Supporting Information

Characterizing Amosamine Biosynthesis in Amicetin Reveals AmiG as a Reversible Retaining Glycosyltransferase

Ruidong Chen,[†‡§] Haibo Zhang,[†§] Gaiyun Zhang,[¶] Sumei Li,[†] Guangtao Zhang,[†] Yiguang Zhu,[†] Jinsong Liu,[⊥] and Changsheng Zhang[†*]

[†] CAS Key Laboratory of Tropical Marine Bio-resources and Ecology, RNAM Center for Marine Microbiology, Guangdong Key Laboratory of Marine Materia Medica, South China Sea Institute of Oceanology, Chinese Academy of Sciences, 164 West Xingang Road, Guangzhou 510301, P. R. China

[‡] Graduate University of the Chinese Academy of Sciences, Beijing 100049, China

[¶] Key Laboratory of Marine Biogenetic Resources, Third Institute of Oceanography, State Oceanic Administration, No.184 Daxue Road, Xiamen 361005,China

[⊥]State Key Laboratory of Respiratory Disease, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, 190 Kaiyuan Avenue, Guangzhou Science Park, Guangzhou 510530, China

[§] These authors contributed equally to this work.

Corresponding author: czhang2006@gmail.com

Content

Table S1. Strains and plasmids used and generated in this study.

Strains/Plasmids	Characteristic(s)	Sources and Refs
Strains		
E. coli		
DH5α	Host strain for cloning	Invitrogen
BW25113	Host strain for PCR targeting	[1]
ET12567	Donor strain for conjugation	[2]
BL21(DE3)	Host strain for protein expression	Novagen
S. vinaceusdrappus		
NRRL 2363	Wild type, amicetin producer	NRRL
AM1004	the *amiG* gene disrupted mutant of NRRL 2363	This study
AM1005	the *amiH* gene disrupted mutant of NRRL 2363	This study
Plasmids		
pET28a	Km^r, expression vector	Novagen
pCSG3201	11650 bp *Bam*HI/*Nde*I fragment from pCSG3104 in pET28a	[3]
pAM1004	pCSG3201derivative where *amiG* was disrupted by *aac(3)IV* using primers AmiGDF and AmiGDR (Table S2)	This study
pAM1005	pCSG3201derivative where *amiH* was disrupted by *aac(3)IV* using primers AmiHDF and AmiHDR (Table S2)	This study
pCSG3247	1.5 kb amiG PCR fragment from genomic DNA was cloned into pET28a (*Nde*I/*Bam*HI) using primer pairs AmiG-P1/P2 (Table S2)	This study

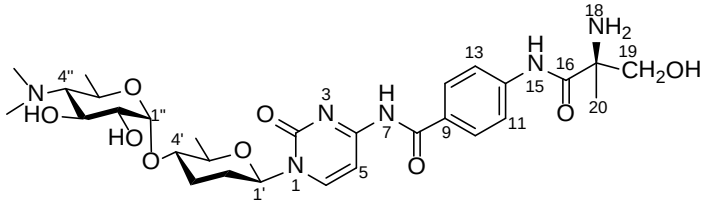
Table S2. Primers used for mutant construction and confirmation, and for *amiG* expression.

Primers	Gene target	Sequences (restriction sites are underlined)
AmiGDF	*amiG*	5’- TTCGTATGCAGCGTGTACTACCCGGTGACCGGTGGTGCGATTCCGGGGATCCGTCGACC -3’
AmiGDR		5’- GTTCCGGATGTATCCGAGGCAGCTGTCGCTGTCGAAGCGTGTAGGCTGGAGCTGCTTC -3’
AmiHDF	*amiH*	5’- GGCTACGCCGCGTACTTGGAGCGCCATGCCGGGTATTTCATTCCGGGGATCCGTCGACC -3’
AmiHDR		5’- CGCCTTCACTACGATGCTCCCGCCGAATCGGCCTCGTTCTGTAGGCTGGAGCTGCTTC -3’
AmiGDJF	*amiG*	5’- ACTGCTGCGCATGGACCG -3’
AmiGDJR		5’- GGTGCCCACGTCGAGGAG -3’
AmiGDJF2	*amiG*	5’- AGACCGTGGGCGGCTAC -3’
AmiGDJR2		5’- ACAGAGCGGAGGTGATG -3’
AmiHDJF	*amiH*	5’- GGCCACCGAAGGGATGCCCG -3’
AmiHDJR		5’- GTTCATGTGGGAGCTGAAG -3’
AmiG-P1	*amiG*	5’- CGCCGCATATGAACATTCTTTTCGTA -3’ (*Nde*I)
AmiG-P2		5’- GCATCGGATCCTTCGGCATTG -3’ (*Bam*HI)

Table S3. [1]H NMR data of compounds **1a, 1c**, **2c**, **1**, **1d** and **2** (multi, *J* in Hz).

No.	**1a**[a]	**1c**[a]	**2c**[a]	**1**[a]	**1d**[a]	**2**[a]
	δ_H	**δ_H**	**δ_H**	**δ_H**	**δ_H**	**δ_H**
5	7.63 (d, 7.5)	7.62 (d, 7.5)	7.61 (d, 7.5)	7.64 (d, 7.5)	7.62 (d, 7.5)	7.59 (d, 7.5)
6	8.20 (d, 7.5)	8.23 (d, 7.5)	8.14 (d, 7.5)	8.19 (d, 7.5)	8.21 (d, 7.5)	8.14 (d, 7.5)
10	8.02 (d, 8.0)	8.01 (d, 8.5)	7.78 (d, 9.0)	8.00 (d, 9.0)	8.02 (d, 8.5)	7.79 (d, 8.6)
11	7.86 (d, 8.0)	7.84 (d, 8.5)	6.72 (d, 9.0)	7.87 (d, 9.0)	7.86 (d, 8.5)	6.72 (d, 8.6)
13	7.86 (d, 8.0)	7.84 (d, 8.5)	6.72 (d, 9.0)	7.87 (d, 9.0)	7.86 (d, 8.5)	6.72 (d, 8.6)
14	8.01 (d, 8.0)	8.01 (d, 8.5)	7.78 (d, 9.0)	8.00 (d, 9.0)	8.02 (d, 8.5)	7.79 (d, 8.6)
19	3.79 (d, 11.0) 4.10 (d, 11.0)	3.83 (d, 11.5) 4.13 (d, 11.5)		3.87 (d, 12.0) 4.11 (d, 12.0)	3.82 (d, 11.5) 4.11 (d, 11.5)	
20	1.62 (s)	1.64 (s)		1.64 (s)	1.65 (s)	
1′	5.81 (d, 8.5)	5.80 (d, 8.5)	5.79 (d, 8.5)	5.75 (d, 8.0)	5.80 (d, 10.0)	5.74 (d, 8.8)
2′	1.70 (brd, 9.5) 2.19 (brd, 9.5)	1.73 (m) 2.20 (m)	1.72 (m) 2.18 (m)	1.73 (m) 2.15 (m)	1.71 (brd, 10.0) 2.20 (brd, 10.0)	1.72 (m) 2.13 (m)
3′	1.70 (m) 2.41 (m)	1.73 (m) 2.43 (m)	1.72 (m) 2.40 (m)	1.69 (m) 2.19 (m)	1.68 (m) 2.44 (m)	1.68 (m) 2.19 (m)
4′	3.44 (td, 3.5, 9.0)	3.46 (td,5.5, 9.5)	3.44 (m)	3.32 (m)	3.45 (m)	3.32 (m)
5′	3.78 (m, overlapped)	3.80 (m)	3.76 (m)	3.54 (m)	3.71 (m)	3.53 (m)
6′	1.39 (d, 6.0)	1.40 (d, 6.0)	1.41 (d, 6.0)	1.36 (d, 6.0)	1.42 (d, 6.5)	1.36 (d, 6.0)
1″	4.97 (d, 3.5)	5.08 (d, 4.0)	5.02 (d, 4.0)		5.05 (d, 4.0)	
2″	3.49 (dd, 3.5, 9.0)	3.57 (dd, 4.0, 9.5)	3.48 (dd, 4.0, 9.5)		3.54 (m)	
3″	3.92 (br. t, 9.0)	3.74 (t, 9.5)	3.56 (t, 9.5)		3.69 (m)	
4″	2.44 (t, 9.0)	2.83 (t, 10.0)	2.61 (t, 10.0)		3.86 (m)	
5″	3.92 (br. t, 9.0)	4.06 (dq, 6.5, 10.0)	3.86 (m)		3.67 (m)	
6″	1.33 (d, 6.0)	1.35 (d, 6.0)	1.30 (d, 6.0)		3.62 (m)	
					3.77 (m)	
7″	2.64 (s)					
8″	2.64 (s)					

[a] recorded at 500 MHz in methanol-*d4*



Atom numbering for amicetin (**1a**)

Table S4. ^{13}C NMR data of compounds **1a, 1c**, **1** and.**2**.

No.	**1a** [a]	**1c** [a]	**1** [a]	**2** [a]
	δ_C	**δ_C**	**δ_C**	**δ_C**
2	157.4 C	157.2 C	157.2 C	157.5 C
4	164.8 C	164.7 C	164.9 C	164.9 C
5	98.8 CH	99.1 CH	98.8 CH	98.7 CH
6	146.4 CH	146.6 CH	146.2 CH	145.9 CH
8	168.4 C	168.8 C	168.2 C	168.6 C
9	130.3 C	130.0 C	130.3 C	121.0 C
10	130.3 CH	130.3 CH	130.3 CH	131.4 CH
11	121.2 CH	121.6 CH	121.3 CH	114.4 CH
12	143.7 C	143.5 C	144.2 C	155.1 C
13	121.2 CH	121.6 CH	121.3 CH	114.4 CH
14	130.3 CH	130.3 CH	130.3 CH	131.4 CH
16	171.9 C	171.6 C	171.3 C	
17	62.8 C	63.0 C	63.3 C	
19	66.5 CH_2	66.1 CH_2	65.8 CH_2	
20	20.1 CH_3	19.7 CH_3	19.5 CH_3	
1′	84.8 CH	84.7 CH	84.8 CH	84.6 CH
2′	31.0 CH_2	30.8 CH_2	32.4 CH_2	31.6 CH_2
3′	27.9 CH_2	27.9 CH_2	31.6 CH_2	32.3 CH_2
4′	75.6 CH	76.2 CH	71.7 CH	71.6 CH
5′	78.5 CH	78.2 CH	80.5 CH	80.4 CH
6′	19.5 CH_3	19.2 CH_3	18.5 CH_3	18.5 CH_3
1″	96.4 CH	96.7 CH		
2″	74.5 CH	73.2 CH		
3″	69.8 CH	70.6 CH		
4″	72.0 CH	58.9 CH		
5″	66.6 CH	66.5 CH		
6″	19.2 CH_3	17.9 CH_3		
7″	42.4 CH_3			
8″	42.4 CH_3			

[a] recorded at 125 MHz in methanol-*d4*

Table S5. ^{1}H, ^{13}C NMR data of compounds **3c, 3** and **3a** (multi, *J* in Hz)

No.	**3c** [a]		**3** [a]		**3a** [a,b]	
	δ_H	**δ_C**	**δ_H**	**δ_C**	**δ_H**	**δ_C**
2		157.5 C		157.4 C		157.2 C
4		167.4 C		167.3 C		167.4 C
5	5.95 (d, 7.0)	96.3 CH	5.94 (d, 7.5)	96.2 CH	5.93 (d, 7.5)	96.3 CH
6	7.71 (d, 7.0)	142.6 CH	7.72 (d, 7.5)	142.7 CH	7.71 (d, 7.5)	142.6 CH
1′	5.74 (d, 9.5)	83.9 CH	5.69 (d, 9.5)	84.0 CH	5.74 (d, 10.5)	83.9 CH
2′	1.73 (m) 2.02 (brd, 11.5)	30.7 CH_2	1.73 (m) 1.99 (brd, 10.0)	32.6 CH_2	1.73 (m) 2.04 (m)	30.7 CH_2
3′	1.65 (m) 2.38 (brd, 9.0)	28.3 CH_2	1.66 (m) 2.15 (dd, 3.5, 12.0)	31.3 CH_2	1.63 (m) 2.38 (dd, 3.5, 12.5)	28.1 CH_2
4′	3.41 (m)	76.6 CH	3.25 (td, 4.5, 9.5)	71.8 CH	3.39 (td, 4.0, 9.5)	75.9 CH
5′	3.70 (m, overlapped)	78.1 CH	3.47 (dt, 6.0, 9.0)	80.3 CH	3.72 (m)	78.2 CH
6′	1.36 (d, 5.5)	19.3 CH_3	1.32 (d, 6.0)	18.5 CH_3	1.37 (d, 6.0)	19.4 CH_3
1″	5.03 (d, 3.0)	97.1 CH			4.97 (d, 3.5)	96.4 CH
2″	3.51 (dd, 3.0, 9.0)	73.6 CH			3.50 (dd, 4.0, 9.5)	74.3 CH
3″	3.70 (brd, overlapped)	70.7 CH			3.94 (t, 10.0)	69.1 CH
4″	2.79 (t, 9.5)	59.3 CH			2.68 (t, 10.0)	72.0 CH
5″	4.02 (br. s)	66.3 CH			3.98 (m)	65.7 CH
6″	1.33 (d, 5.5)	17.9 CH_3			1.35 (d, 6.0)	19.3 CH_3
7″					2.77 (s)	42.4 CH_3
8″					2.77 (s)	42.4 CH_3

[a] recorded at 500 MHz in methanol-*d4*

[b] the NMR data of **3a** were previously reported in the supplemental reference 3.[3]

Experimental procedures

Construction of amiH-inactivation mutant AM1005. The mutant AM1005 was constructed by replacing a 591 bp internal *amiH* fragment with a 1369 bp DNA fragment containing *oriT* and *acc3(IV)* in pAM1005 (Figure S1A). The apramycin resistant exconjugants, resulting from the conjuation experiments between *E. coli* ET12567/pUZ8002 & pAM1006 (Table S1) and WT *S. vinaceus-drappus* NRRL 2363, were randomly selected for double crossover mutants by using the primer pairs AmiIDJF and AmiIDJR (Figure S1A, Table S2). The sizes of PCR products were as follows: WT, 1021 bp (Figure S4B, lane 1); double crossover mutant AM1005, 1799 bp (Figure S1B, lane 2). The PCR product from AM1005 (Figure S1B, lane 2) were subjected to sequence analysis for further confirmation of a double crossover event.

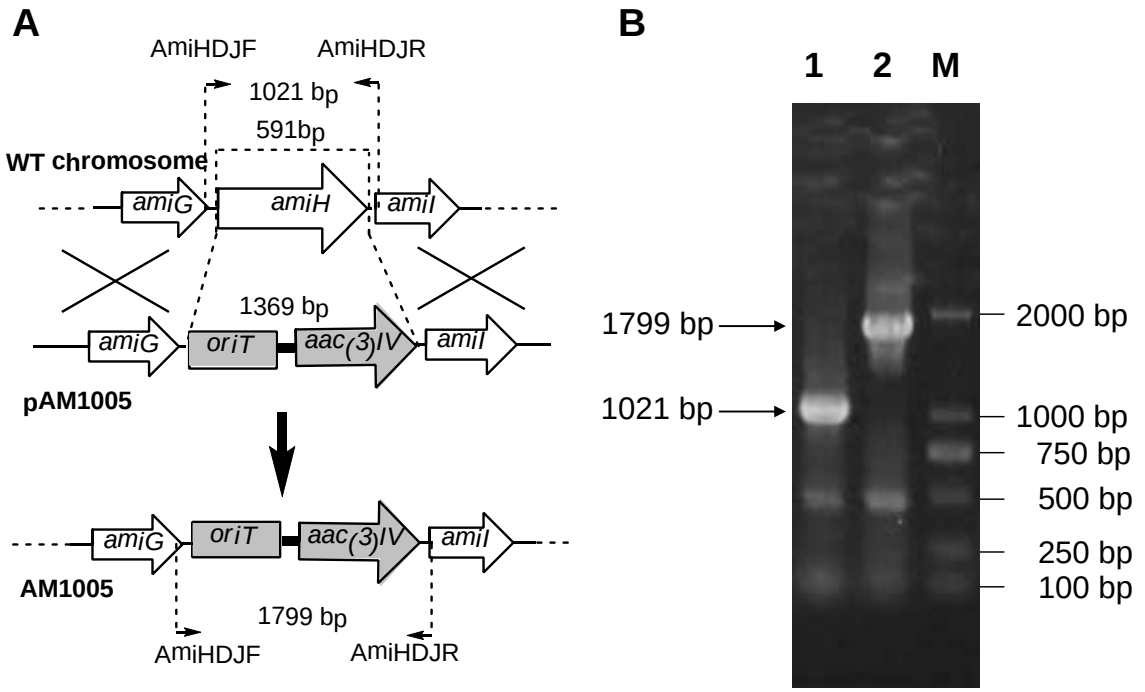


Figure S1. Construction of *amiH*-inactivation mutant AM1005. (**A**) **Depiction of *amiH*-inactivation**. Location of the diagnostic PCR primers AmiIDJF and AmiIDJR (Table S2) was indicated. Sizes of PCR products were also indicated: 782 bp for the wild type strain *S. vinaceus-drappus* NRRL 2363 and 1698 bp for the mutant AM1001. (**B**) **Gel electrophoresis of PCR products**. DNA templates were from: AM1011 (lane 1); *S. vinaceus-drappus* NRRL 2363 (lane 2); and DNA marker DL 2000 (Takara, lane M).

