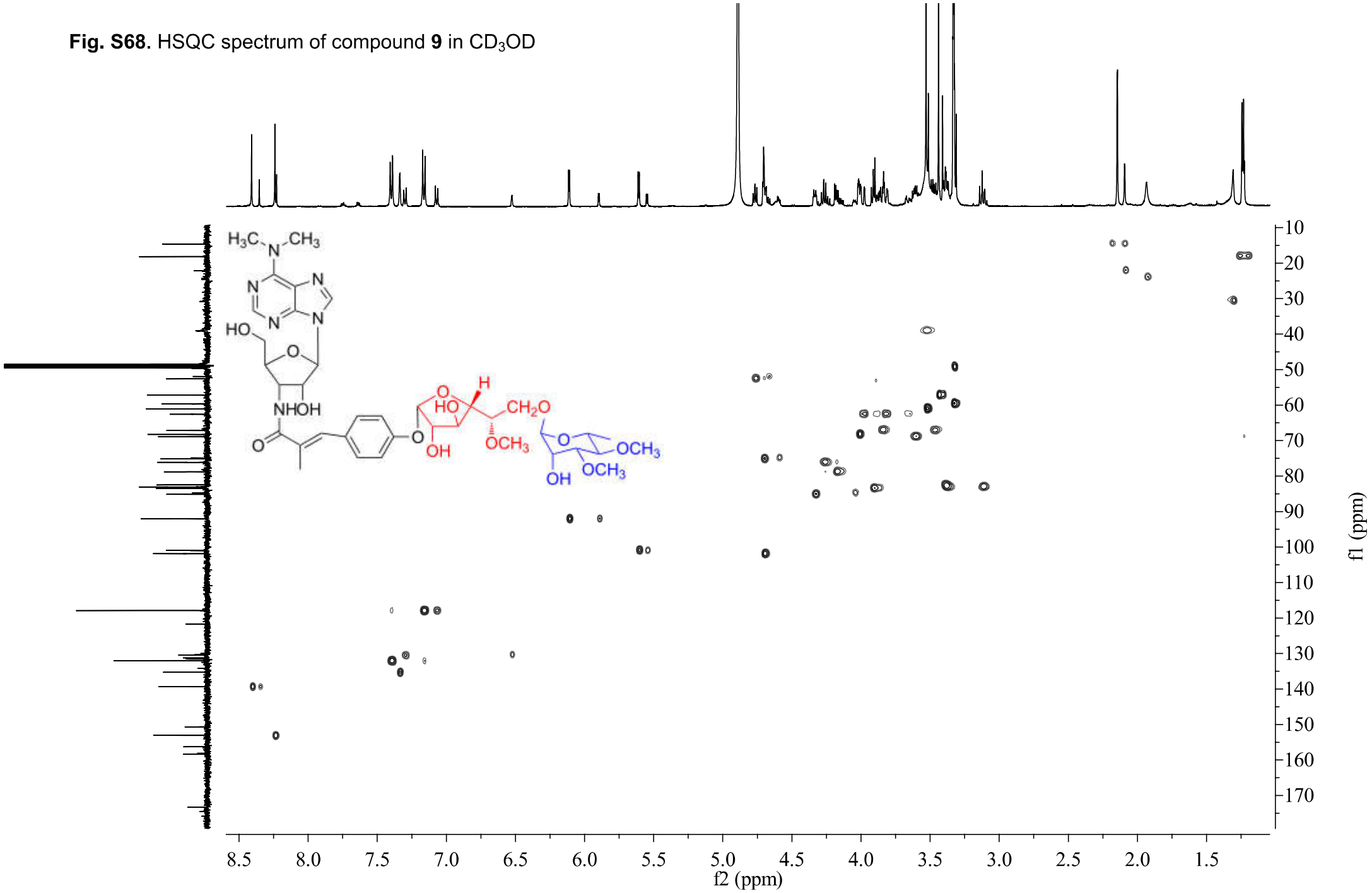


Fig. S69. ^{1}H-^{1}H COSY spectrum of compound **9** in CD_3OD

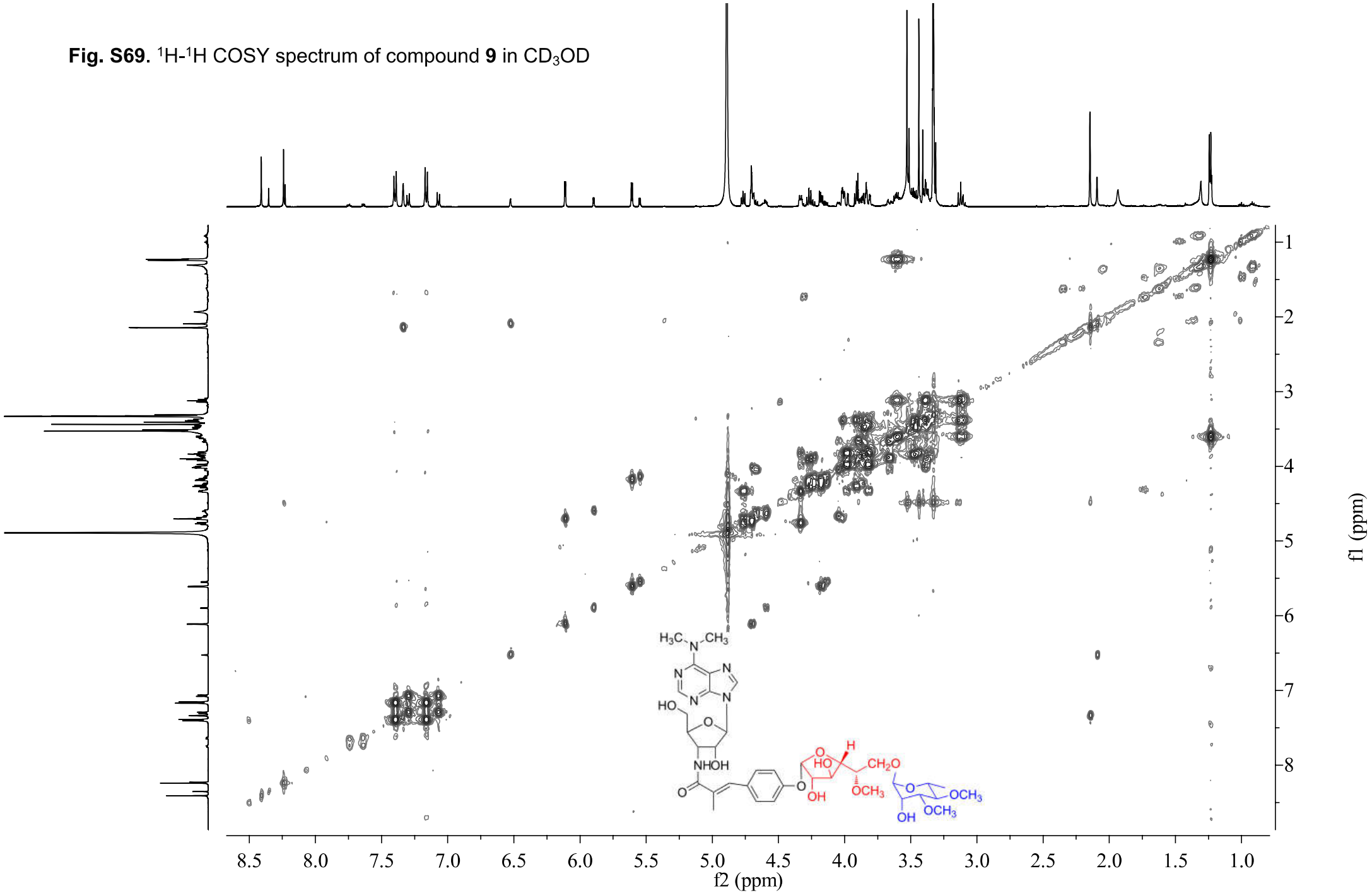


Fig. S70. HMBC spectrum of compound **9** in CD_3OD

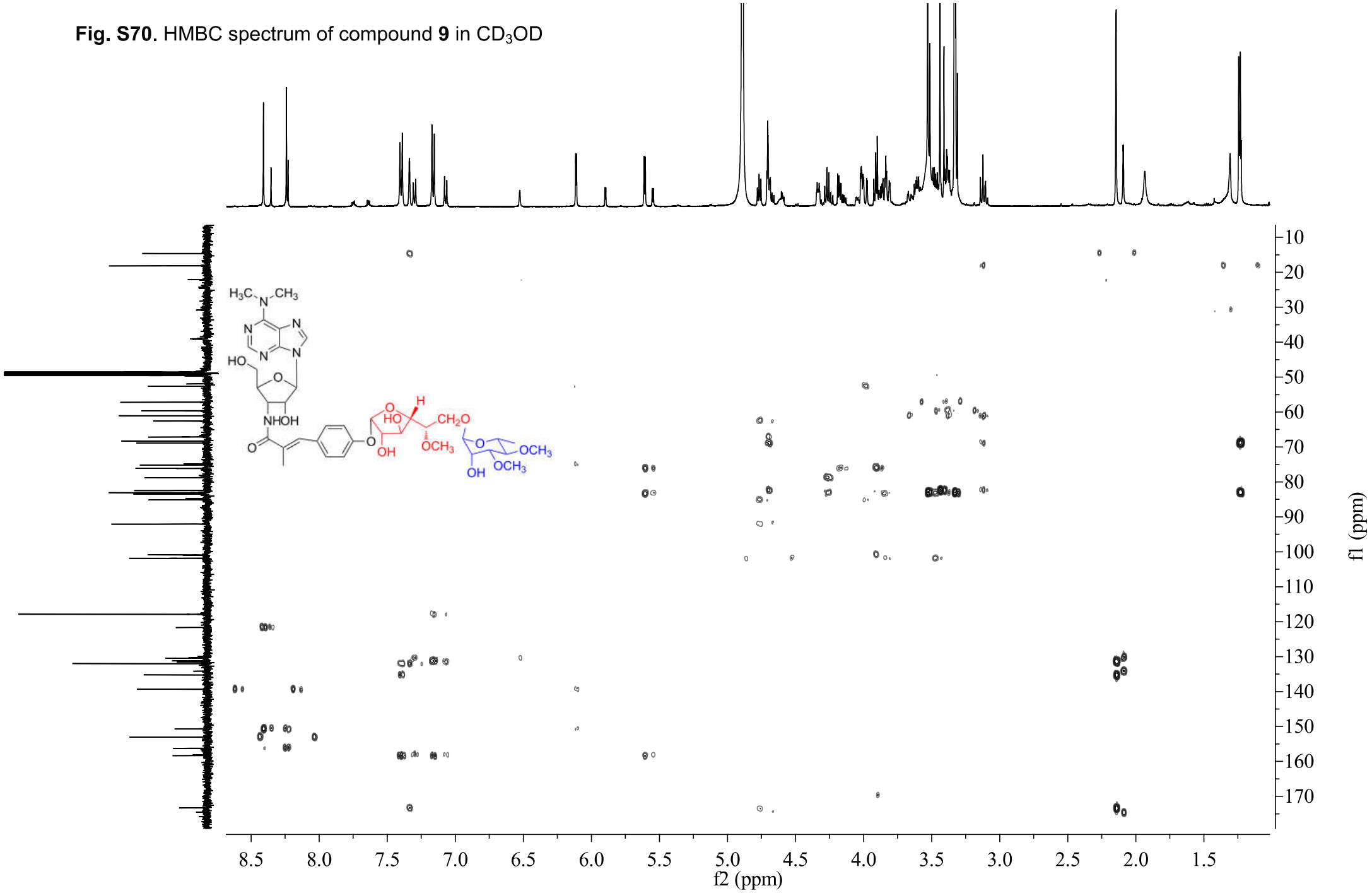


Fig. S71. ^{1}H NMR (500 MHz) spectrum of compound **10** in CD_3OD