Construction of *amiG*-inactivation mutant AM1004. The mutant AM1004 was constructed by replacing a 1395 bp internal *amiI* fragment with a 1369 bp DNA fragment containing *oriT* and *acc3(IV)* in pAM1004 (Figure S2A). The apramycin resistant exconjugants, resulting from the conjuation experiments between *E. coli* ET12567/pUZ8002 & pAM1004 (Table S1) and WT *S. vinaceusdrappus* NRRL 2363, were randomly selected for double crossover mutants by using the primer pairs AmiGDJF and AmiGDJR (Figure S2A, Table S2). The sizes of PCR products were as follows: double crossover mutant AM1006, 1887 bp (Figure S2B, lane 1); single crossover mutant, 1913 bp and 1887 bp (Figure S2B, lane 2); WT, 1913 bp (Figure S2B, lane 3). Since we could not

differentiate the double and single crossover mutants (both were apramycin resistant), a pair of primers AmiGDJIF2 and AmiGDJR2 were designed to target on an internal 834 bp fragment of *amiG*. Using this pair of primers, no PCR products were amplified from the double crossover mutant AM1006 (Figure S2C, lane 1), while 834 bp PCR products were detectable from single crossover mutant (Figure S2C, lane 2) and wild type strain (Figure S2C, lane 3). The PCR product from AM1004 (Figure S2B, lane 1) were subjected to sequence analysis for further confirmation of a double crossover event.

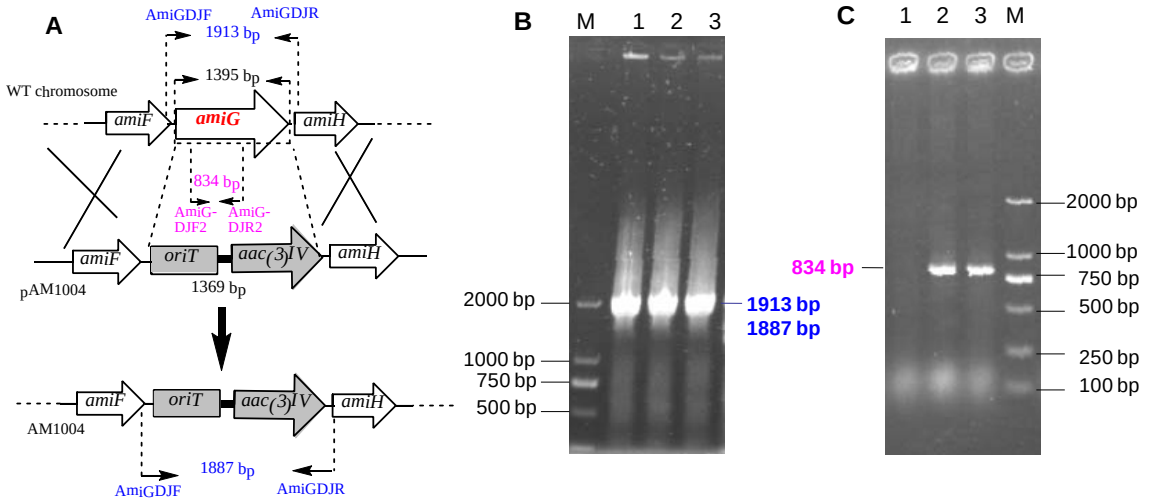


Figure S2. Construction of *amiG*-inactivation mutant AM1004. (**A**) **Depiction of *amiG*-inactivation**. Location of the diagnostic PCR primers AmiGDJF/AmiGDJR and AmiGDJF2/AmiGDJR2 (Table S2) was indicated. Sizes of PCR products for PCR primers AmiGDJF/AmiGDJR were indicated: 1913 bp for the wild type strain *S. vinaceusdrappus* NRRL 2363; 1887 bp for the mutant AM1004; sizes of PCR products for PCR primers AmiGDJF2/AmiGDJR2 were indicated: 834 bp for both WT *S. vinaceus-drappus* NRRL 2363 and single crossover mutant; none for the double crossover mutant AM1004. (**B**) **Gel electrophoresis of PCR products for the primer pair AmiFDJF/AmiFDJR**. DNA templates were from: double crossover mutant AM1004 (lane 1); single crossover mutant (lane 2); WT (lane 3); DNA marker DL 2000 (lane M). (**C**) **Gel electrophoresis of PCR products for the primer pair AmiFDJF2/AmiFDJR2**. DNA templates were from: double crossover mutant (lane 1); single crossover mutant AM1003 (lane 2); WT (lane 3); DNA marker DL 2000 (lane M).

General HPLC analysis. General HPLC analysis for metabolite analysis and enzyme assays was carried out on a reversed phase column Luna C18 (Phenomenex, 150 × 4.6 mm, 5 μm) with UV detection at 254 nm under the following program: solvent system (solvent A, 10% CH_3CN in water supplementing with 0.8% trifluoroacetic acid (TFA); solvent B, 90% CH_3CN in water); process: 0 -40% B (0 – 15 min); 40 - 80% B (15 -18 min); 80 - 0% B (18-19 min); 0% B (19 – 25 min), flow rate at 1 mL min^{-1}.

Isolation of compounds from the Δ*amiH* mutant AM1005. The 10 L fermentation broth of the Δ*amiH* mutant AM1005 was extracted 4 times with 16 L n-butanol. Given that n-butanol was hard to be removed by conventional evaporation in vacuum, an equal volume of water and some methanol were added to the extracts to facilitate the removal of organic solvents under in vacuum

by heating at ~55^{0}C for days, affording residue I by this way. The mycelia cake was extracted 3 times with 3 L acetone. After removing acetone in vacuum, the residue was re-extracted by 1.5 L n-butanol to afford residue II upon removal of the solvent with addition of equal volume of water and some methanol for evaporation under vacuum and heating at 55^{0}C. Residues I and II were combined as the crude extracts for further isolation. The crude extract (8.1 g) was subjected to C18 reverse phase column (YMC*GEL ODS-A, 12 nm S-50 μm, 30 × 2.5 cm I.D.) using MPLC to give 7 fractions (A-G), with a linear gradient elution (0-70 min, 0-100% CH_3OH/H_2O; 20 mL min^{-1}; UV detection at 268 nm). Fraction A (900 mg) was found to contain the target molecule and was loaded on Sephadex LH-20, eluting with $CHCl_3/CH_3OH$ (1:1) to yield sub-fractions A1-A3. Fraction A2 (31 mg) was purified by preparative TLC (20 × 20 cm), under development with $CHCl_3/CH_3OH/H_2O$/formic acid, 5:5:0.5:0.1) to obtain A2-1 (19 mg). A2-1 was purified by semi-preparative HPLC eluting with gradient solvents (phase A: 10 % CH_3CN/H_2O, 1‰ formic acid; phase B: 90 % CH_3CN/H_2O, 2.5 mL min^{-1}, UV detection at 268 nm) to get **1c** (freeze-drying, 7.6 mg). Fraction A3 (26.6 mg) was purified by preparative silica gel column (300-400 mesh), developed with $CHCl_3/CH_3OH/H_2O$ (5:1:0.1) to obtain **2c** (2.6 mg). Fraction A1 (221 mg) was subjected to C18 reverse phase column using MPLC to give 6 fractions (A1-1~A1-6), with a linear gradient elution (0%-100% CH_3OH/H_2O; 20 mL min^{-1}; UV detection at 268 nm). A1-1 was purified by semi-preparative HPLC eluting with gradient solvents (Gemini 5μm C-18, 250×10 mm, Phenomenex company; phase A: 2 ‰ ethylenediamine in water; phase B: methanol; 2.5 mL min^{-1}; UV detection at 260 nm) to get A1-1-2, A1-1-2 was purified by semi-preparative HPLC with isocratic elution (Gemini 5μm C-18, 250×10 mm, Phenomenex company; phase A: 3‰ formic acid in H_2O; phase B: 20% CH_3CN; 2.5 mL min^{-1}; UV detection at 250 nm) to get **3c** (2.2 mg).

Isolation of compounds from the Δ*amiG* mutant AM1004. Around 10.5 g of crude extract was obtained from 9 L AM1004 fermentation by following the extraction procedure described above for the mutant AM1005. The crude extract was subjected to reverse phase medium pressure liquid chromatography (RP-MPLC) on ODS column (YMC*GEL ODS-A, 12 nm S-50 μm, 40 × 3.5cm) with linear gradient elution (0-70 min, CH_3OH/H_2O 0-100%, 20 mL min^{-1}, UV detection at 268 nm and 300 nm) to get 8 fractions (Fr. A-Fr. H). Fraction C (483 mg) was found to contain the desired product, and was subsequently purified by Sephadex LH-20 eluting with $CHCl_3/CH_3OH$ (1:1) to give sub-fractions C1-C3. Solvent in Fraction C2 was naturally evaporated in fuming hood at room temperature, and the residues were re-crystallized to afford **1** (27 mg). Fraction A (105 mg) was purified by Sephadex LH-20, eluting with $CHCl_3/CH_3OH$ (1:1) to give sub-fractions A1-A5. A-1 (14 mg) was purified by semi-preparative HPLC eluting with gradient solvents (Gemini 5μm C-18, 250×10 mm, Phenomenex company; phase A: 1‰ formic acid in H_2O; phase B: CH_3OH, 2.5 mL min^{-1}; UV detection at 275 nm) to get **3** (freeze-drying, 3.0 mg).

Isolation of compounds from the Δ*amiR* mutant AM1009. The 12 L fermentation broth of the Δ*amiR* mutant AM1009[3] was extracted 3 times with 5 L n-butanol. Since that n-butanol was hard to be removed by conventional evaporation under vacuum, an equal volume of water were added to facilitate the removal of organic solvents under vacuum by heating at ~55^{0}C for days, affording a residue I. The mycelia cake was extracted 3 times with 3 L acetone. After removing acetone under vacuum, the residue was re-extracted by 1.5 L n-butanol to afford residue II upon removal of the solvent with addition of equal volume of water under vacuum by heating at 55^{0}C. Residues I and II

were combined as the crude extracts for further isolation. The crude extract (13.5 g) was subjected to C18 reversed-phase column (YMC*GEL ODS-A, 12 nm S-50 μm, 30 × 2.5 cm I.D., YMC Company, Japan) using MPLC to give 5 fractions A-E, with a linear gradient elution (0-70 min, 0-100% CH_3CN/H_2O; 20 mL min^{-1}; UV detection at 324 nm). Fraction B (1.1 g) was subjected to semi-preparative HPLC (YMC-Pack ODS-A, 12 nm S-5 μm, 250 × 20 mm), eluting with gradient program (0-30 min, 0-40% CH_3CN/H_2O; 30-50 min, 40-100% CH_3CN/H_2O; 50-60 min, 100% CH_3CN; 10 mL min^{-1}; UV detection at 324 nm) to yield sub-fractions B1-B7. Fractions B5 and B6 (631 mg) were combined and purified again by preparative TLC (20 × 20 cm, 20 plates), under development with $CHCl_3/CH_3OH/H_2O$ (4:1:0.1, UV detection at 254 nm) to obtain compounds **2a** ($R_f \approx 0.56$, 289.4 mg) and **2b** ($R_f \approx 0.22$, 170.3 mg).

Cloning, expression, and purification of AmiG. The *amiG* gene was amplified from the genomic DNA of *S. vinaceus-drappus* NRRL 2363 by PCR using high fidelity DNA polymerase with the primers pairs AmiG-P1/P2 (Table S2): 5' – CGCCGCATATGAACATTCTTTTCGTA - 3' (forward, *Nde*I) and 5' – GCATCGGATCCTTCGGCATTG - 3' (reverse, *Bam*HI). The 1.5 kb PCR product was digested with *Nde*I and *Bam*HI, and cloned into pET28a to afford pCSG3247 after sequence confirmation. A single transformant of *E. coli* BL21 (DE3)/pCSG3247 was inoculated into 3 mL LB medium supplemented with 50 μg/mL of kanamycin and grown at 37 ºC overnight. The precultures were inoculated into 1 L LB medium containing 50 μg/mL of kanamycin and was grown at 28 ºC until the OD_{600} achieved 0.5 – 0.7. The *amiG* expression was induced with the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) at a final concentration of 0.4 mM for 6 h at 28 ºC. The cells obtained from 1 L of culture were then harvested by centrifugation and resuspended in 30 mL of the binding buffer (20mm sodium phosphate, 10 mM imidazole; 300 mM NaCl, pH 7.4). The cells were lysed by sonication on ice and the insoluble material was removed by centrifugation at 10000 g for 30 min at 4^0C, the supernatant was applied to a HisTrap HP column (1 ml, GE Healthcare) and the N-$(His)_6$-tagged AmiG was eluted with a linear gradient of imidazole (10 – 500 mM) in the binding buffer by a ÄKTA Purifier system (GE Healthcare). The purified protein was desalted through PD-10 column (GE Healthcare), and stored in the buffer containing 50 mM Tris-HCl (pH 8.0) and 10 % glycerol at -80 ºC until use. Protein concentration was measured by Bradford assay.[4]

***In vitro* AmiG assays.** Methods for optimizing AmiG *in vitro* assays were described in **Figure S13** legend (Page S32). General AmiG assays were performed in a total volume of 50 μL in MOPS buffer (50 mM, pH 6.5) containing 10 mM $MgCl_2$. Individual AmiG assay conditions were described in legends of **Figures S18-S21** (Pages S40-S43). Kinetic analysis of AmiG reactions were carried out according to **Figure S22** (Pages S44). AmiG reactions were quenched by the addition of 50 μL MeOH, and denatured proteins removed by centrifugation. A 5 ml enzymatic conversion of **1** (5 mg) to **1d** was carried out and the product was purified by semi-preparative HPLC eluting with gradient solvents (Gemini 5μm C-18, 250×10 mm, Phenomenex company; phase A: 1‰ formic acid in H_2O; phase B: CH_3OH, 2.5 mL min^{-1}, UV detection at 254 nm) to get **1d** (freeze-drying, 1.0 mg). A 20 mL AmiG reverse reaction was performed to convert **2a** (20 mg) to **2**. The product **2** (4.8 mg) was isolated by isocratic elution (30% B) on semi-preparative reverse phase column (Gemini-NX 5μm C-18, 250×10 mm, Phenomenex company; phase A: 1‰ formic acid in H_2O, phase B: CH_3OH, 3.0 mL min^{-1}; UV detection at 317 nm).

Figure S3. Spectral data of amicetin (**1a**).

Figure S3. (A) The ^{1}H NMR spectrum of amicetin (**1a**) in methanol-*d4*.

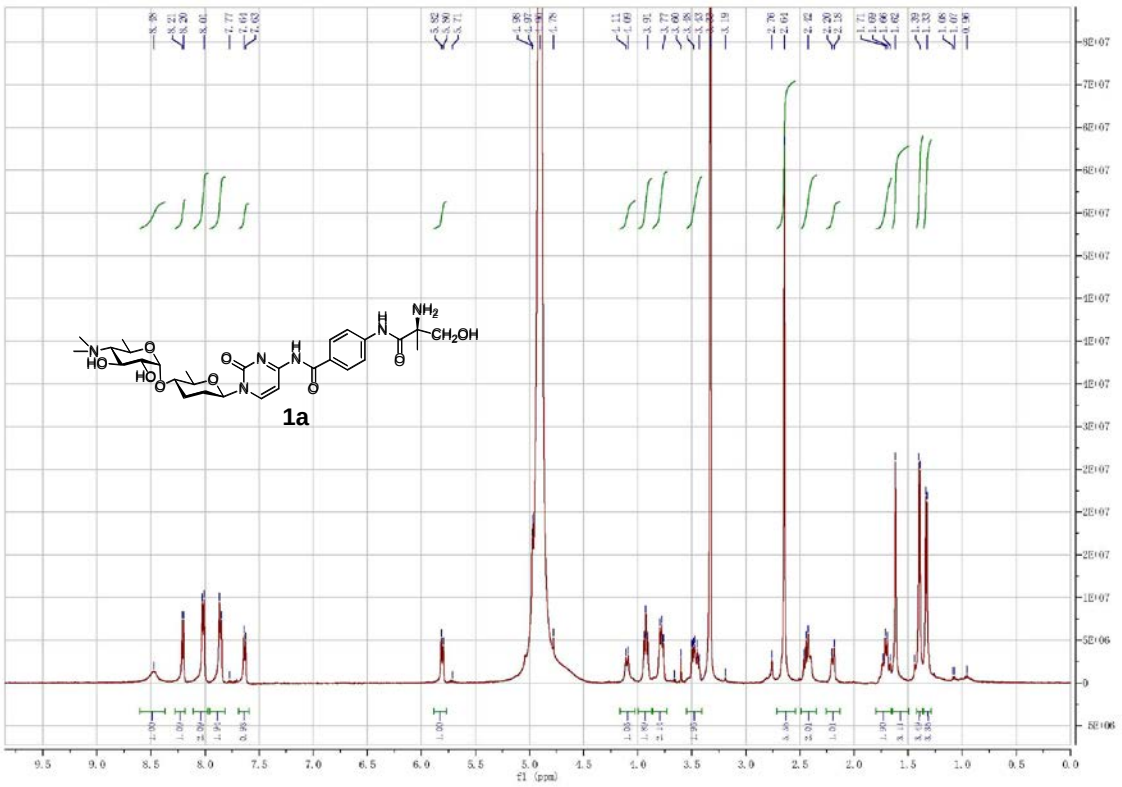


Figure S3. (B) The ^{13}C NMR spectrum of amicetin (**1a**) in methanol-*d4*.

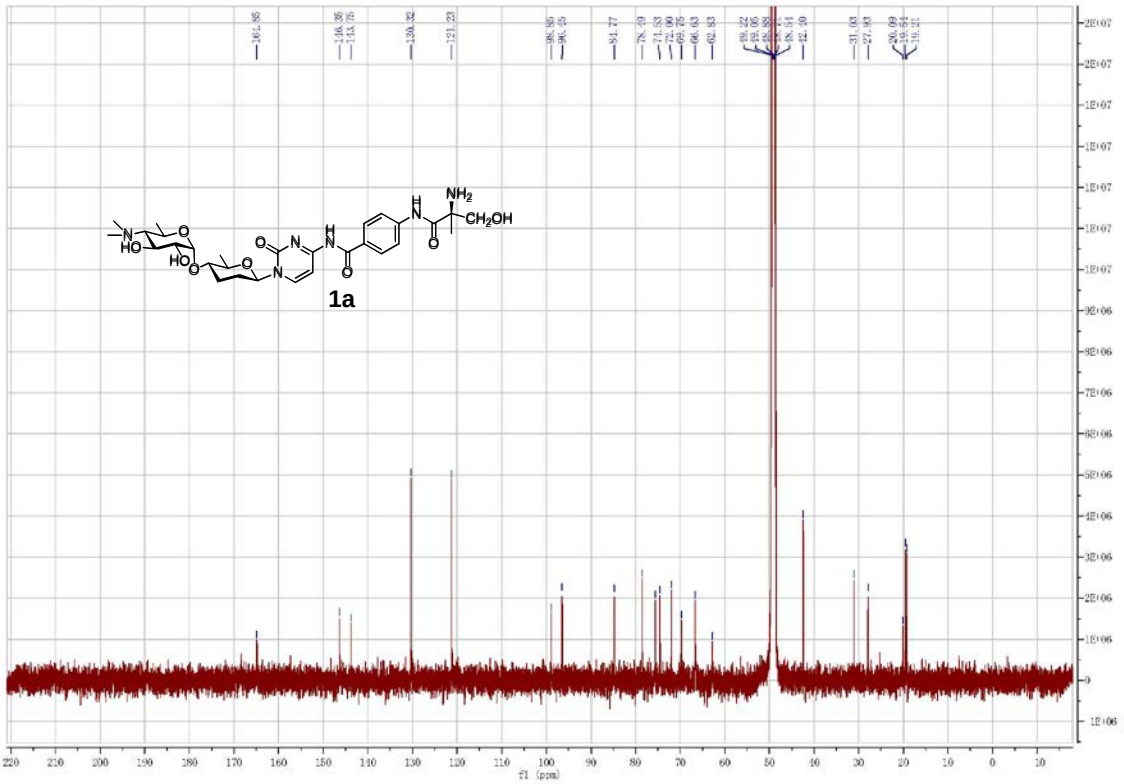


Figure S3. (C) The HCQC spectrum of amicetin (**1a**) in methanol-*d4*.

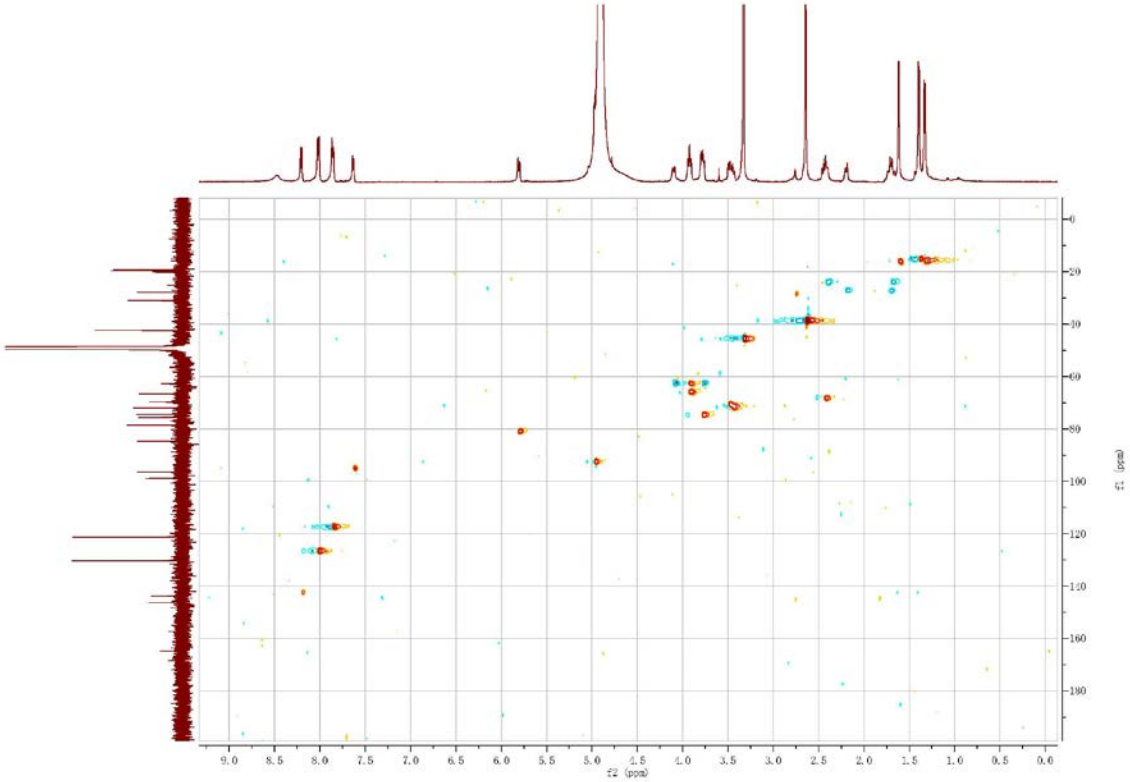


Figure S3. (D) The COSY spectrum of amicetin (**1a**) in methanol-*d4*.

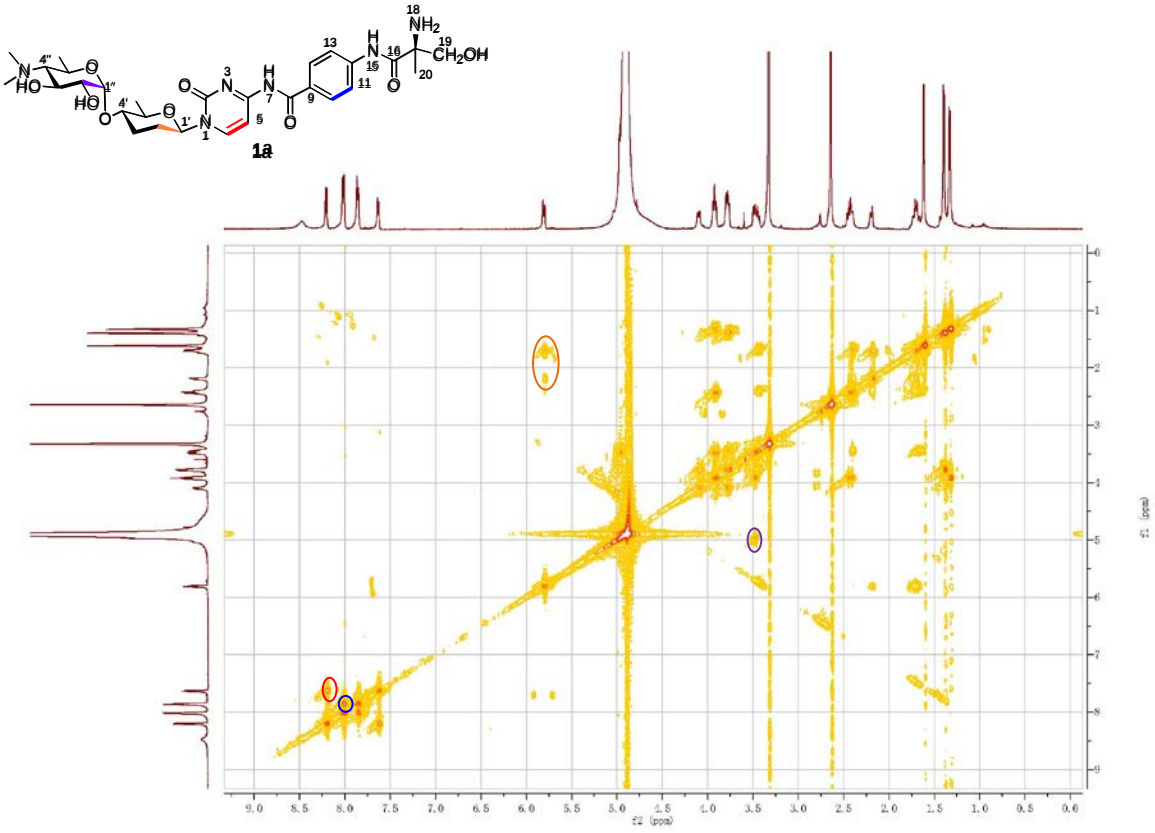


Figure S3. (E) The HMBC Spectrum of amicetin (**1a**)

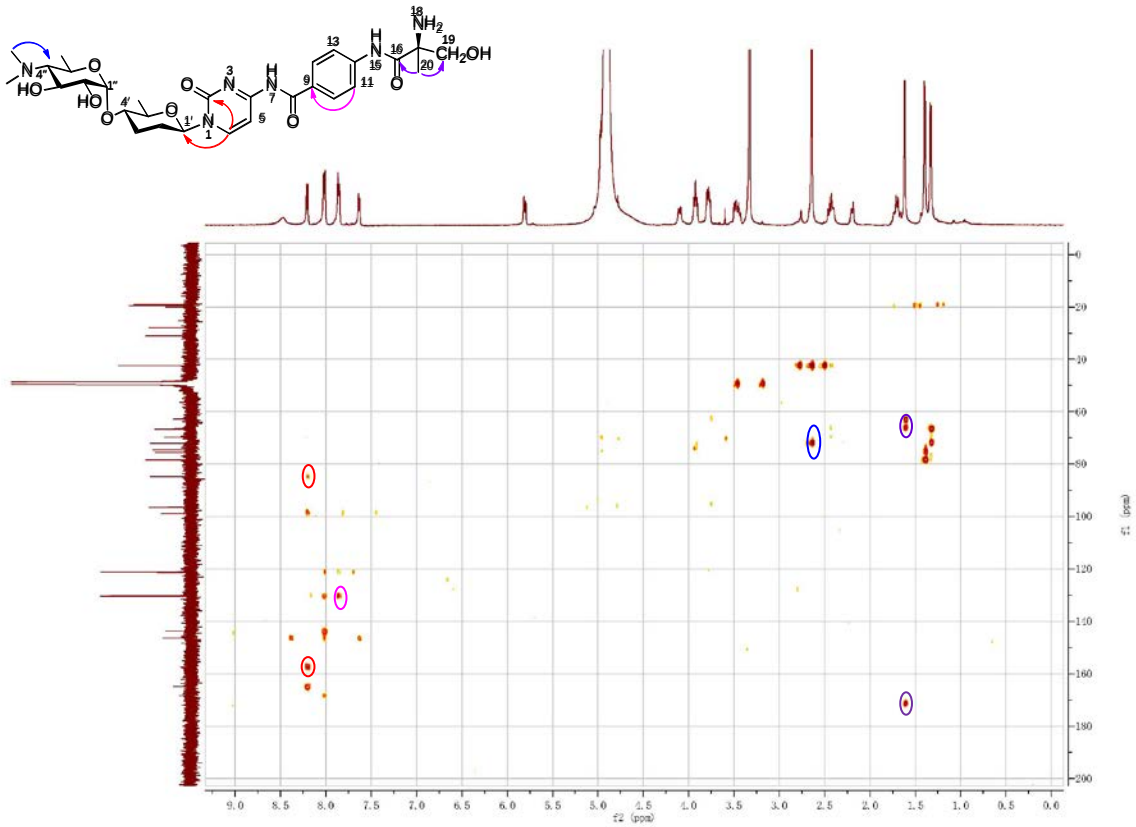


Figure S3. (F) The NOESY spectrum of amicetin (**1a**)

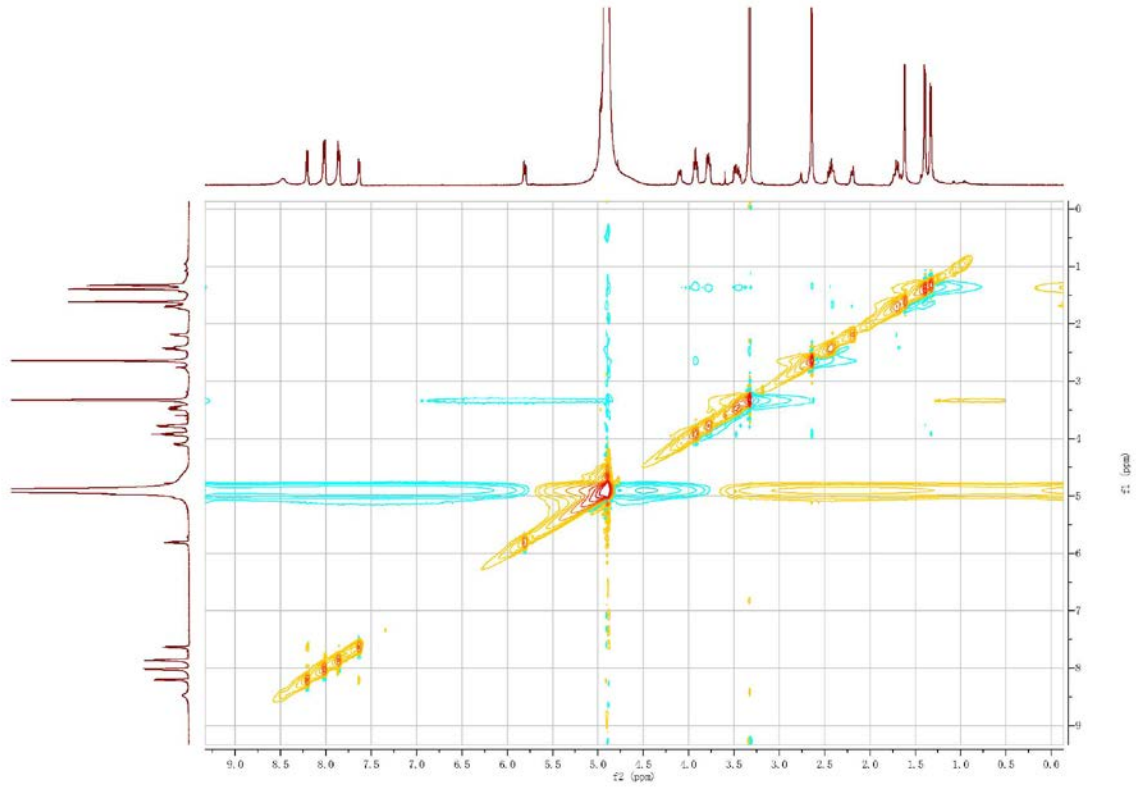


Figure S4. Spectral data of di-demethyl-amicetin (**1c**).

Figure S4. (A) The ^{1}H NMR spectrum of di-demethyl-amicetin (**1c**) in methanol-*d4*.