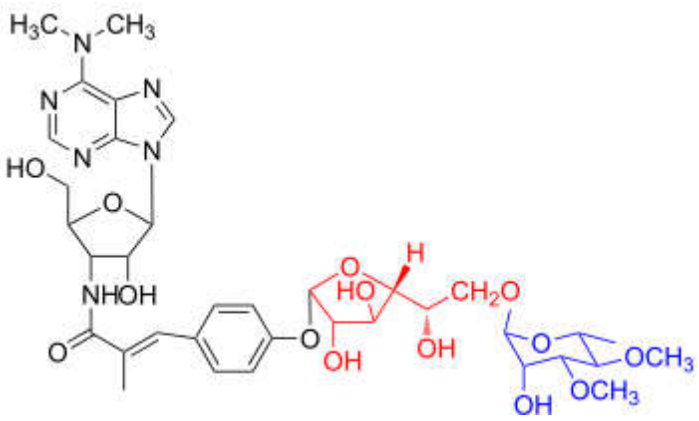


Fig. S72. ^{13}C NMR (125 MHz) spectrum of compound **10** in CD_3OD

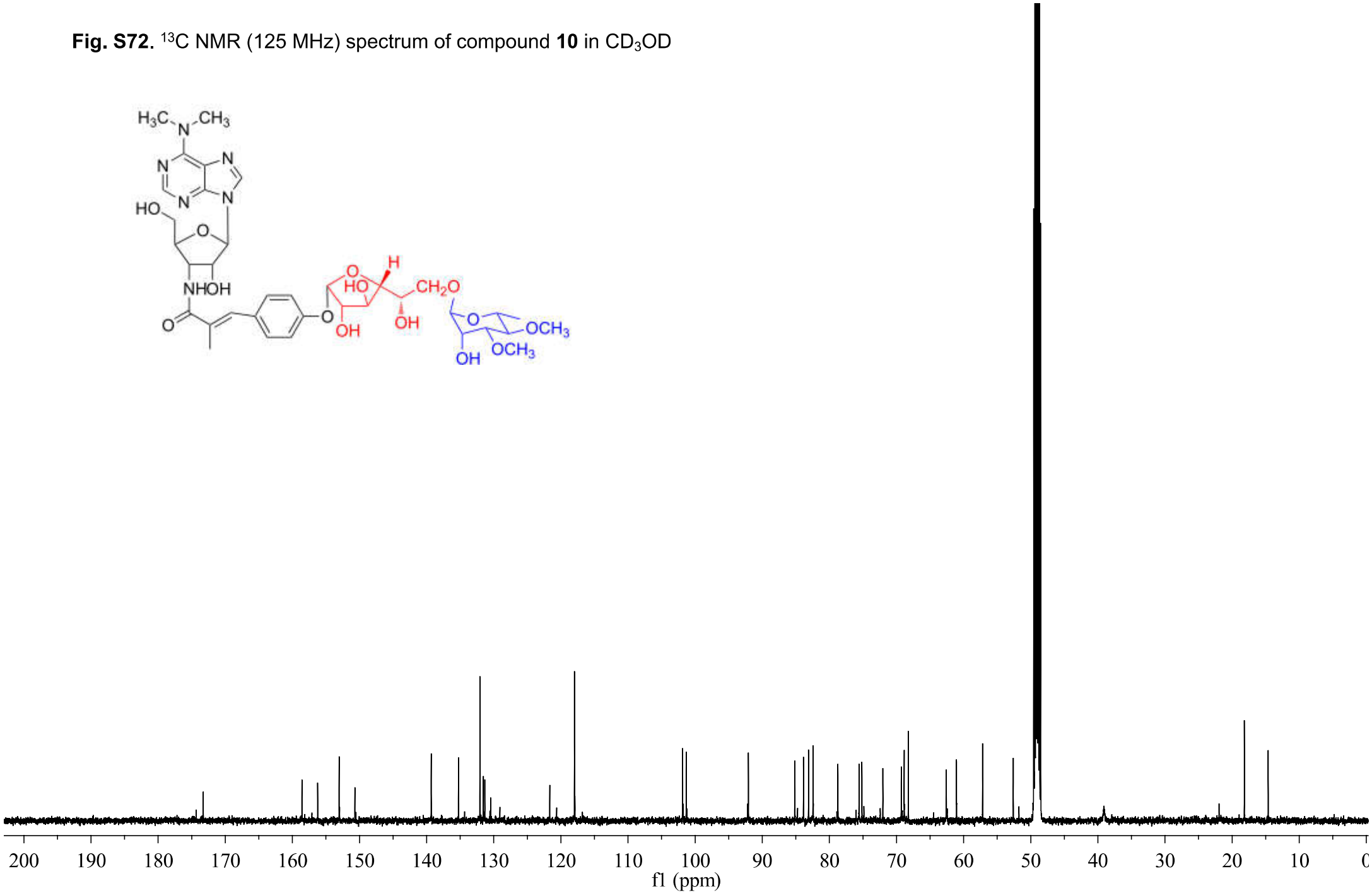


Fig. S73. HSQC spectrum of compound **10** in CD_3OD

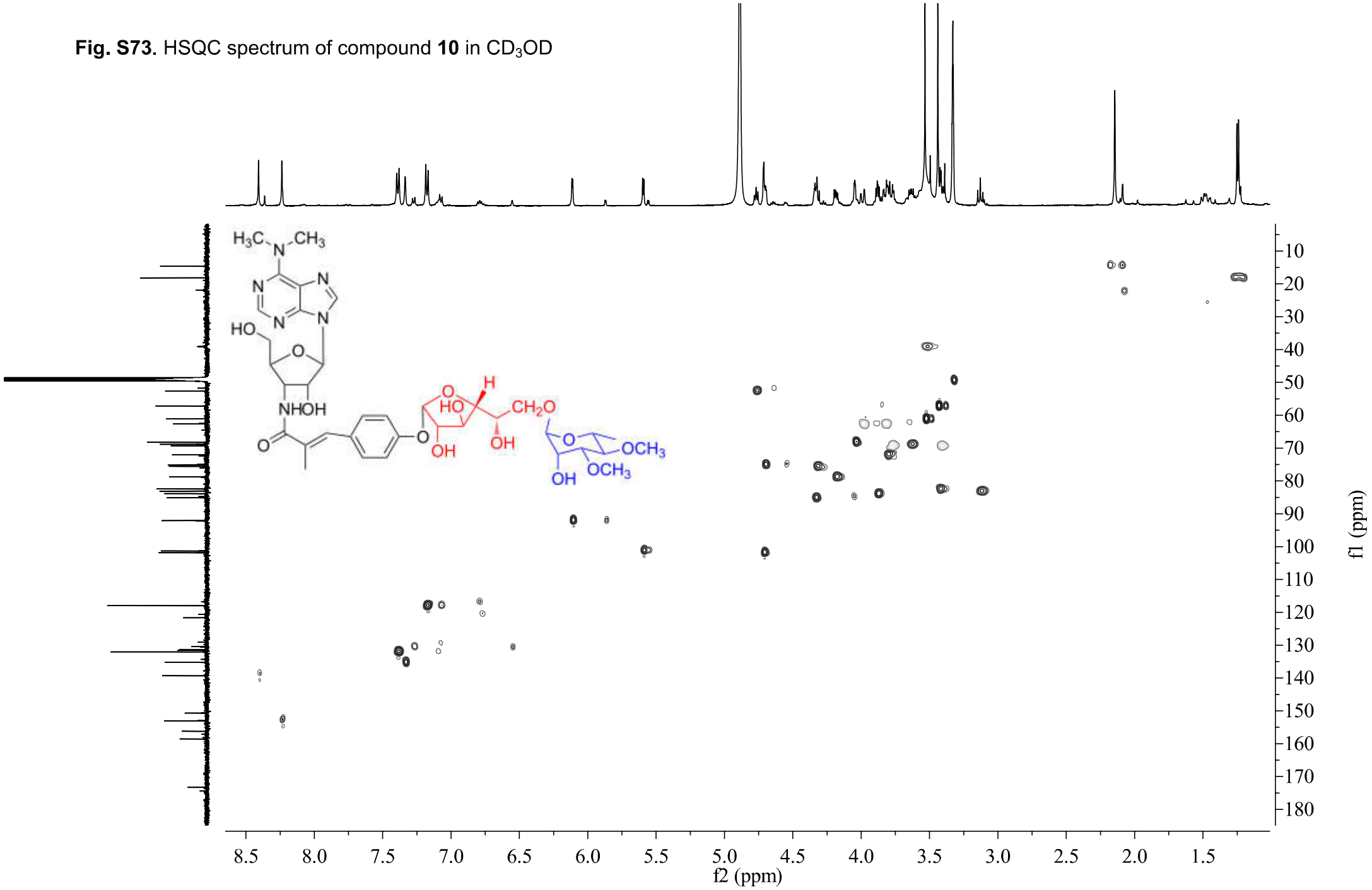


Fig. S74. ^{1}H-^{1}H COSY spectrum of compound **10** in CD_3OD

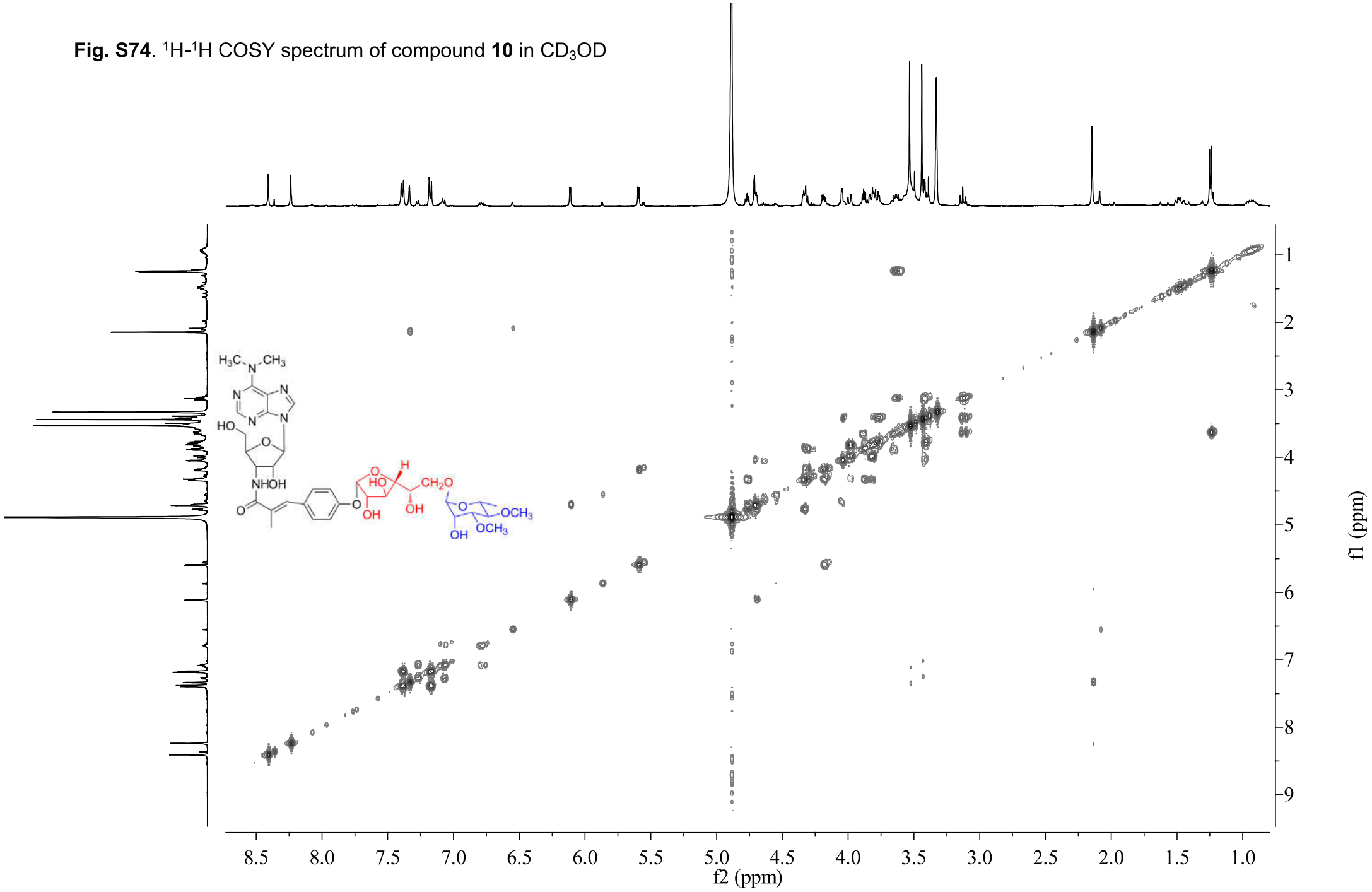


Fig. S75. HMBC spectrum of compound **10** in CD_3OD

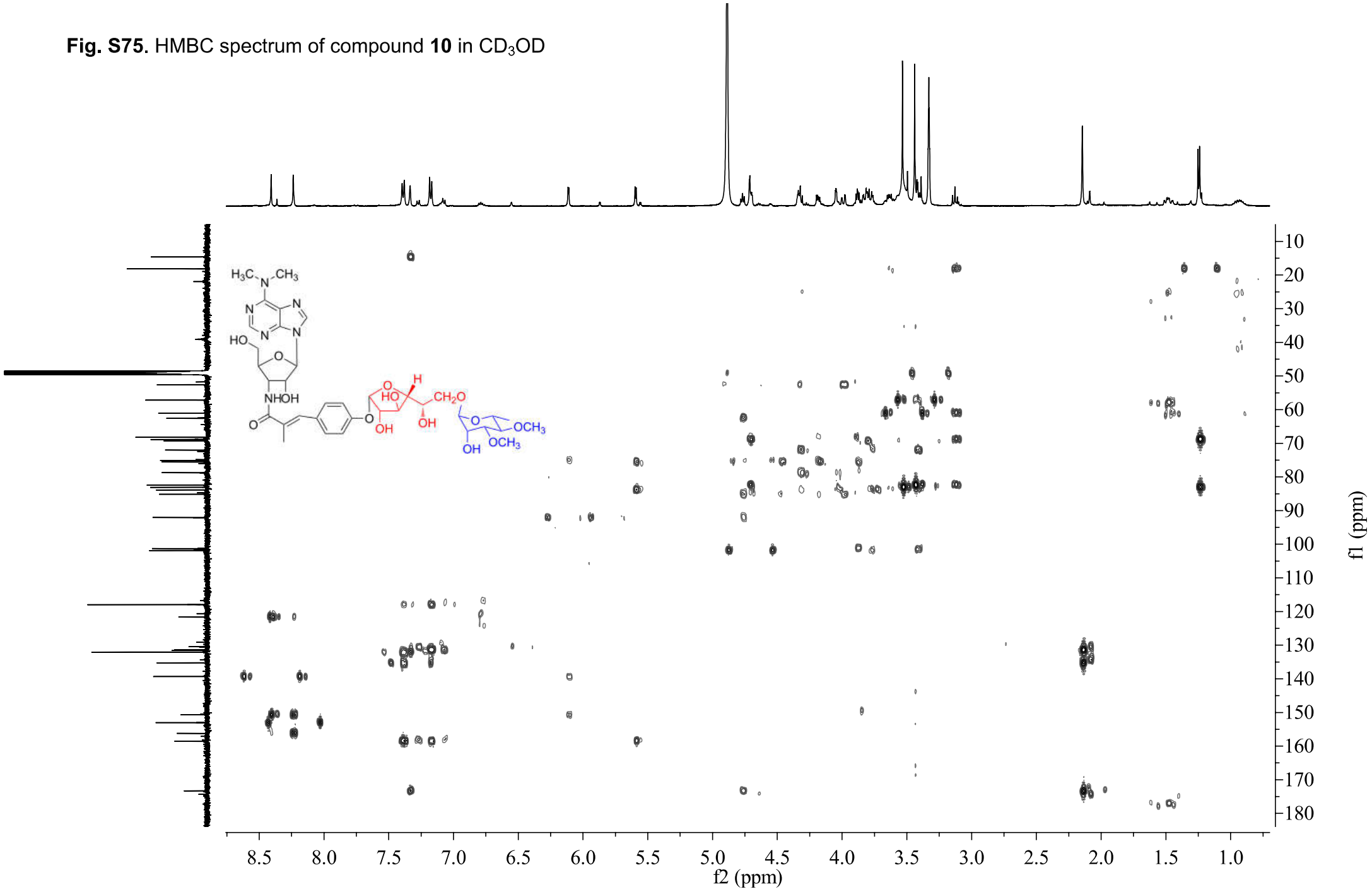


Fig. S76. ^{1}H NMR (500 MHz) spectrum of compound **11** in CD_3OD

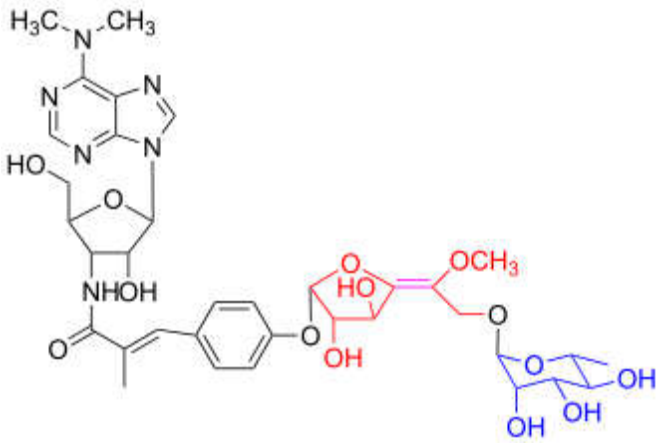


Fig. S77. ^{13}C NMR (125 MHz) spectrum of compound **11** in CD_3OD