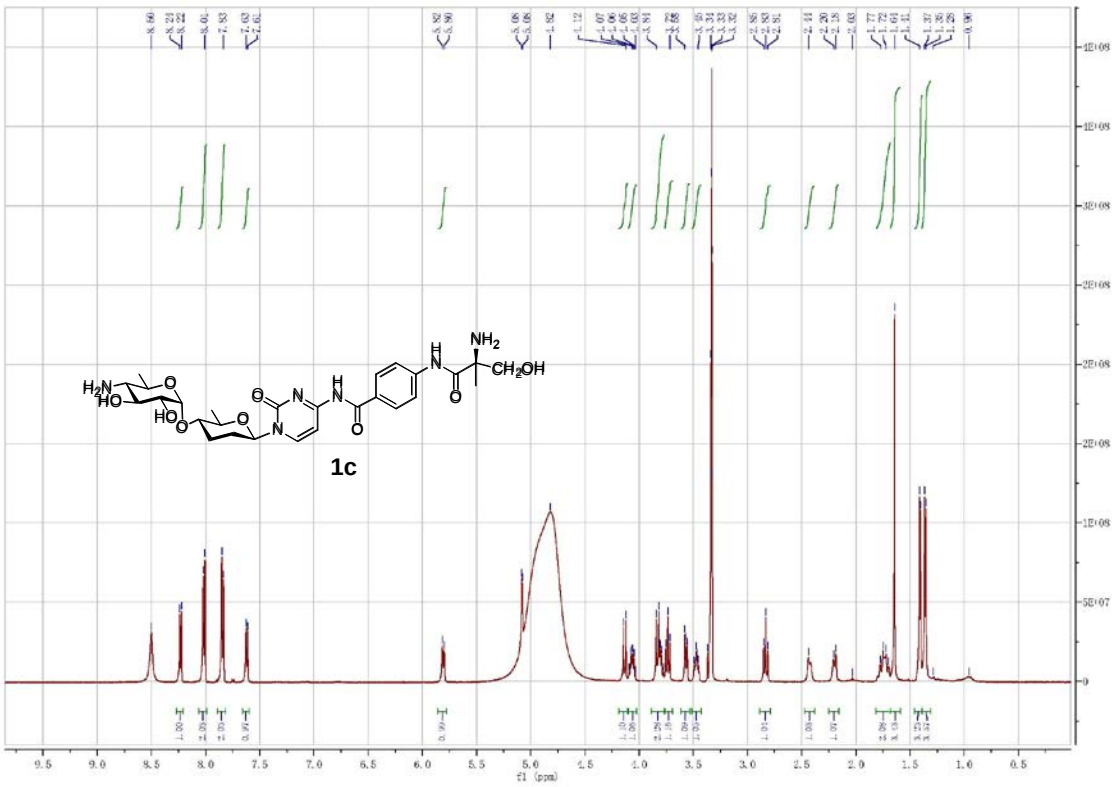


Figure S4. (B) The ^{13}C NMR spectrum of di-demethyl-amicetin (**1c**) in methanol-*d4*.

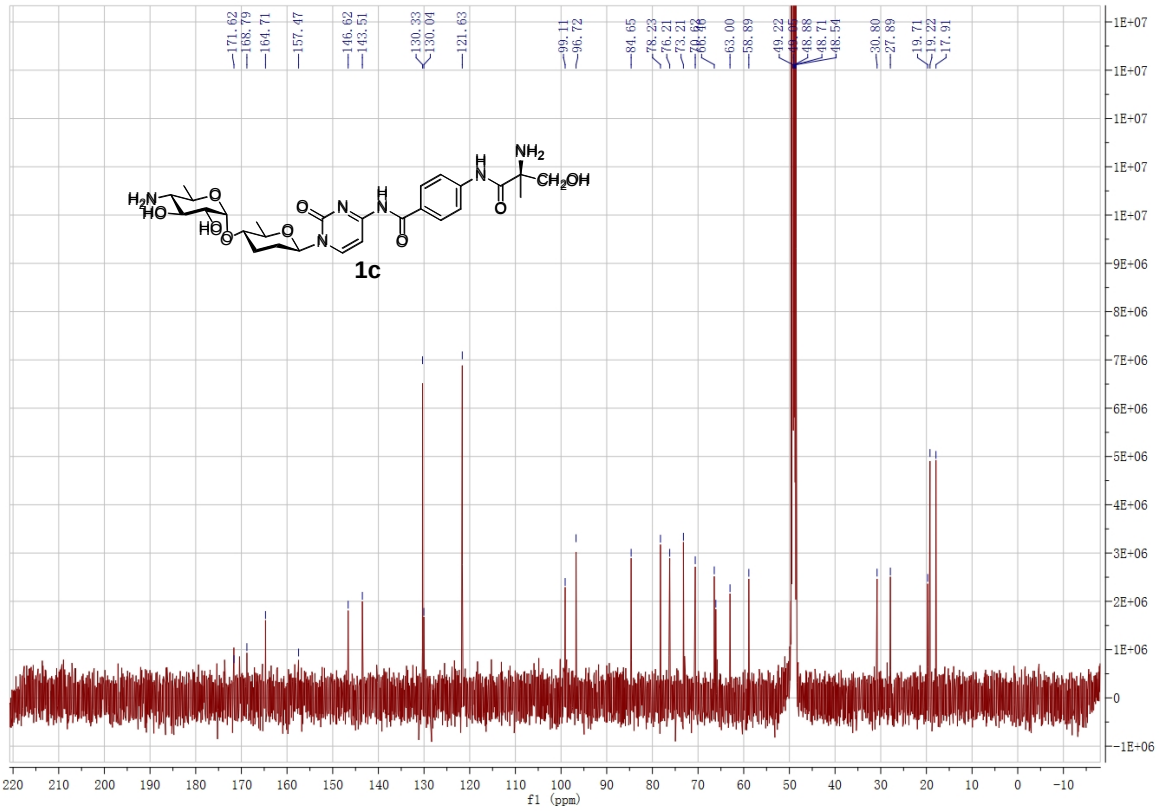


Figure S4. (C) The DEPT-135 NMR spectrum of di-demethyl-amicetin (**1c**).

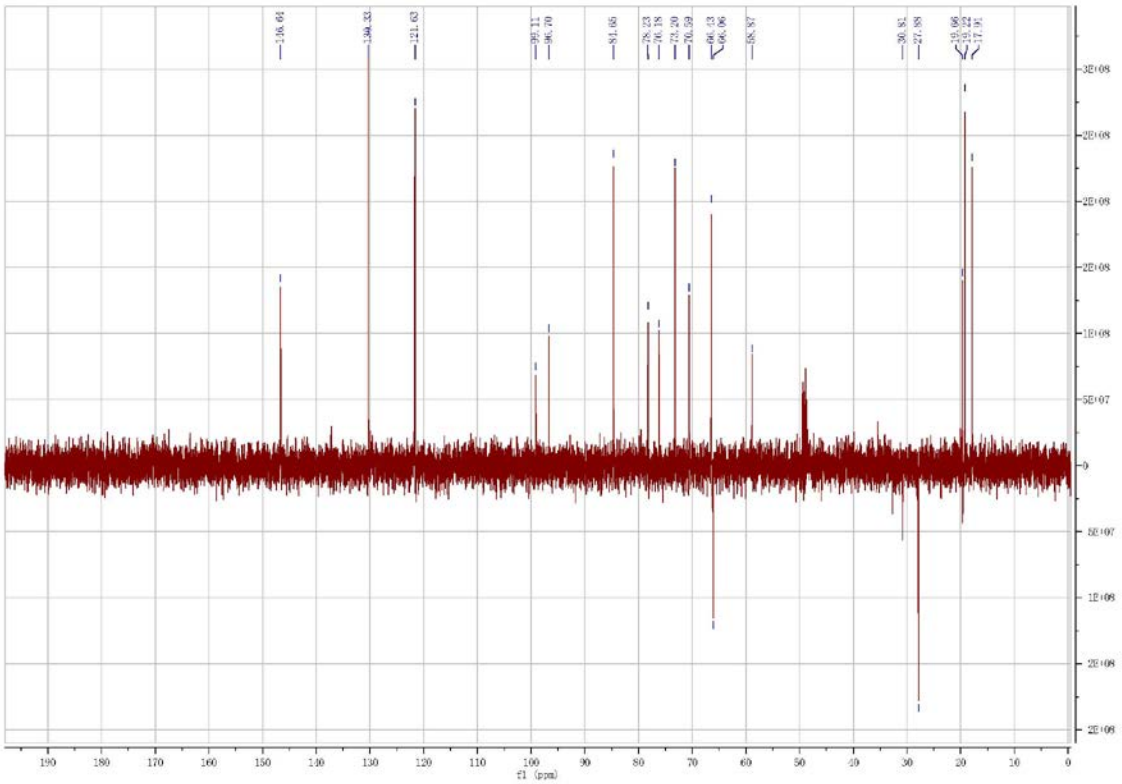


Figure S4. (D) The HSQC spectrum of di-demethyl-amicetin (**1c**).

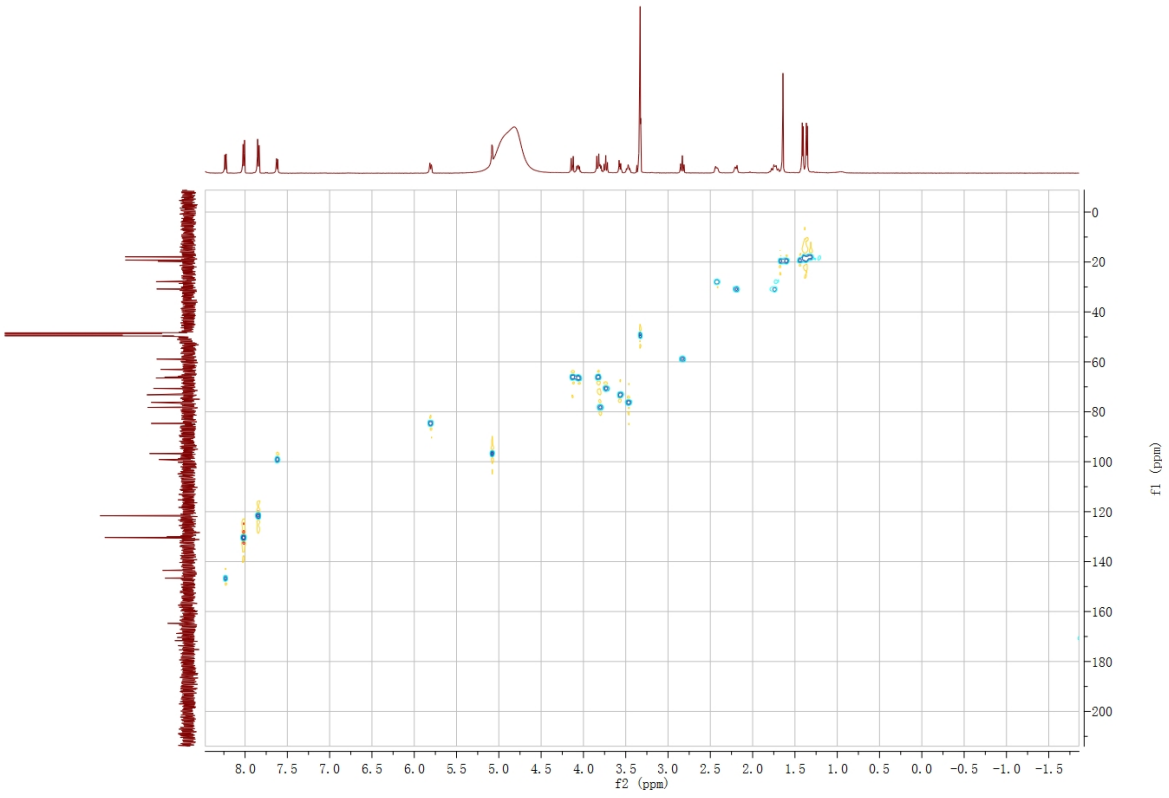


Figure S4. (E) The COSY spectrum of di-demethyl-amicetin (**1c**).

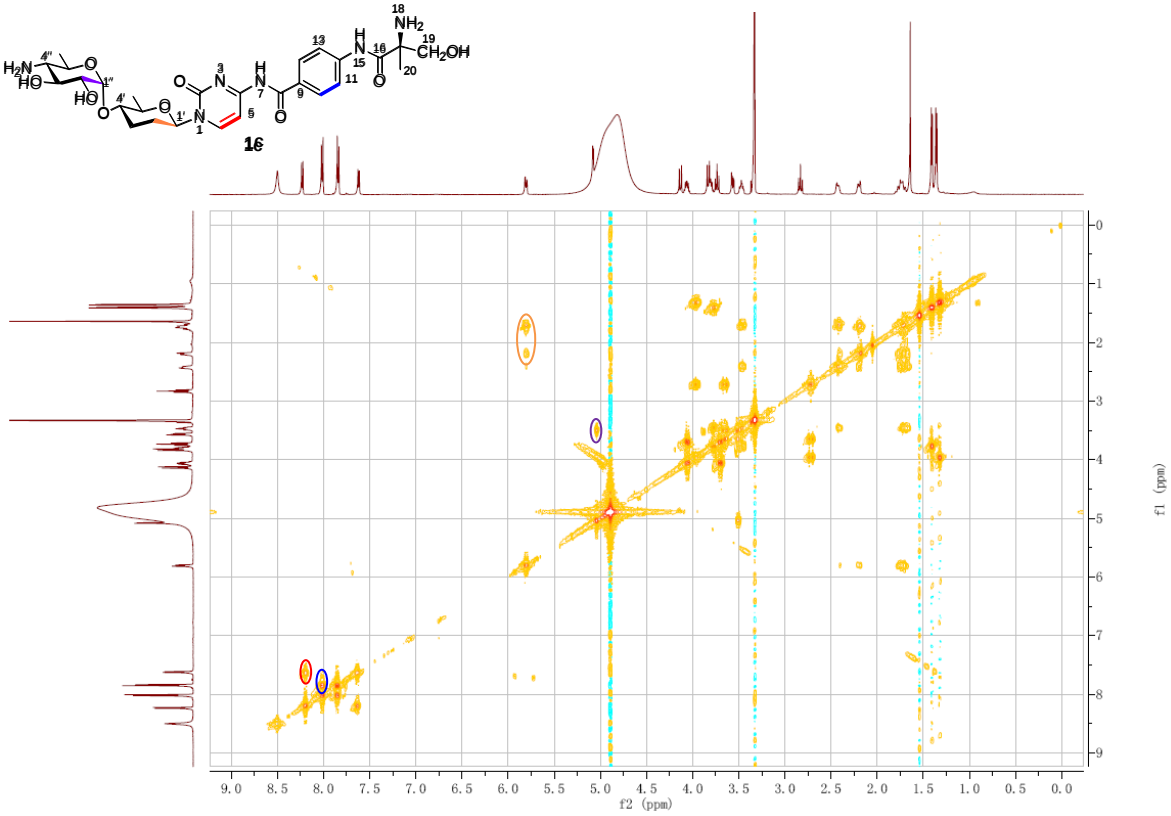


Figure S4. (F) The HMBC Spectrum of di-demethyl-amicetin (**1c**).

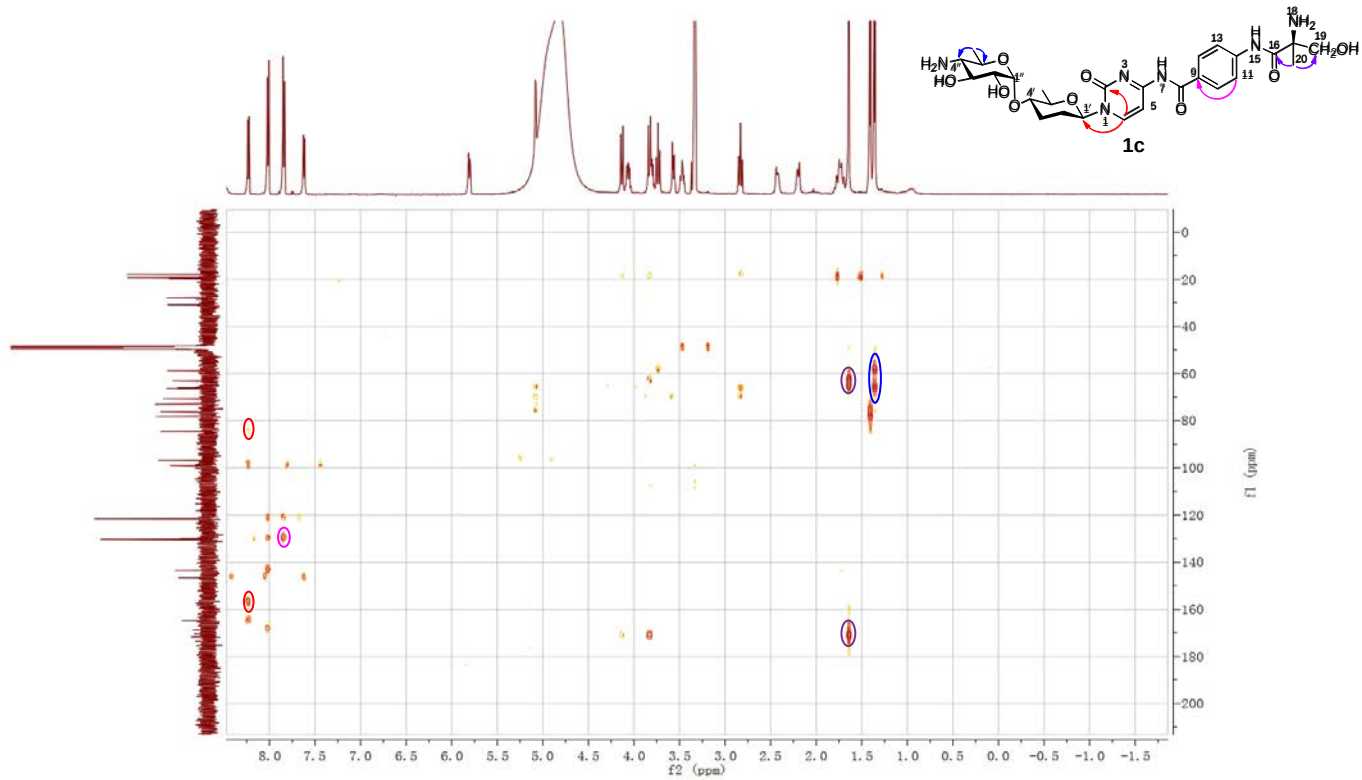


Figure S4. (F) The NOESY Spectrum of di-demethyl-amicetin (**1c**).