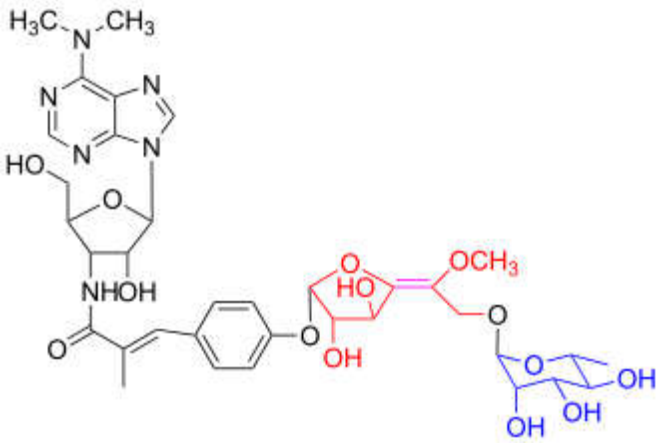


Fig. S78. HSQC spectrum of compound **11** in CD_3OD

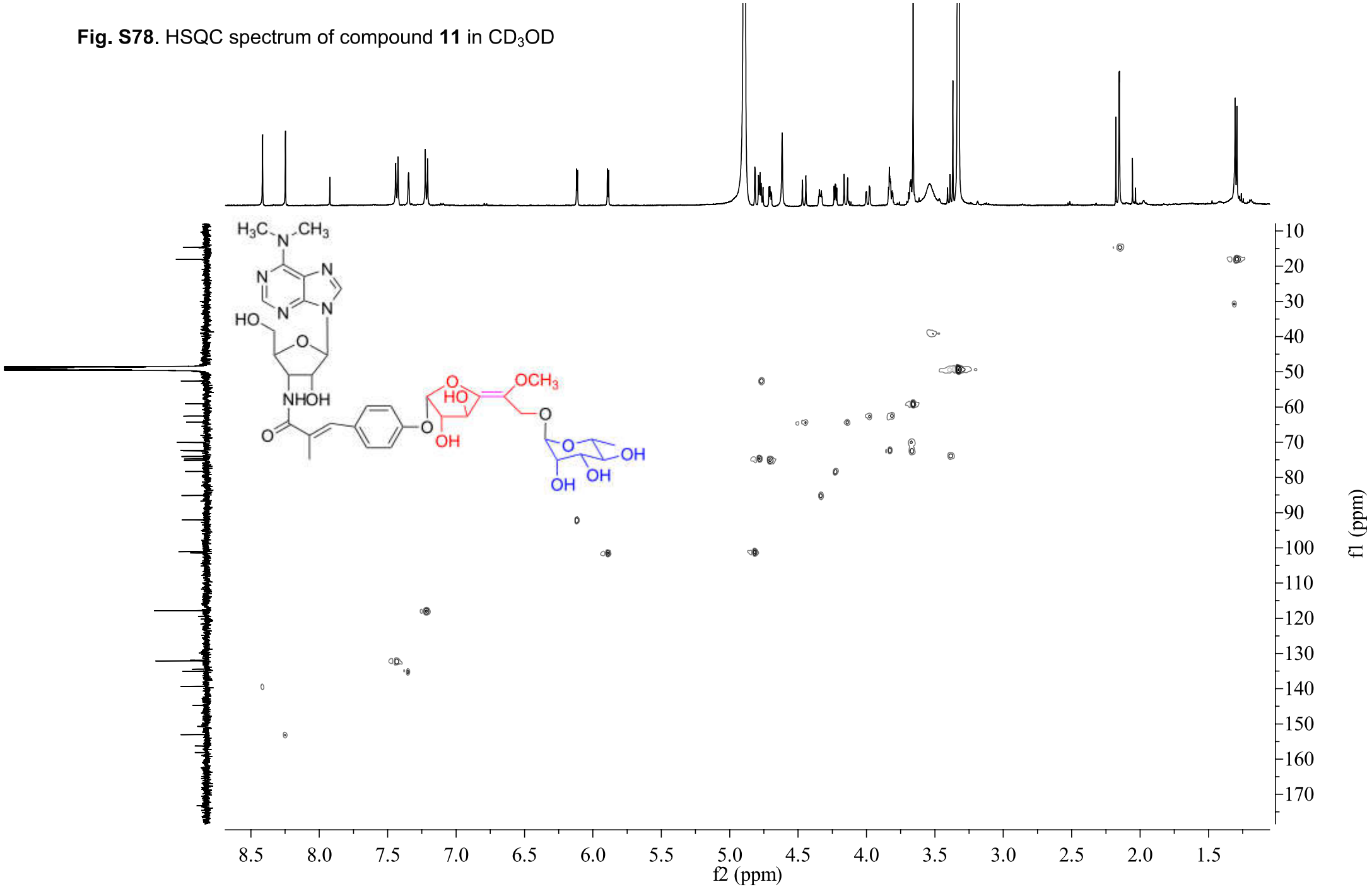


Fig. S79. ^{1}H-^{1}H COSY spectrum of compound **11** in CD_3OD

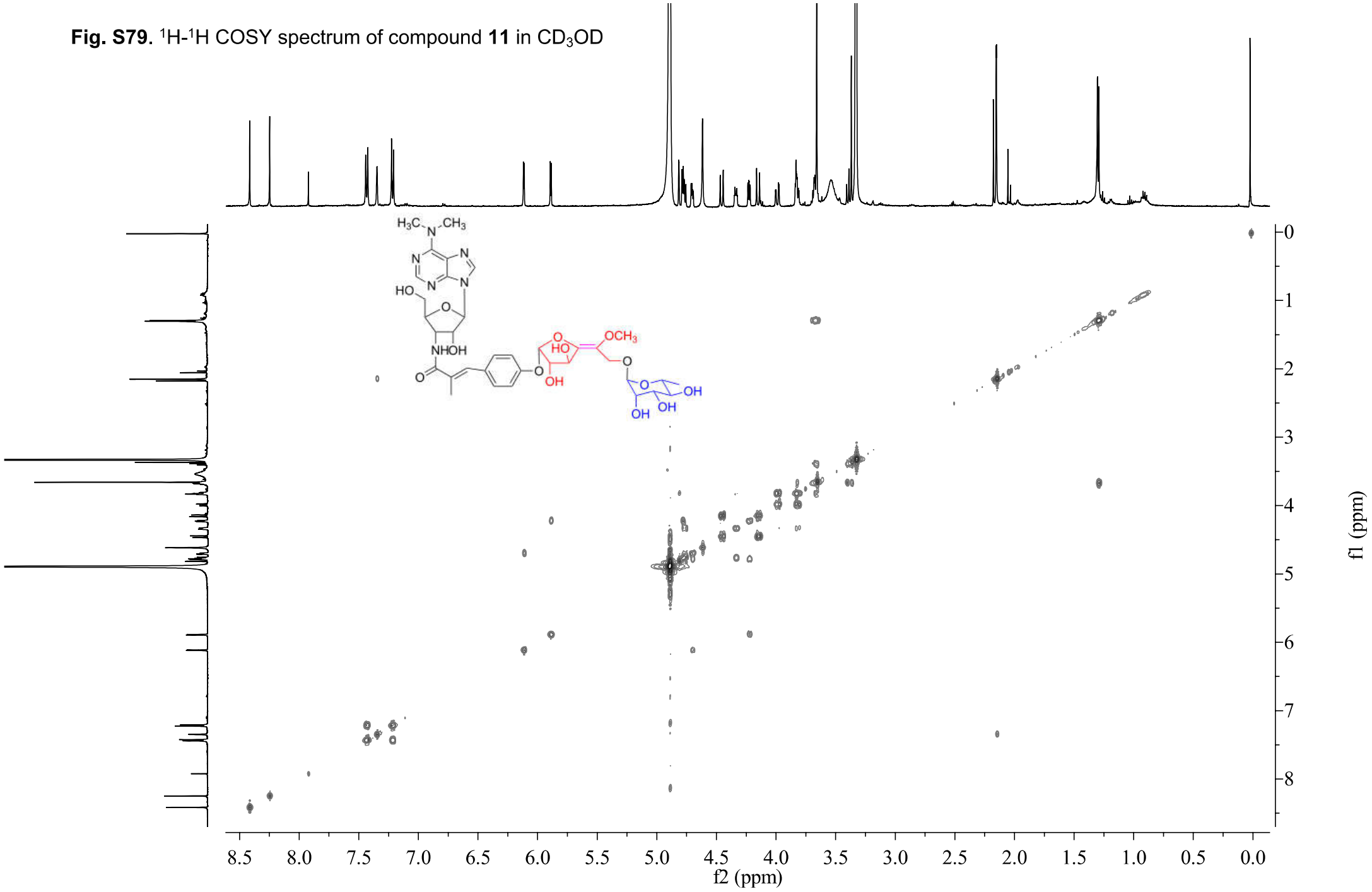


Fig. S80. HMBC spectrum of compound **11** in CD_3OD

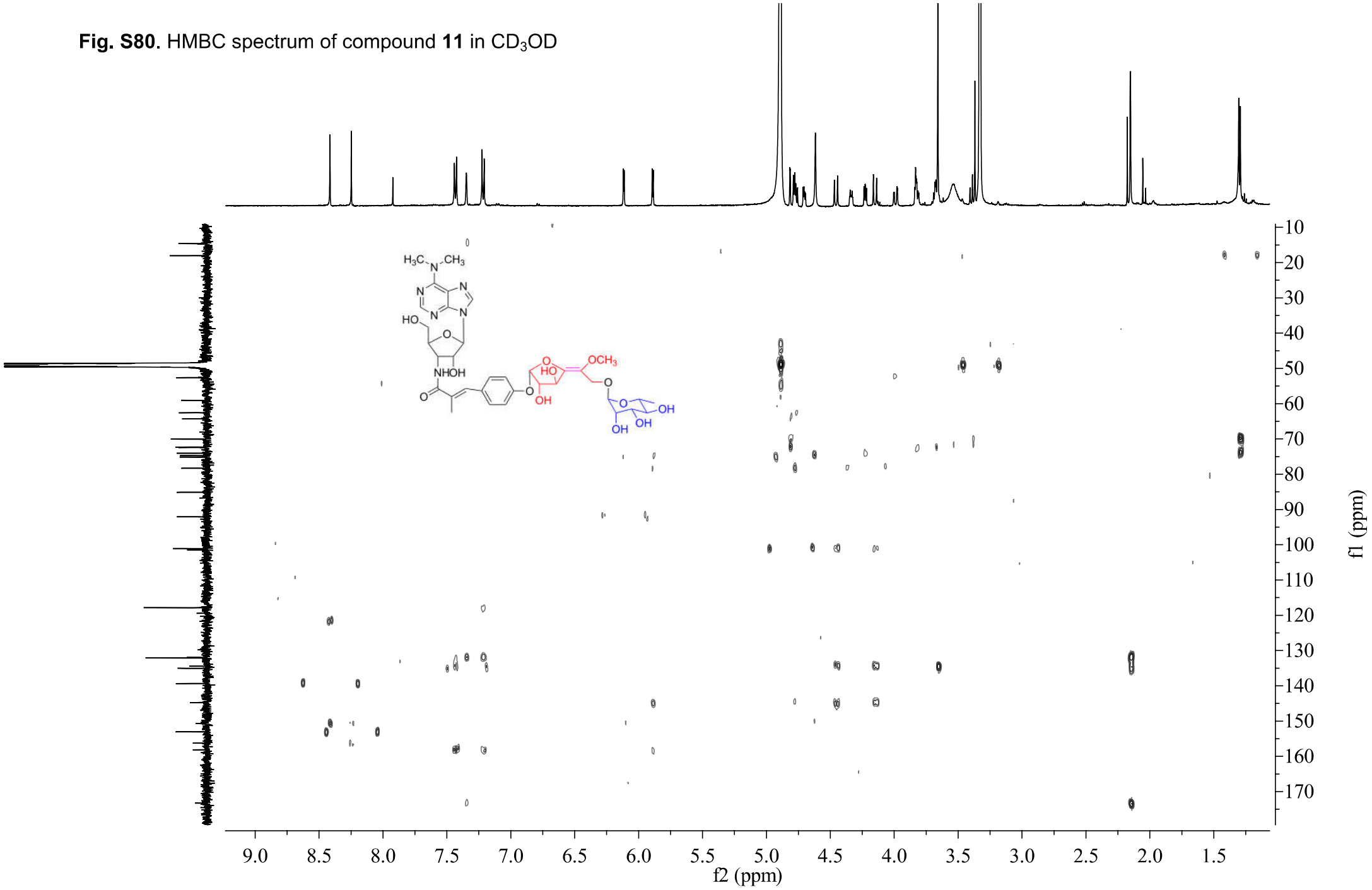


Fig. S81. ^{1}H NMR (500 MHz) spectrum of compound **12** in $CDCl_3$

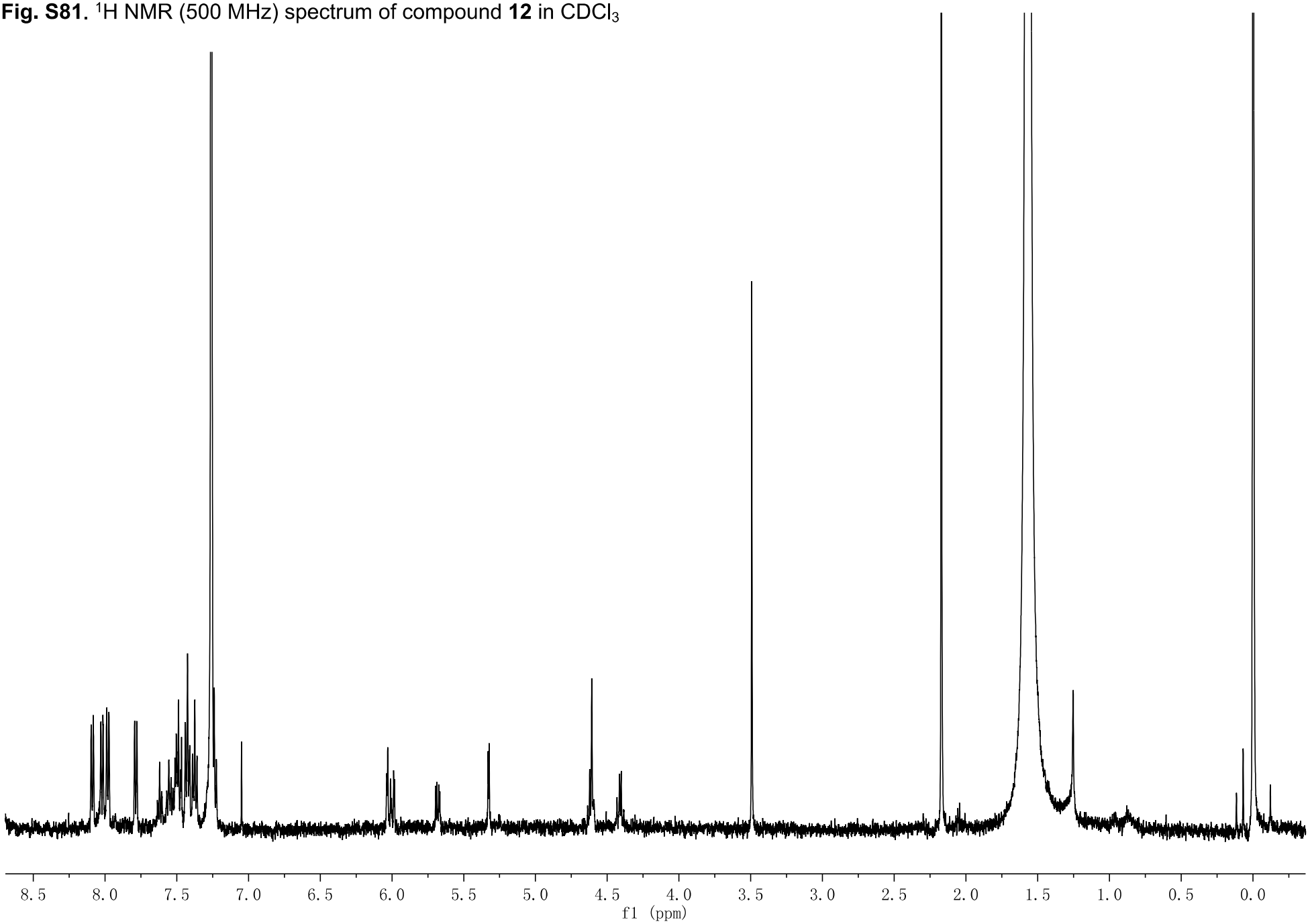


Fig. S82. ^{1}H NMR (500 MHz) spectrum of compound **13** in $CDCl_3$

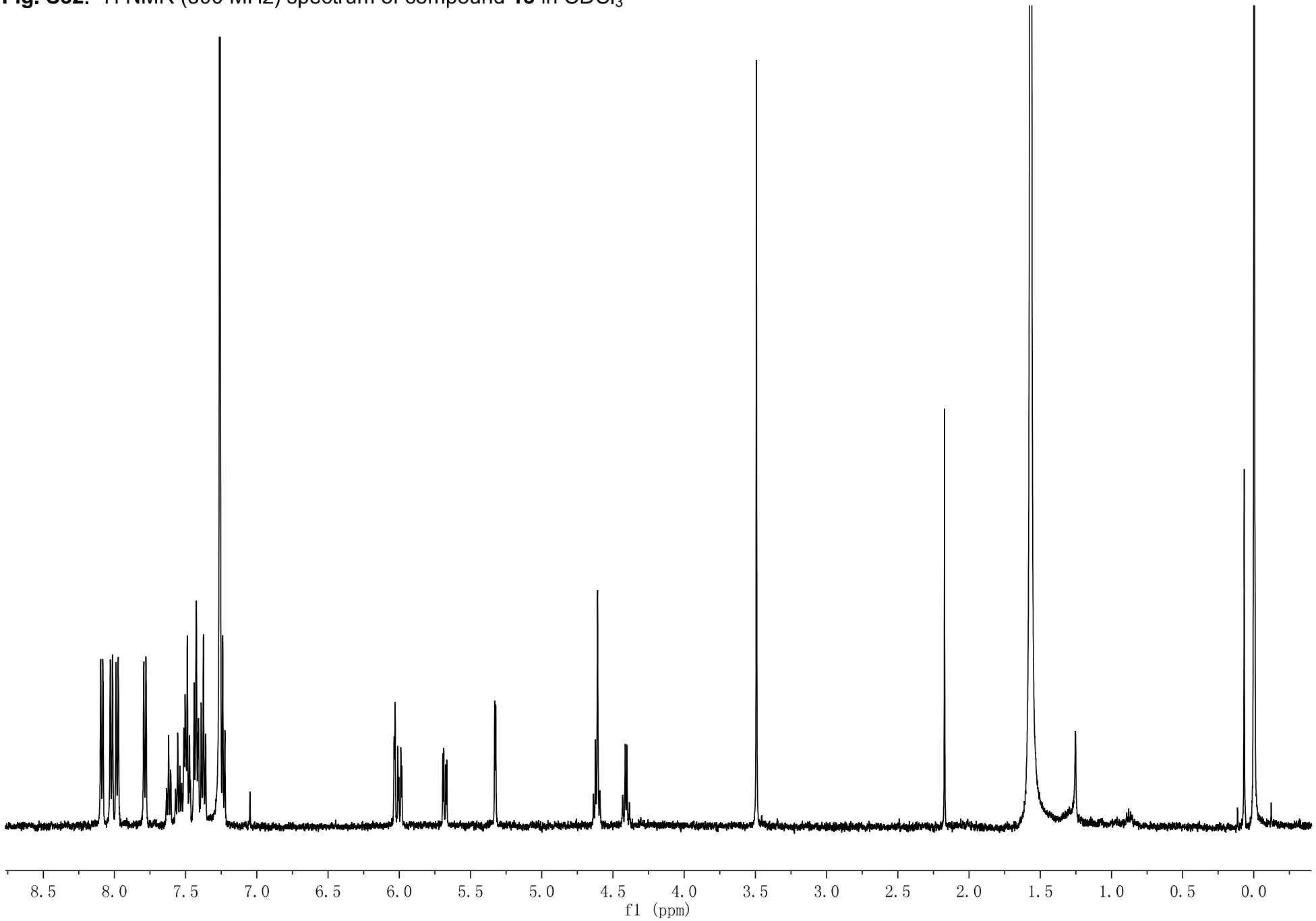


Fig. S83. ^{1}H NMR (500 MHz) spectrum of compound **14** in $CDCl_3$