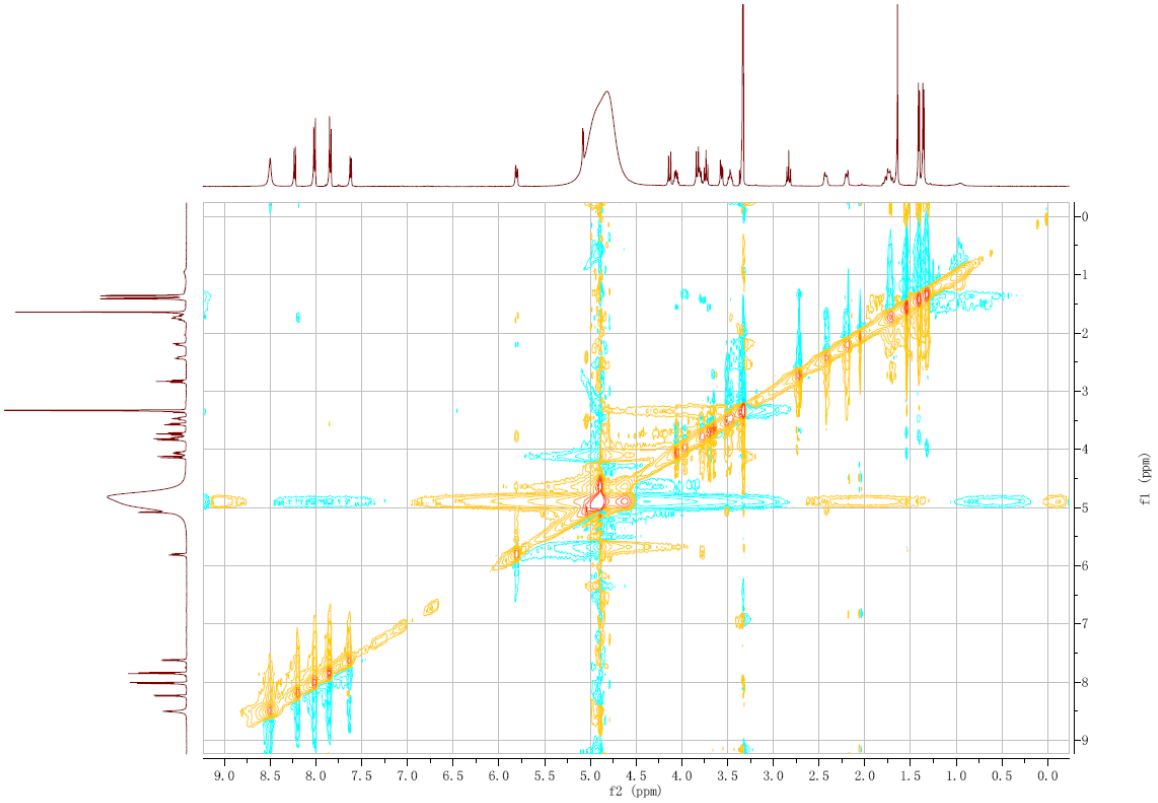


Figure S4. (G) The HR-ESI-MS Spectrum of di-demethyl-amicetin (**1c**).

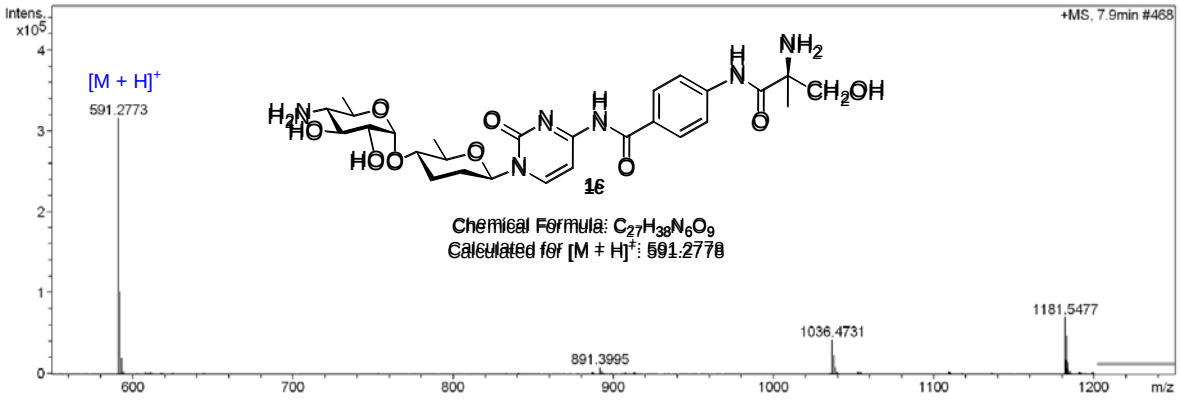


Figure S5. Spectral data of di-demethyl-plicacetin (**2c**).

Figure S5. (A) The ^{1}H NMR spectrum of di-demethyl plicacetin (**2c**) in methanol-*d4*.

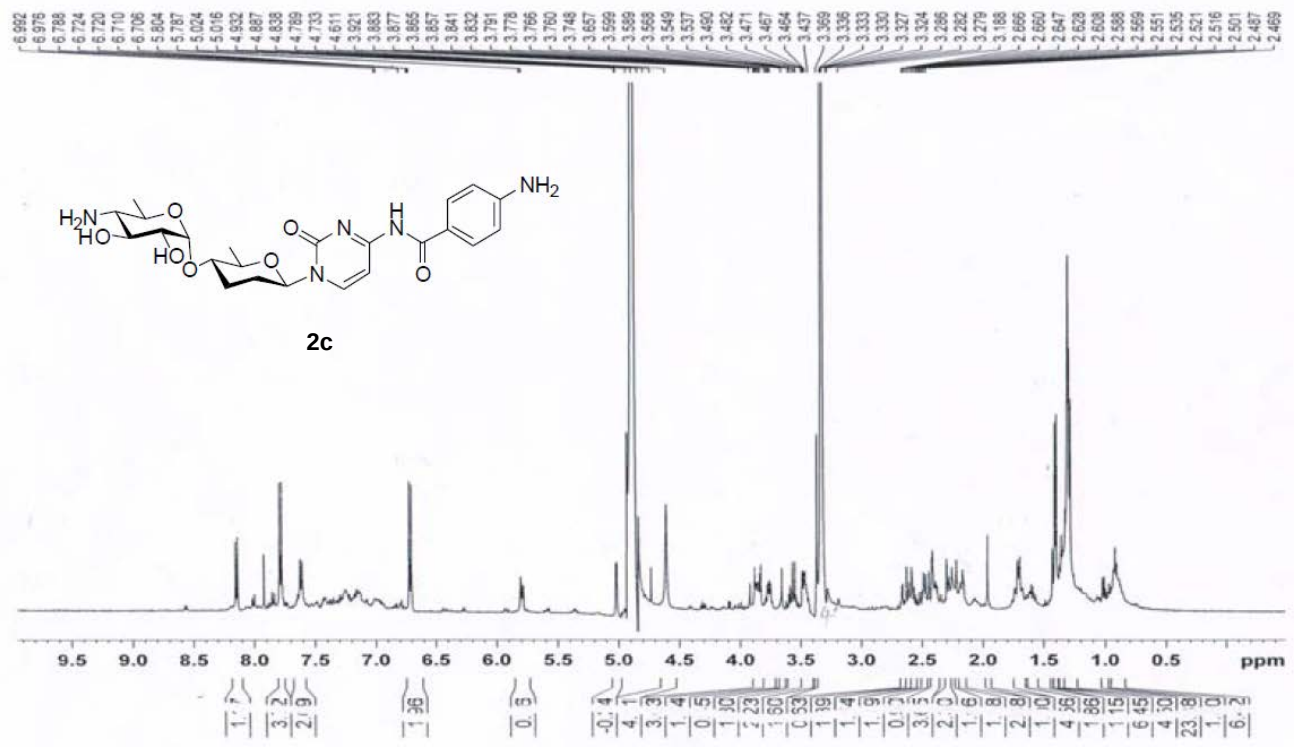
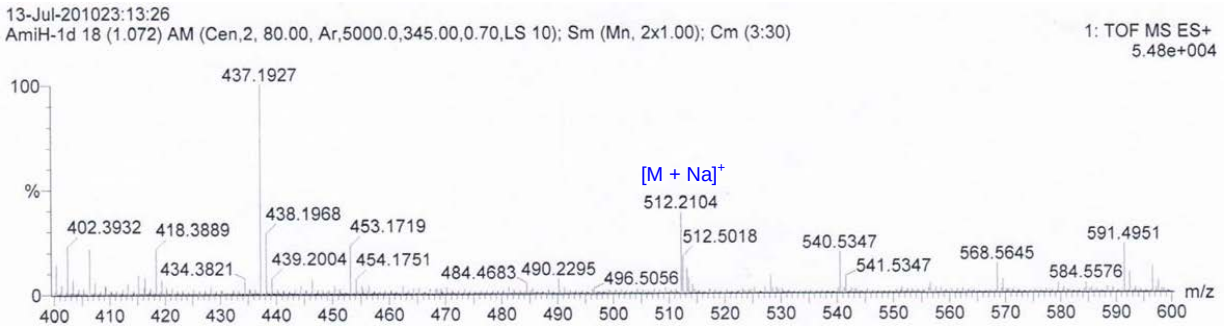
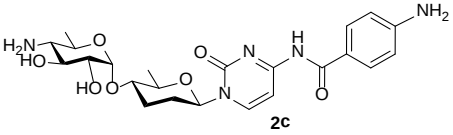


Figure S5. (B) The HRESIMS spectrum of di-demethyl-plicacetin (**2c**).





2c

Chemical Formula: $C_{23}H_{31}N_5O_7$
Calculated for $[M + Na]^+$: 512.2121

Figure S6. Spectral data of di-demethyl-cytosamine (**3c**).

Figure S6. (A) The ^{1}H NMR Spectrum of di-demethyl cytosamine (**3c**) in methanol-*d4*.

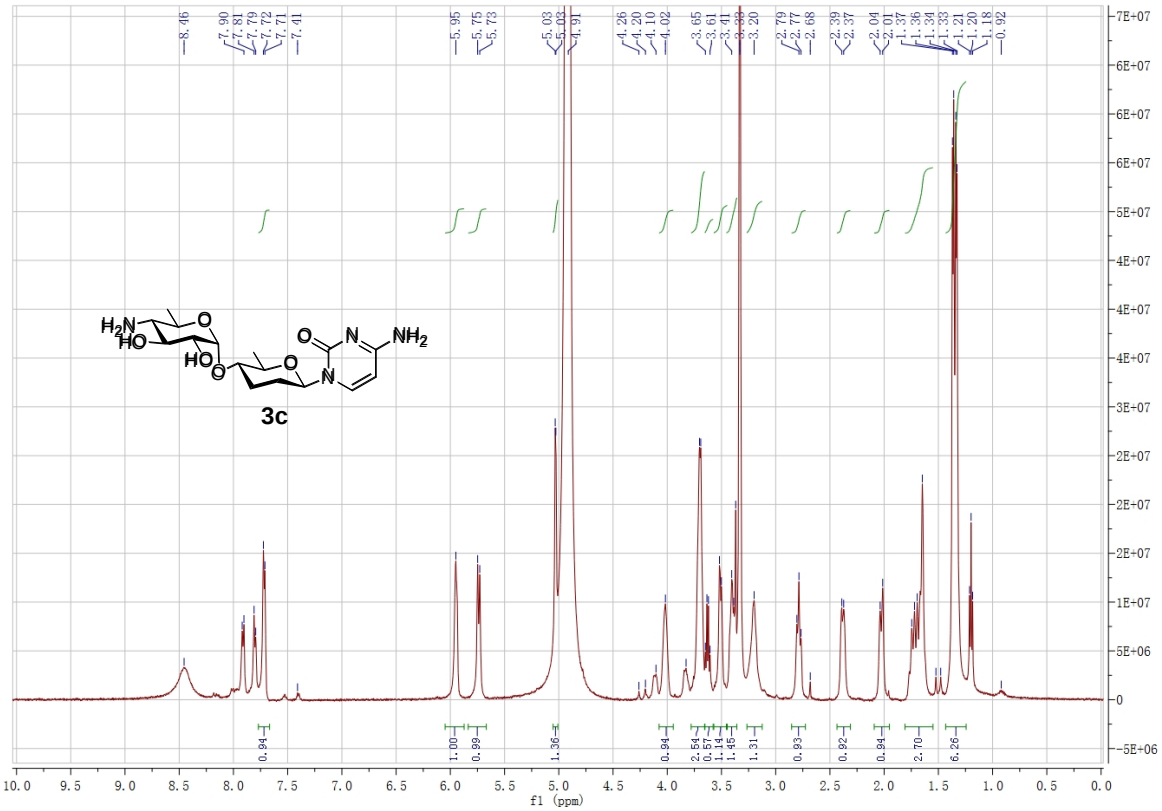


Figure S6. (B) The ^{13}C NMR spectrum of di-demethyl-cytosamine (**3c**) in methanol-*d4*.

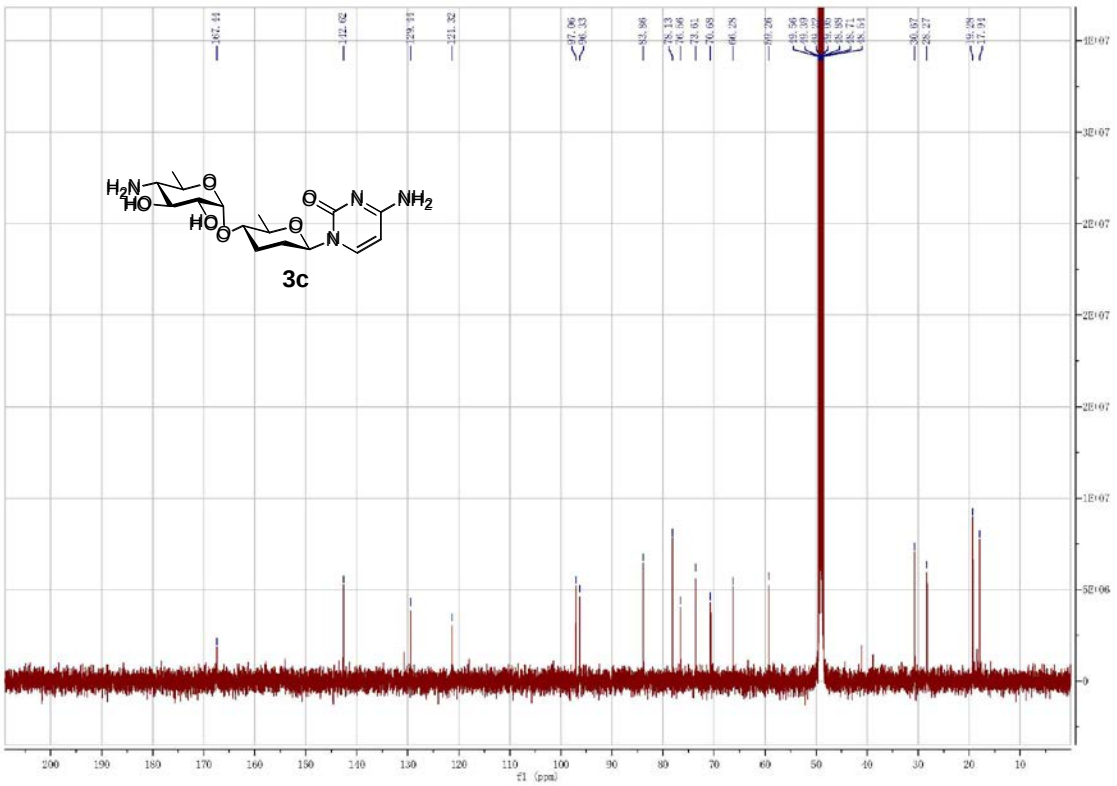


Figure S6. (C) The DEPT-135 NMR spectrum of di-demethyl-cytosamine (**3c**)

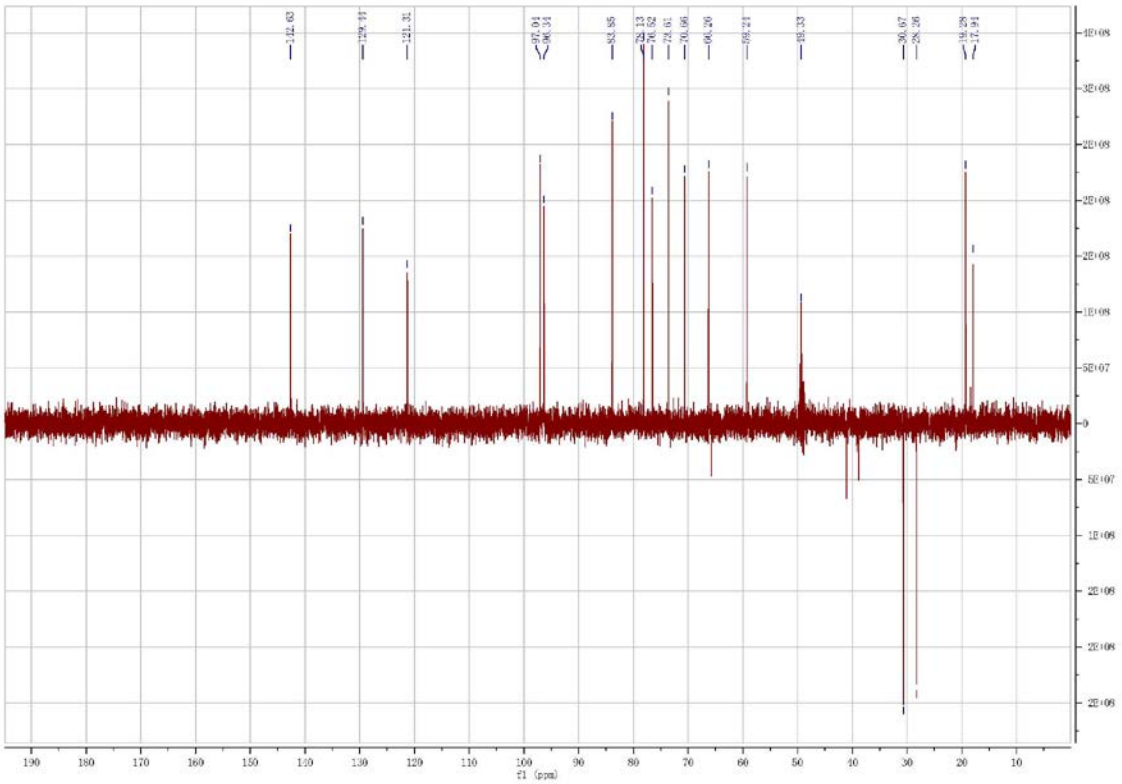


Figure S6. (D) The HSQC NMR Spectrum of di-demethyl-cytosamine (**3c**)

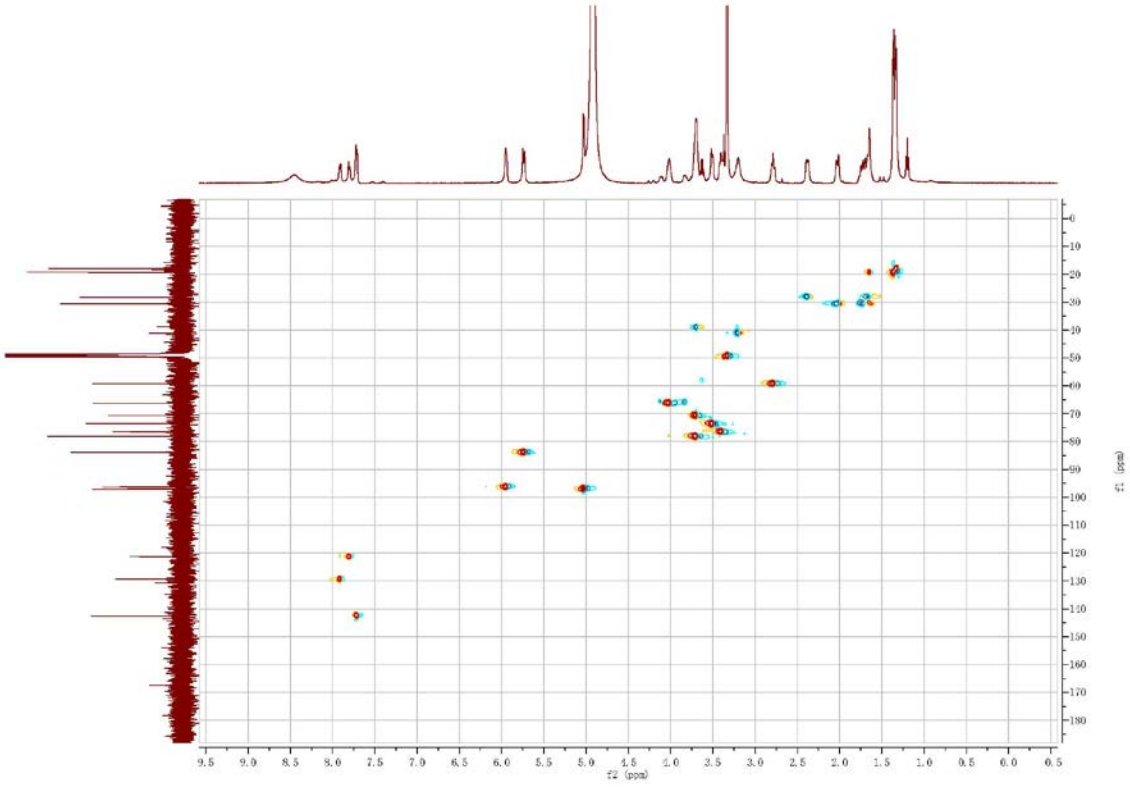


Figure S6. (E) The COSY NMR Spectrum of di-demethyl cytosamine (3c)

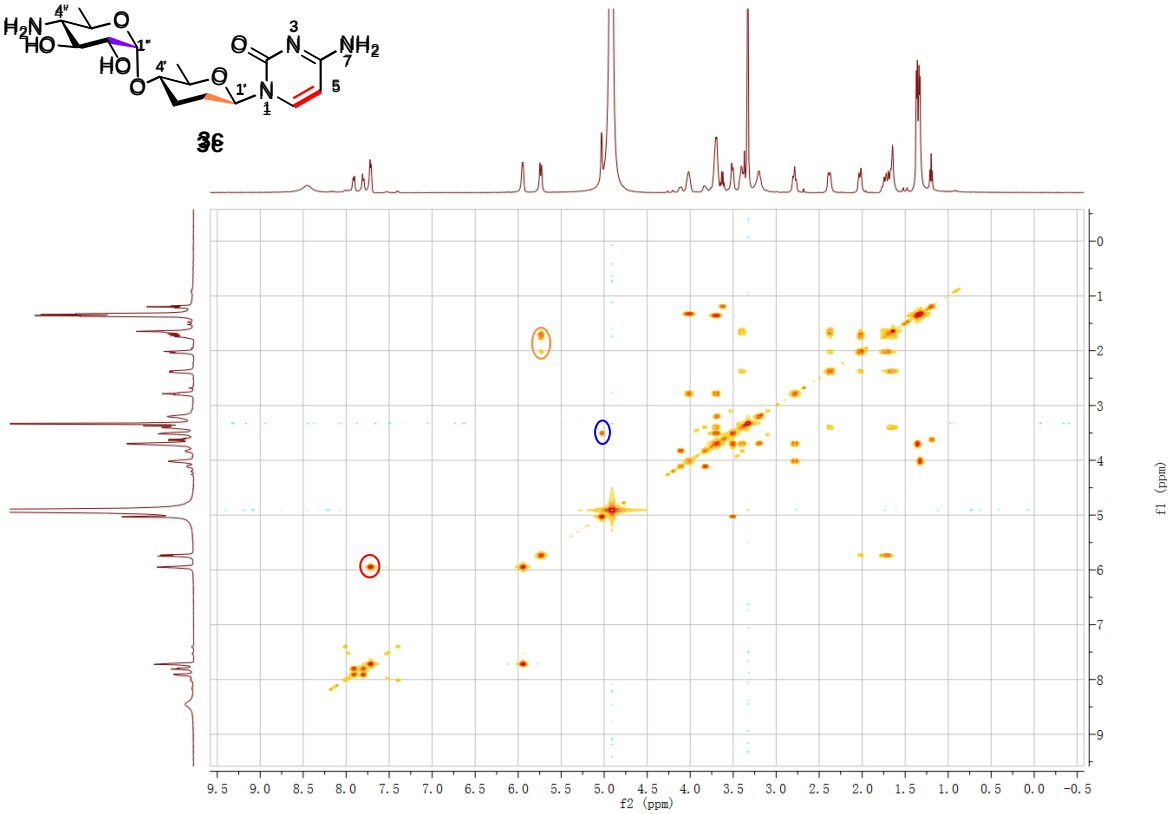


Figure S6. (F) The HMBC NMR Spectrum of di-demethyl-cytosamine (**3c**)

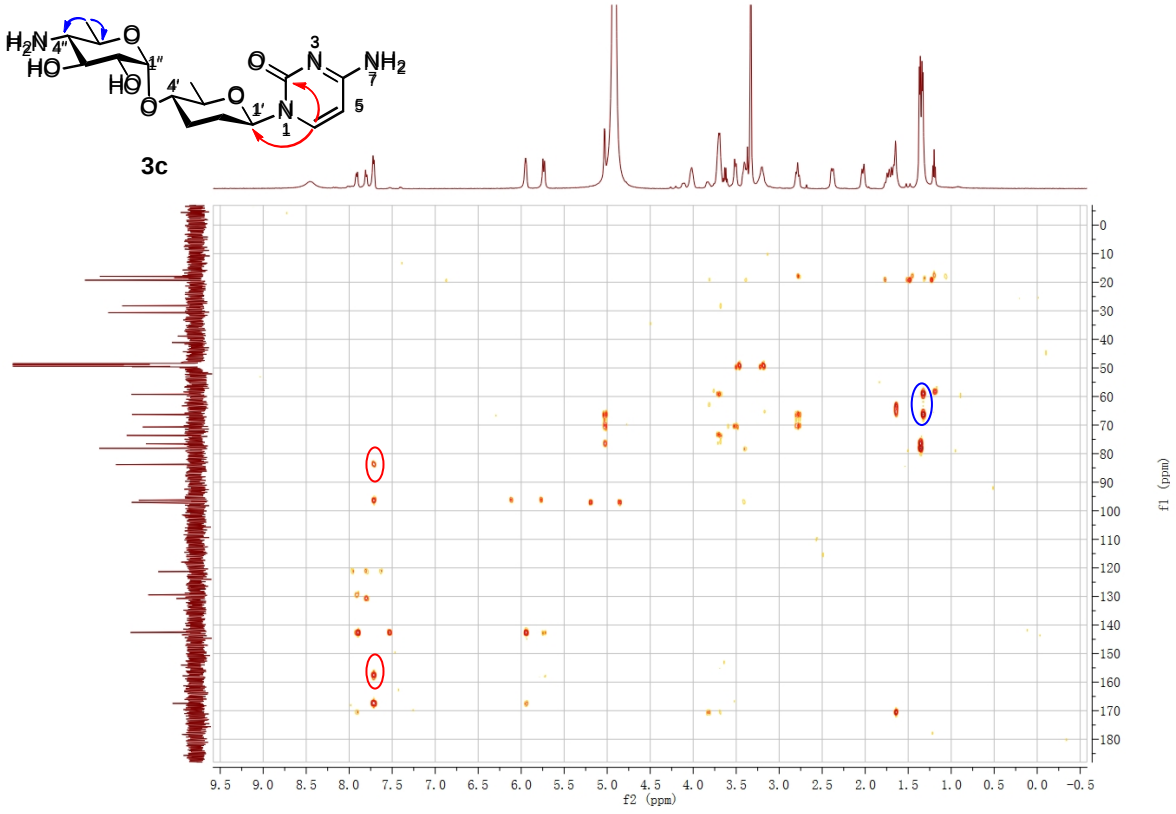


Figure S6. (G) The ROESY NMR Spectrum of di-demethyl-cytosamine (**3c**)

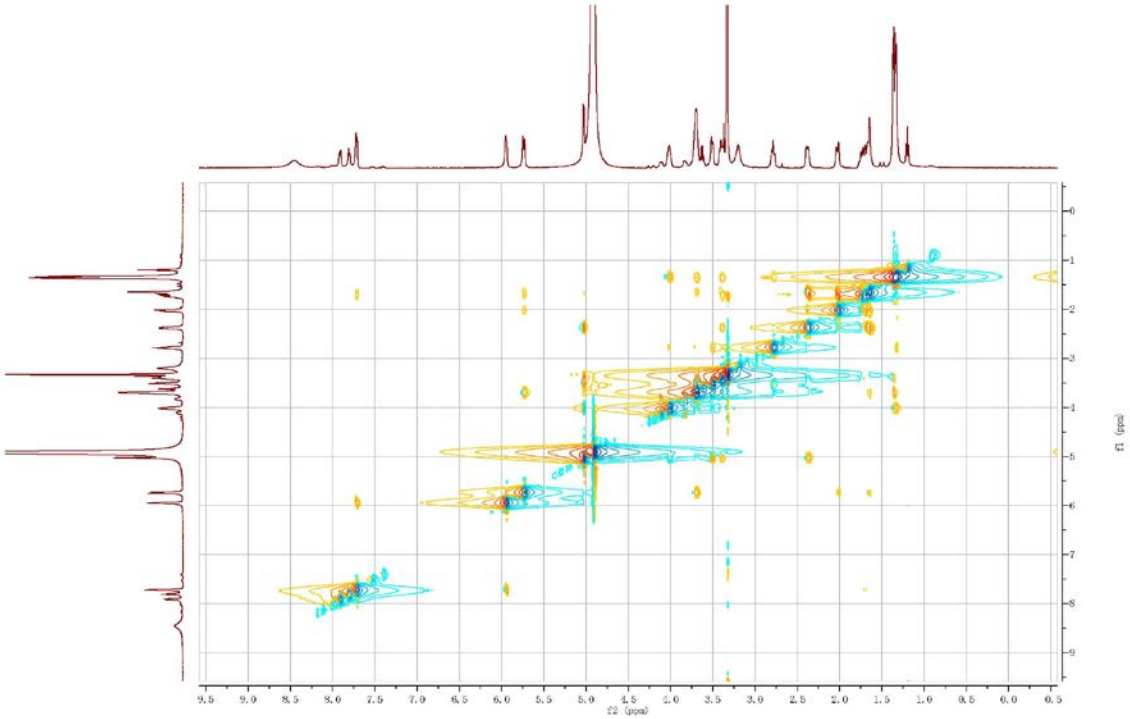
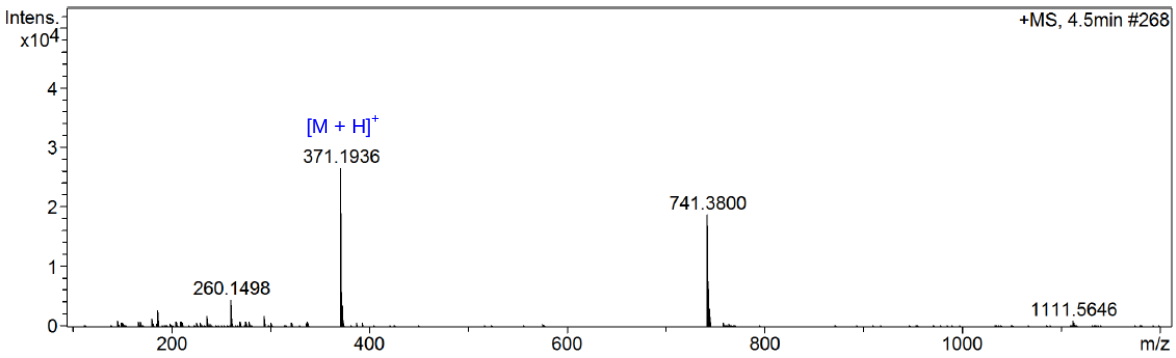
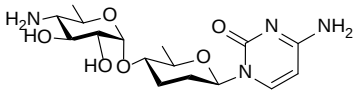


Figure S6. (H) The HRESIMS spectrum of di-demethyl-cytosamine (**3c**)





3c

Chemical Formula: $C_{16}H_{26}N_4O_6$
Calculated for $[M + H]^+$: 371.1930

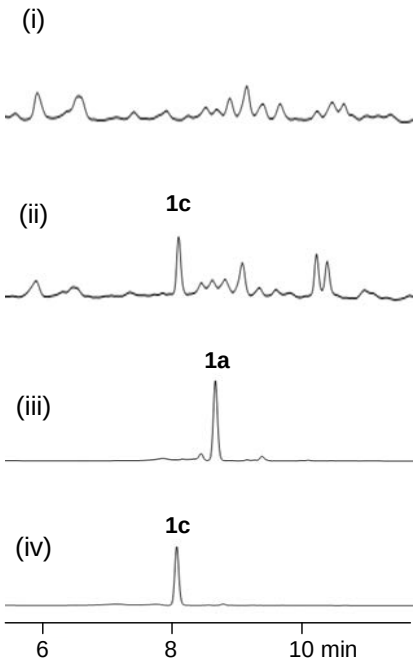


Figure S7. HPLC analysis of biotransformation of **1c** in the Δ*amiI* mutant AM1006. (i) feeding DMSO to the the Δ*amiI* mutant AM1006, in which the nucleoside 2-deoxyribosyltransferase-encoding gene *amiI* was inactivated and the production of amicetin and its analogues were abolished; (ii) feeding **1c** with a final concentration of 25 μM to AM1006; (iii) **1a** standard; (iv) **1c** standard. No conversion of **1c** to **1a** was observed in the Δ*amiI* mutant AM1006 (trace ii).