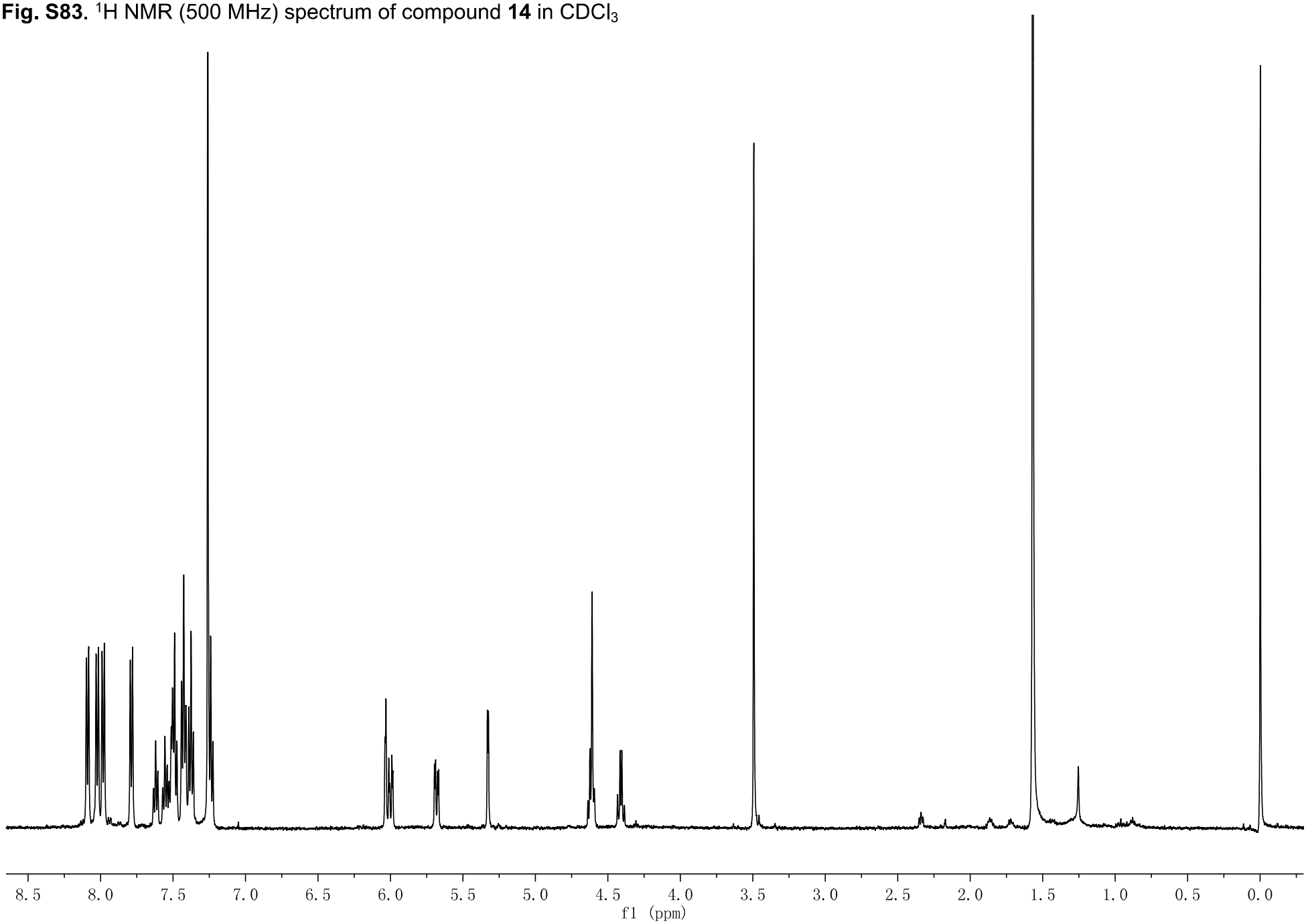


Fig. S84. ^{13}C NMR (125MHz) spectrum of compound **1** ([1-^{13}C] D-mannose-derived) in CD_3OD

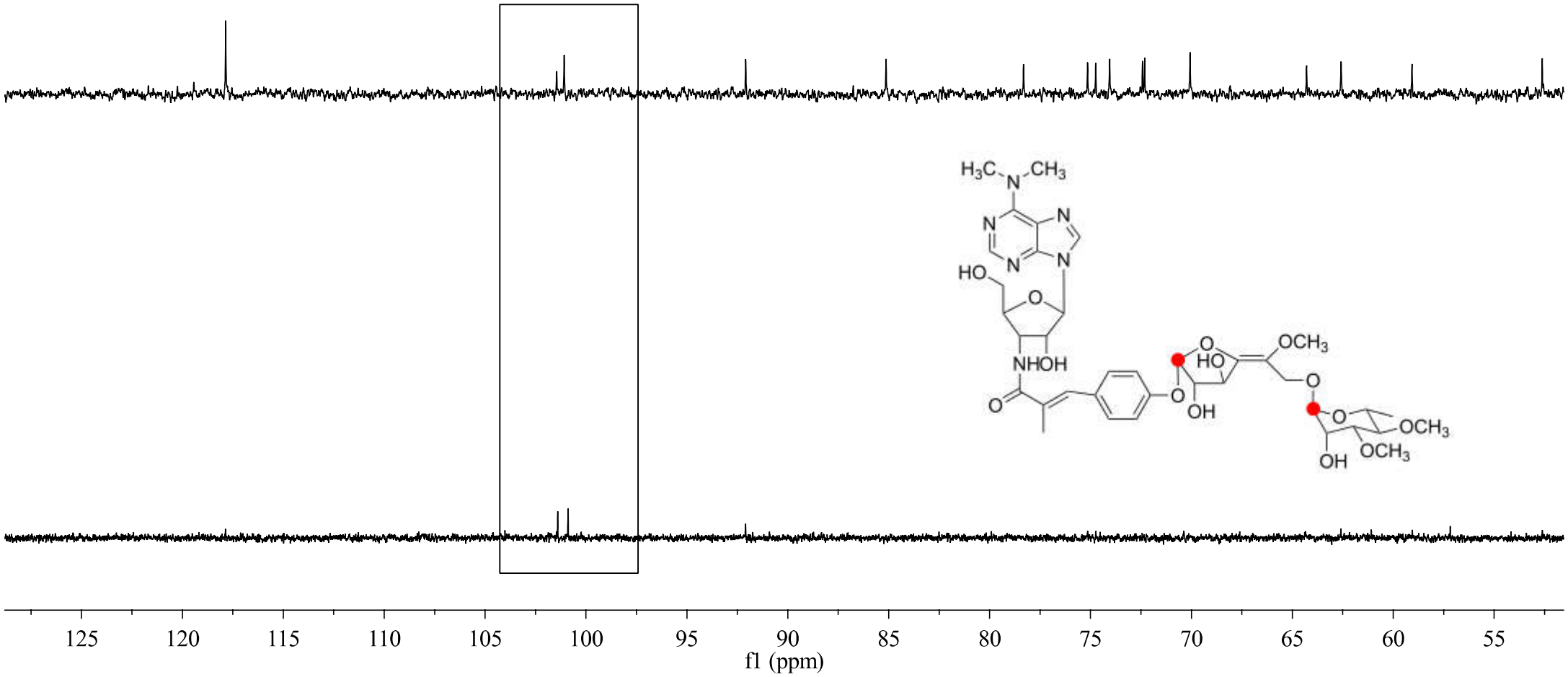


Fig. S85. ^{13}C NMR (125MHz) spectrum of compound **1** ([1-^{13}C] D-mannose-derived) in CD_3OD