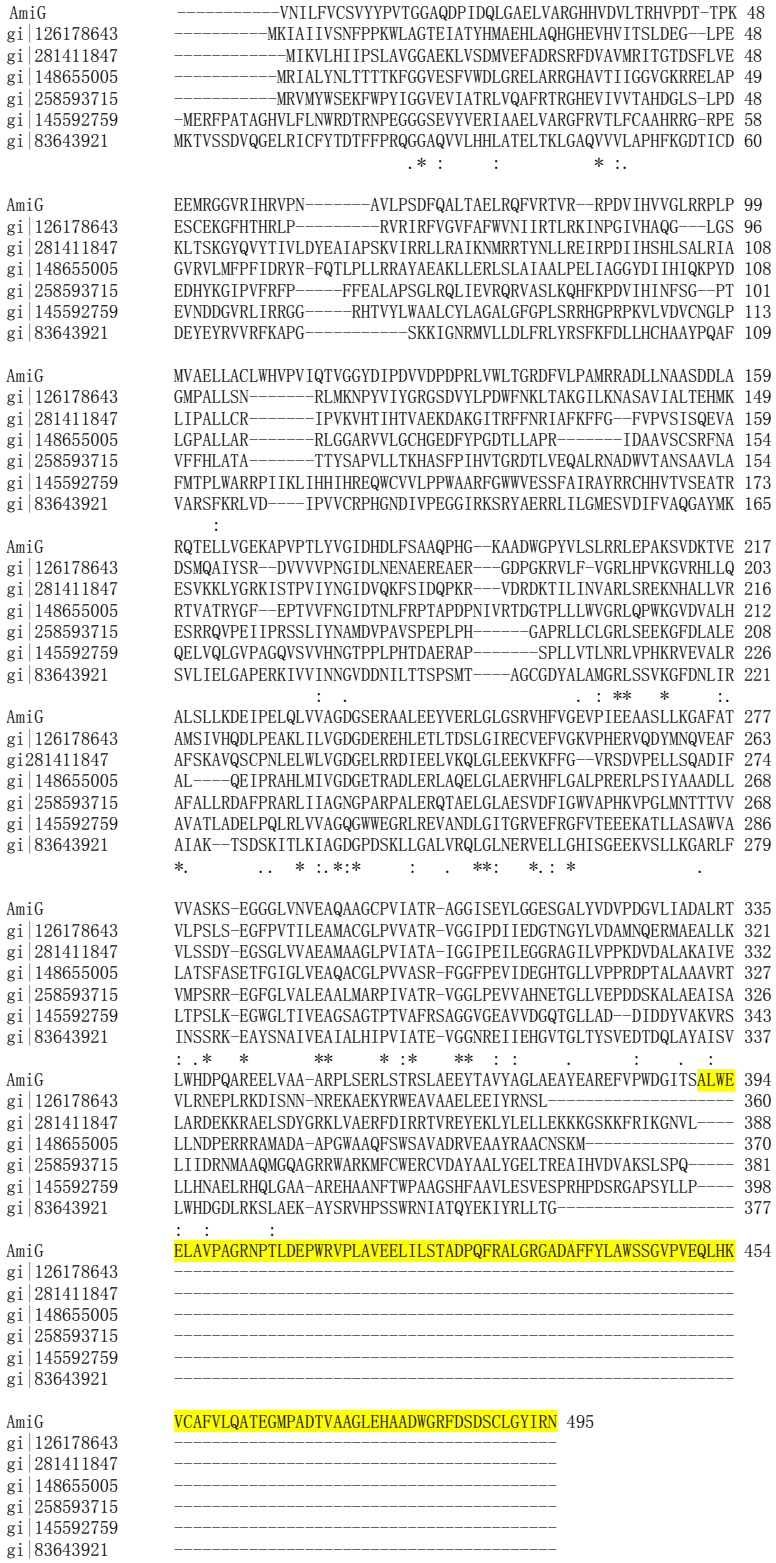


Figure S8. Alignment of AmiG with bacterial GTs from the group 1 family. AmiG from *S. vinaceus-drappus* NRRL 2363 (gi|333109246); gi|258593715, putative GT from *Candidatus Methylomirabilis oxyfera*; gi|126178643, GT from *Methanoculleus marisnigri* JR1; gi|148655005, GT from *Roseiflexus* sp. RS-1; gi|145592759, GT from *Salinispora tropica* CNB-440; gi|83643921, GT from Hahella chejuensis KCTC 2396; gi|281411847, GT from *Thermotoga naphthophila* RKU-10. The about 100 AA longer C-terminal end of AmiG was highlighted in yellow background.

Figure S9. Spectral data of de-amosaminyl-amicetin (**1**).

Figure S9. (A) The ^{1}H NMR Spectrum of de-amosaminyl-amicetin (**1**) in methanol-*d4*.

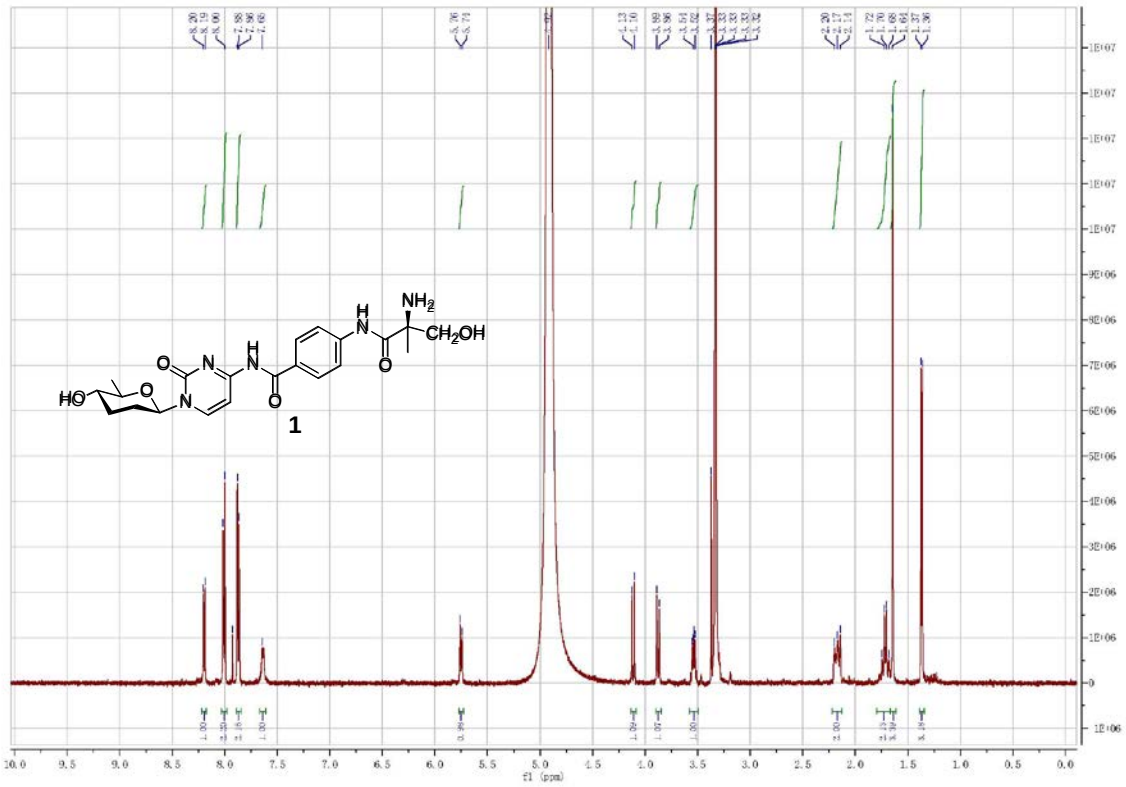


Figure S9. (B) The ^{13}C NMR Spectrum of de-amosaminyl-amicetin (**1**) in methanol-*d4*.

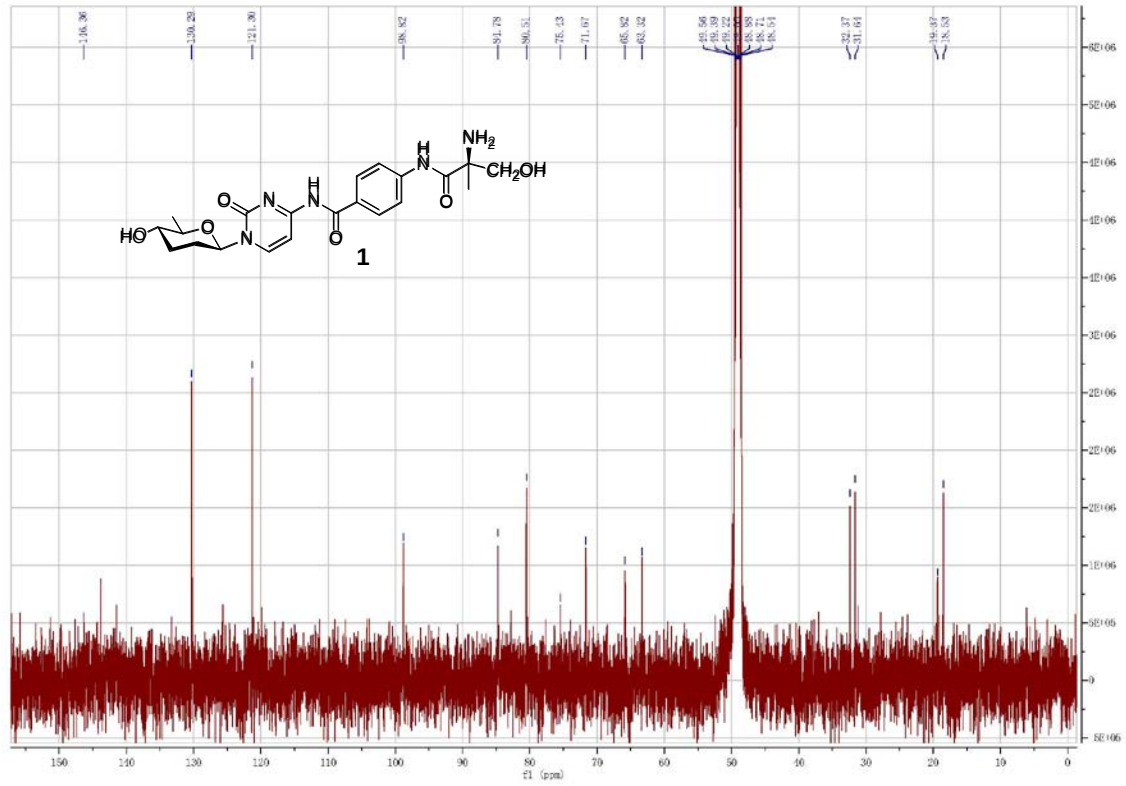


Figure S9. (C) The HSQC spectrum of de-amosaminyl-amicetin (**1**).

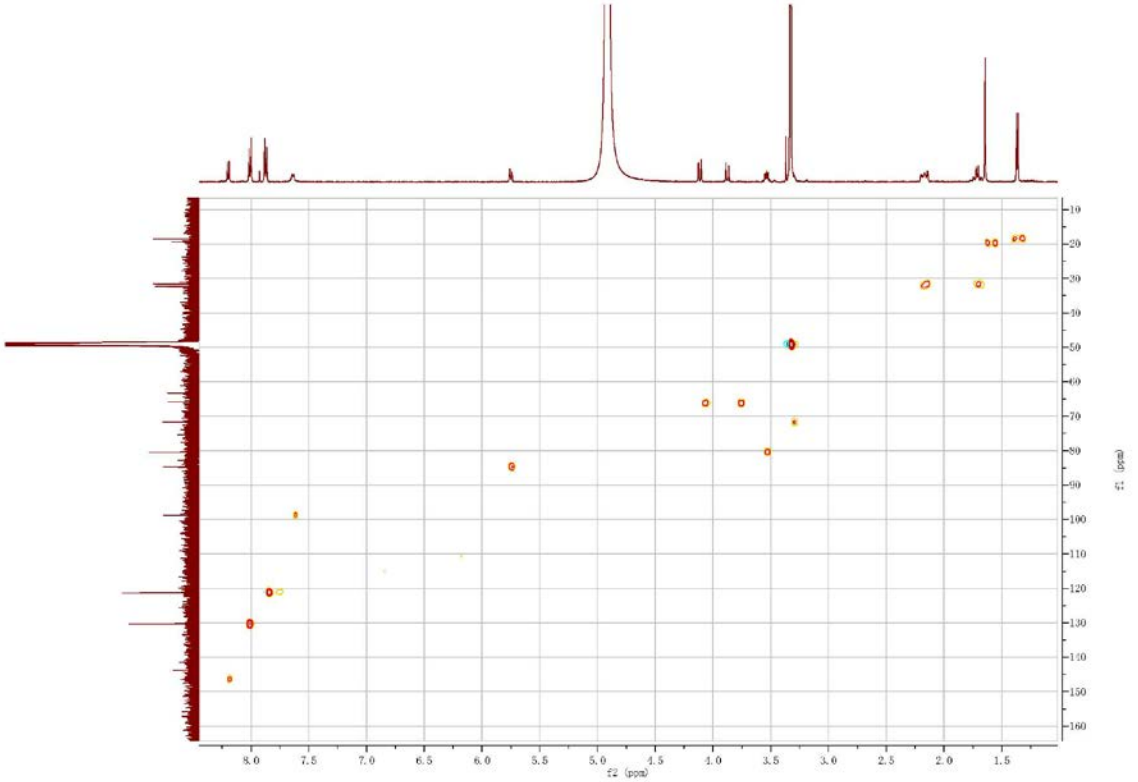


Figure S9. (D) The COSY spectrum of de-amosaminyl-amicetin (**1**).

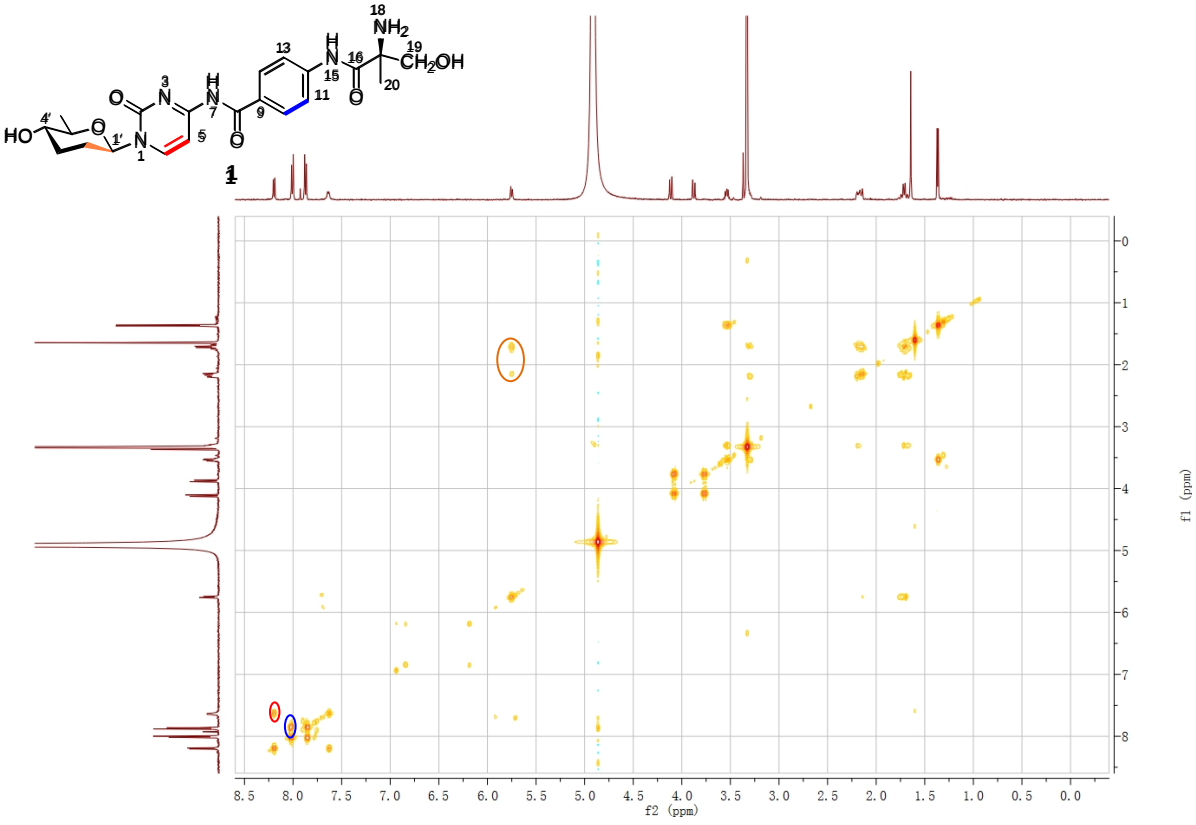


Figure S9. (E) The HMBC Spectrum of de-amosaminyl-amicetin (**1**).