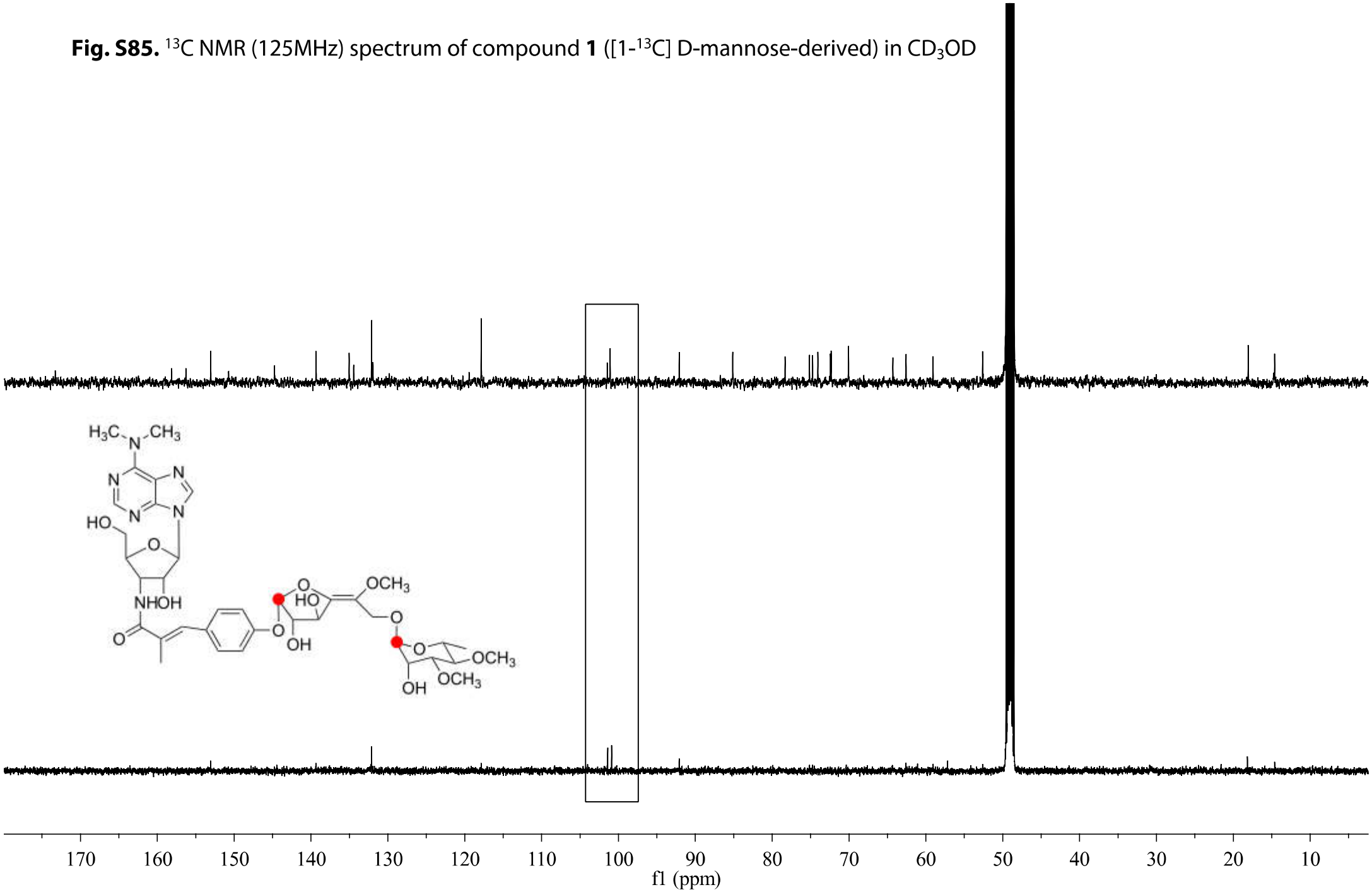


Fig. S86. ^{1}H NMR (500 MHz) spectra of MtdL-mediated enzymatic mixture and GDP-β-L-galactose in D_2O

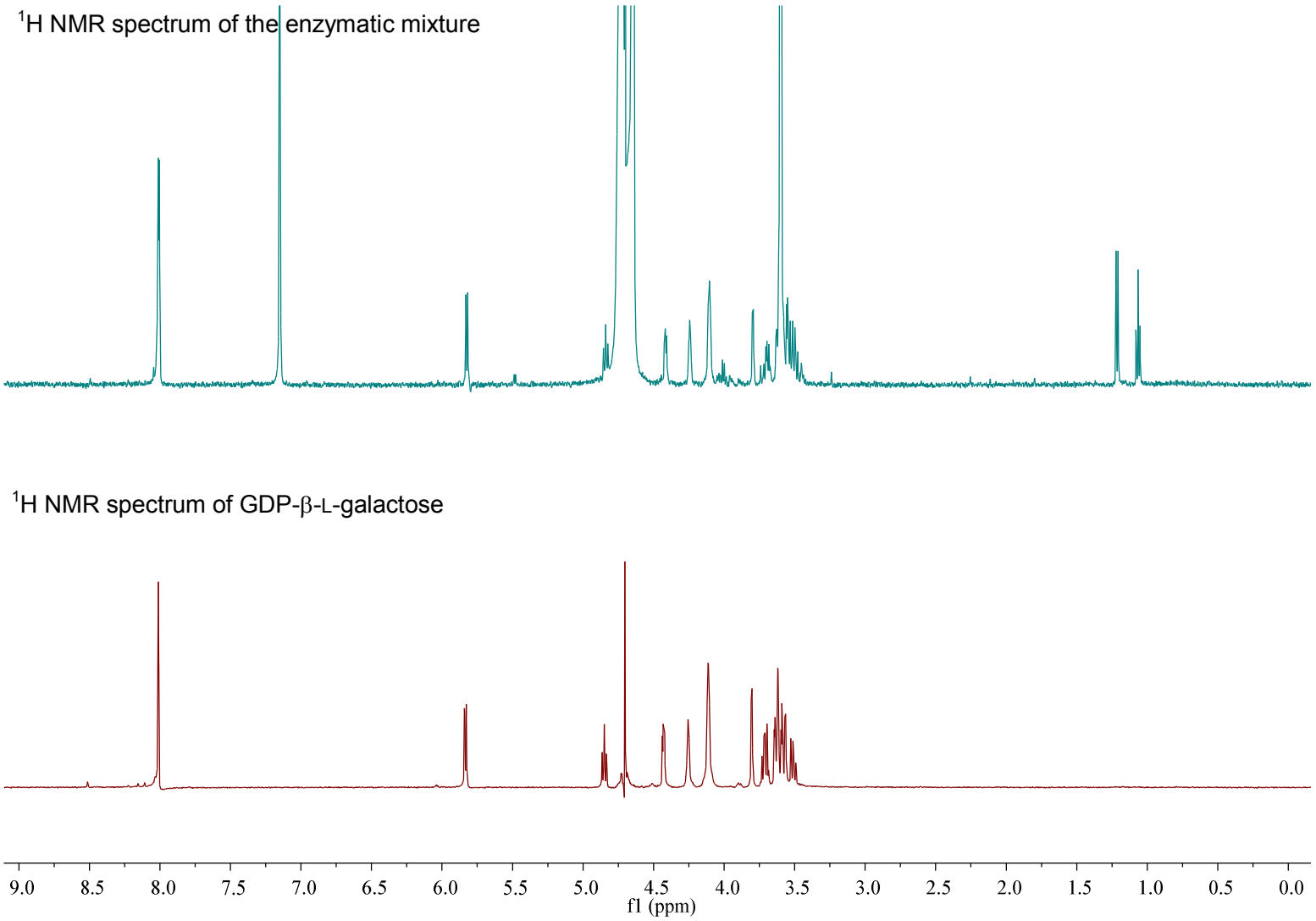
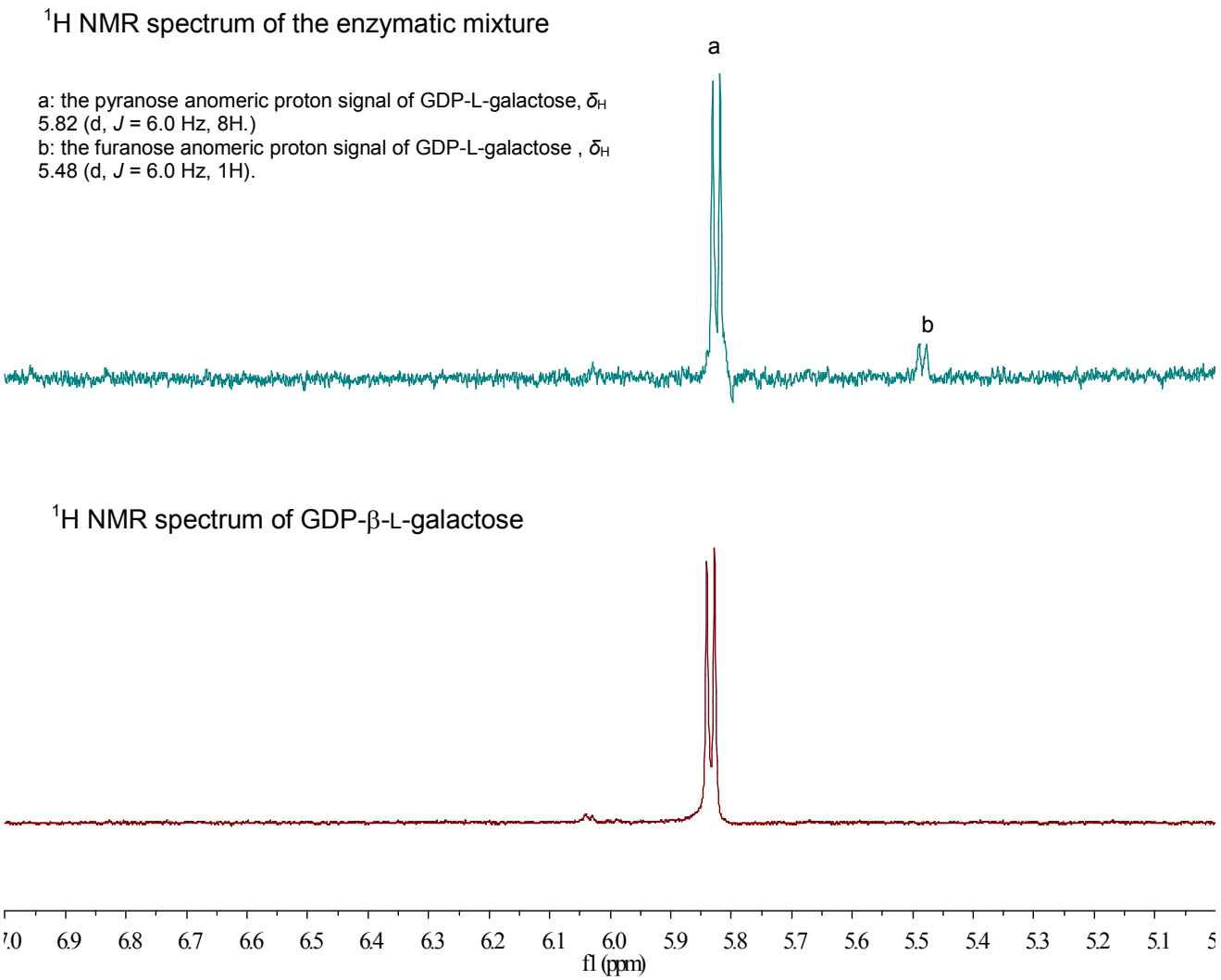


Fig. S87. ^{1}H NMR (500 MHz) spectra of MtdL-mediated enzymatic mixture and GDP-β-L-galactose in D_2O



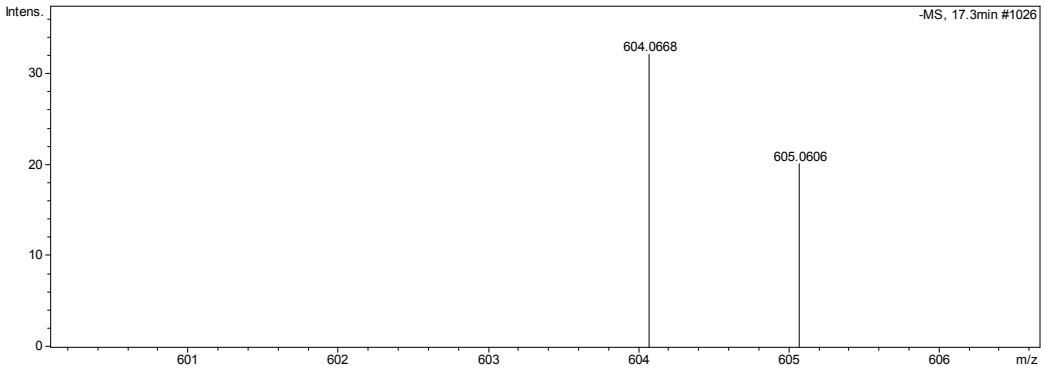


Fig. S88. The LC-MS analysis of MtdL-mediated enzymatic mixture.
([M - H]$^-$ =604.0668, calcd for $C_{16}H_{24}N_5O_{16}P_2^-$, 604.0699, err 5.0 ppm)

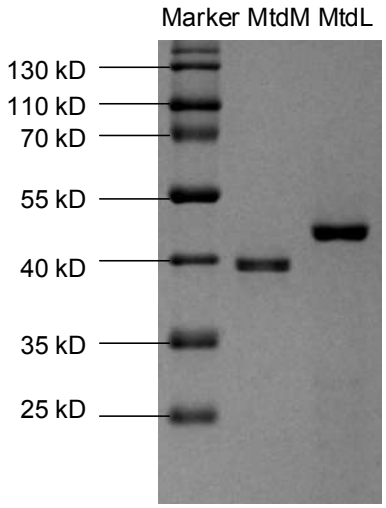


Fig. S89. SDS-PAGE for MtdM and MtdL

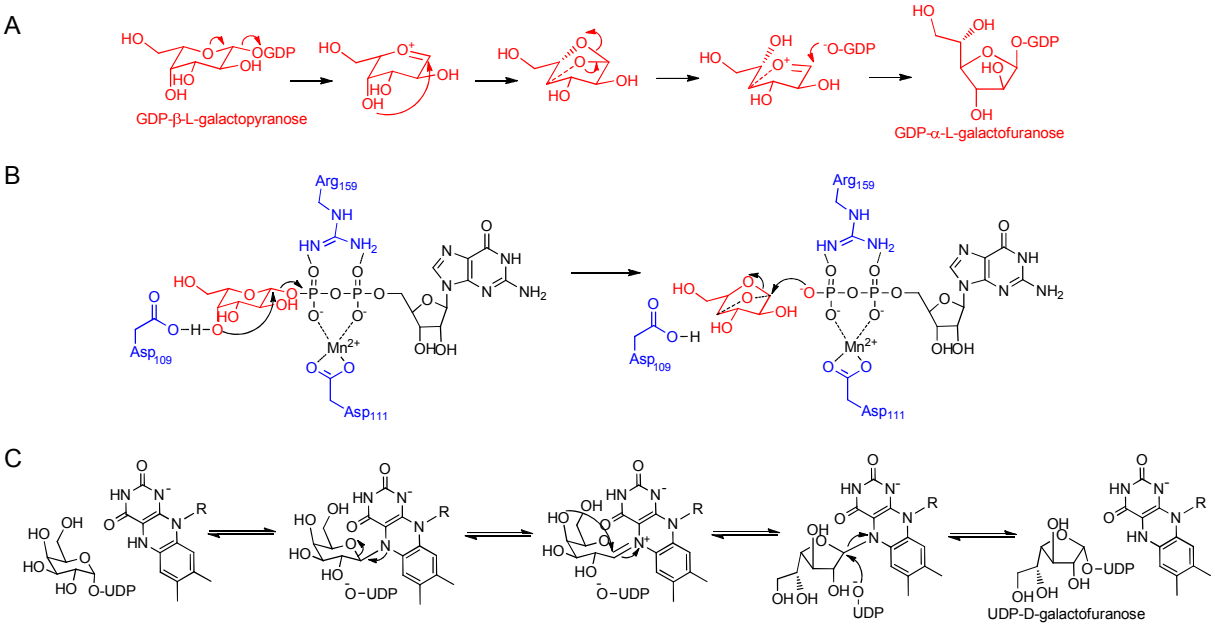


Fig. S90. Plausible chemical mechanisms for MtdL and UGM. (A) Mechanism involving bond cleavage and formation *en route* to GDP-α-L-galactofuranose. (B) MtdL mediates flavin-independent GDP-β-L-galactose pyranose–furanose transformation. (C) UGM mediates flavin-dependent UDP-Gal*p* and UDP-Gal*f* transformation. Nucleophilic attack by the reduced flavin leads to a flavin-galactose adduct. Sugar ring contraction occurs by attack of the C_4 hydroxyl to the C_1-carbon.

Fig. S91. Phylogenetic tree of MtdL with its homologues. The amino acid sequences were aligned using ClustalW and the phylogenetic tree was generated using Molecular Evolutionary Genetics Analysis (MEGA) 6.0.

Supplementary References

(1) Zhu Q, et al. (**2012**) Discovery and engineered overproduction of antimicrobial nucleoside antibiotic A201A from the deep-sea marine actinomycete *Marinactinospora thermotolerans* SCSIO 00652. *Antimicrob Agents Chemother* 56(1): 110–114.

(2) Ma J, et al. (**2012**) Characterization of a single gene cluster responsible for methylpendolmycin and pendolmycin biosynthesis in the deep sea bacterium *Marinactinospora thermotolerans*. *Chembiochem* 13(4): 547–552.

(3) Zhang Y, et al. (**2013**) Identification of the grincamycin gene cluster unveils divergent roles for GcnQ in different hosts, tailoring the L-rhodinose moiety. *Org Lett* 15(13): 3254–3257.

(4) Gust B, Kieser T, Chater KF (**2002**) REDIRECT technology: PCR-targeting system in *Streptomyces coelicolor*. The John Innes Centre, Norwich, United Kingdom.

(5) Kirst HA, et al. (**1985**) The structure of A201A, a novel nucleoside antibiotic. *J Antibiot (Tokyo)* 38(5): 575–586.

(6) Costantino V, et al. (**2008**) Corrugoside, a new immunostimulatory α-galactoglycosphingolipid from the marine sponge *Axinella corrugata*. *Bioorg Med Chem* 16(4): 2077–2085.

(7) Costantino V, Fattorusso E, Imperatore C, Mangoni A (**2004**) Glycolipids from sponges. 13. Clarhamnoside, the first rhamnosylated α-galactosylceramide from *Agelas clathrodes*. Improving spectral strategies for glycoconjugate structure determination. *J Org Chem* 69(4): 1174–1179.

(8) Wikler MA (**2003**) Methods for dilution antimicrobial susceptibility tests for bacteria that grow aerobically; Approved Standard–Eighth Edition.

(9) Hopwood DA, Kieser T, Wright HM, Bibb MJ (**1983**) Plasmids, recombination and chromosome mapping in *Streptomyces lividans* 66. *J Gen Microbiol* 129(7): 2257–2269.