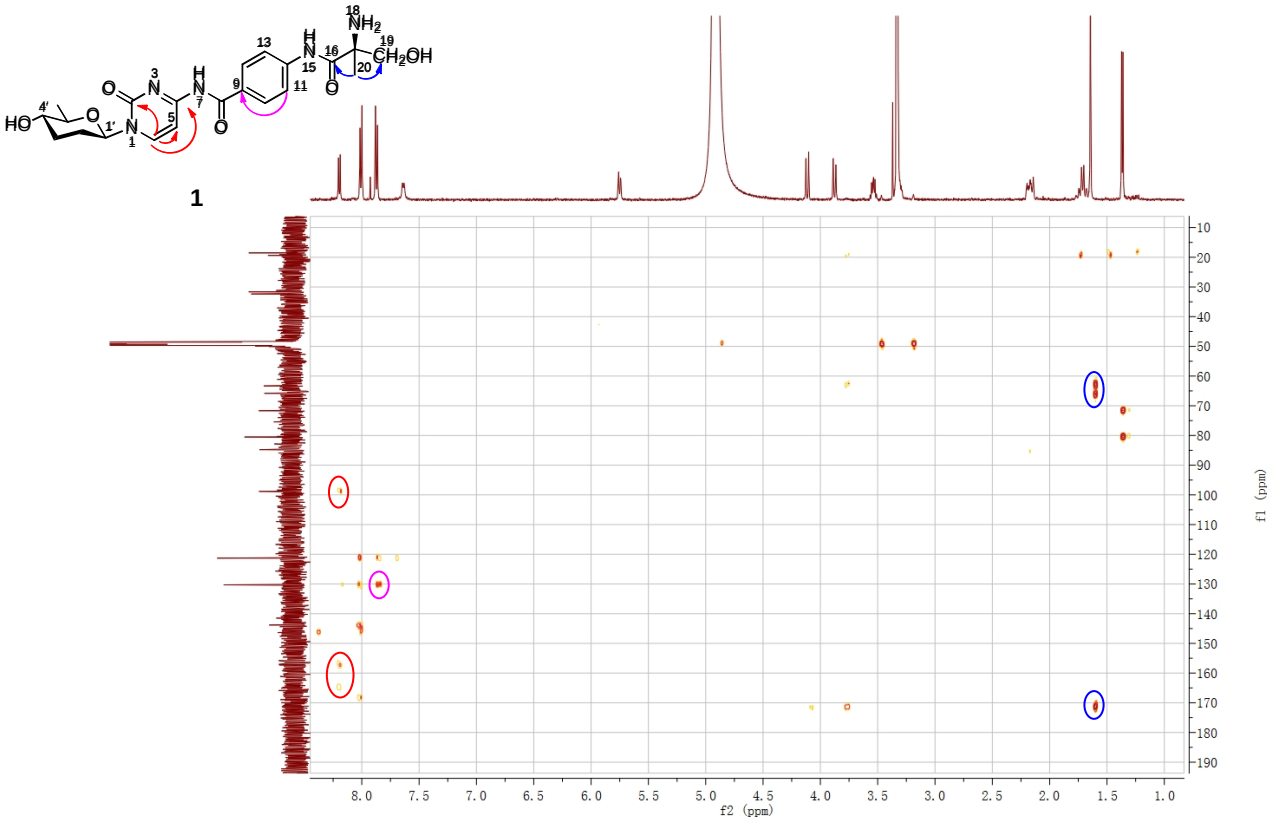
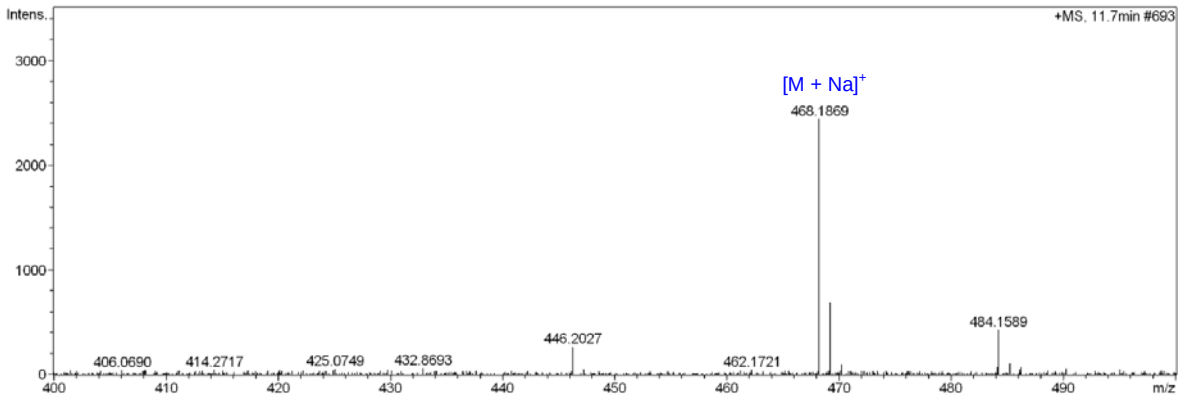
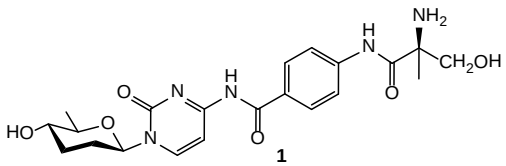


Figure S9. (F) The HRESIMS Spectrum of de-amosaminyl-amicetin (**1**).





Chemical Formula: $C_{21}H_{27}N_5O_6$
Calculated for $[M + Na]^+$: 468.1859

Figure S10. Spectral data of de-amosaminyl-cytosamine (**3**).

Figure S10. (A) The ^{1}H NMR Spectrum of de-amosaminyl-cytosamine (**3**) in methanol-*d4*.

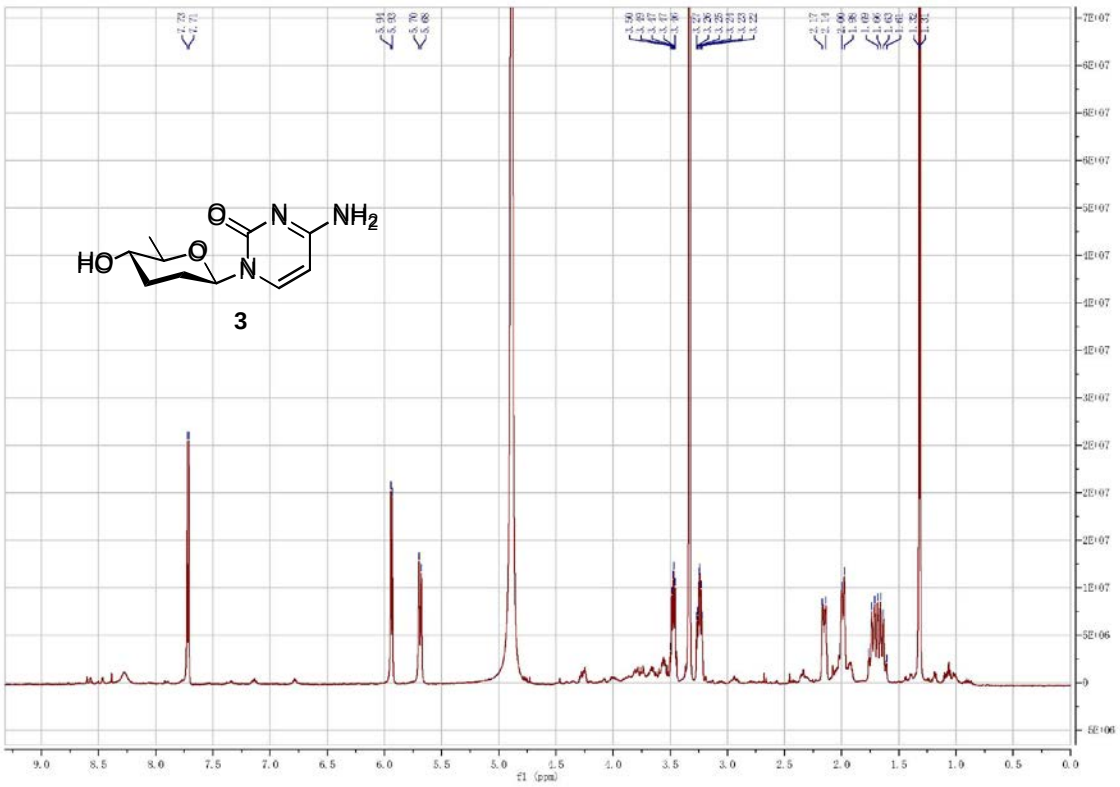


Figure S10. (B) The ^{13}C NMR spectrum of de-amosaminyl cytosamine (**3**) in methanol-*d4*.

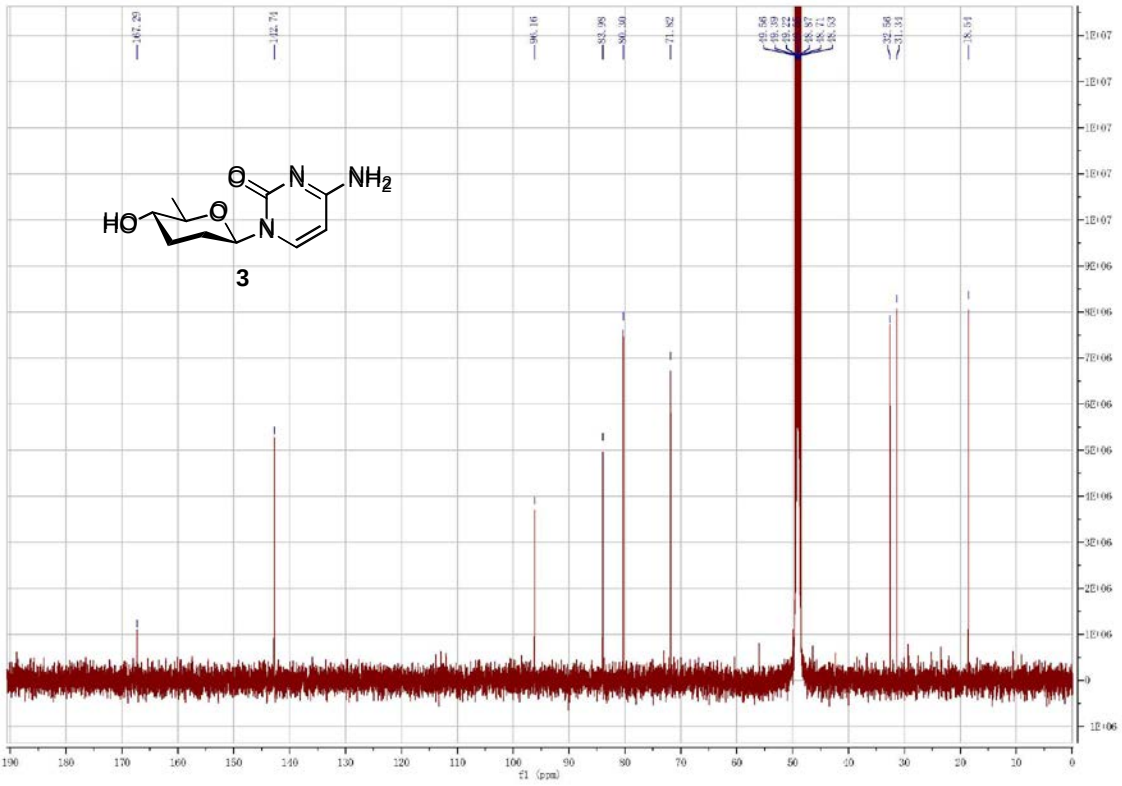


Figure S10. (C) The DEPT-135 spectrum of de-amosaminyl-cytosamine (**3**)

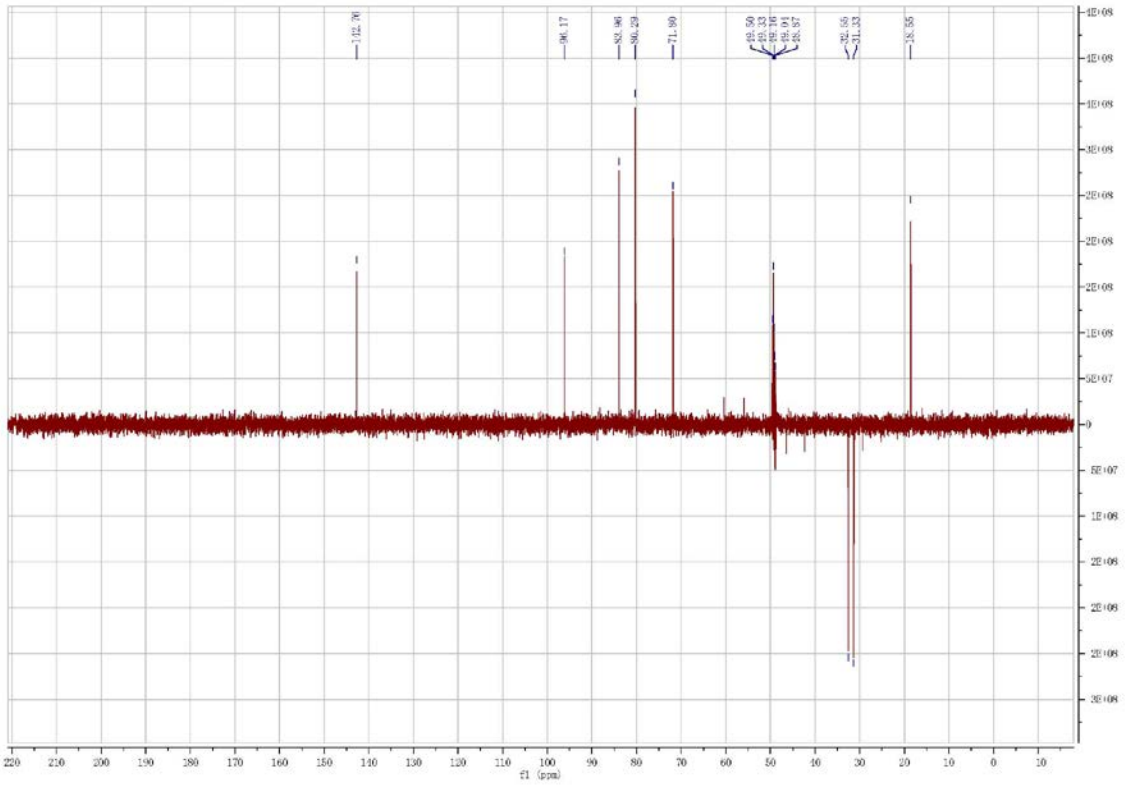
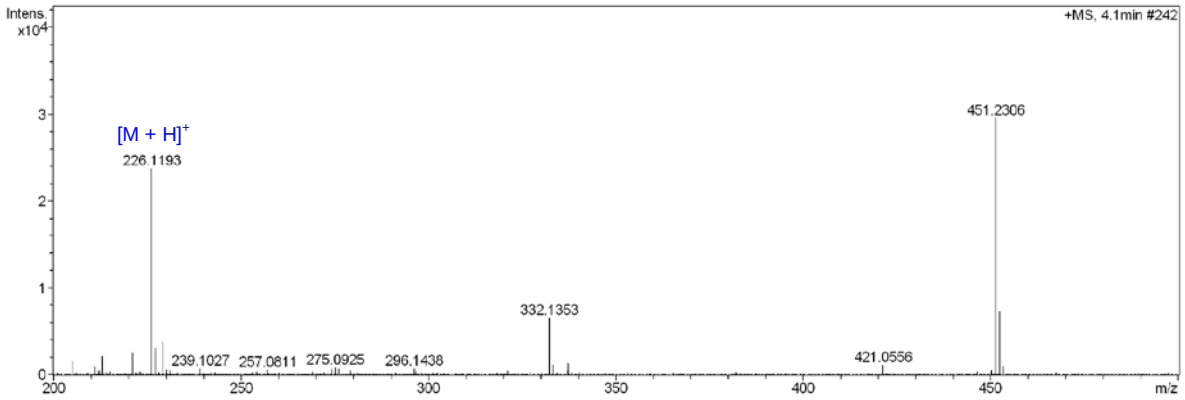
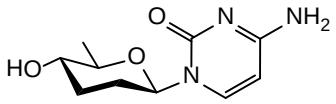


Figure S10. (D) The HRESIMS spectrum of de-amosaminyl-cytosamine (**3**)





3

Chemical Formula: $C_{10}H_{15}N_3O_3$
Calculated for $[M + H]^+$: 226.1186

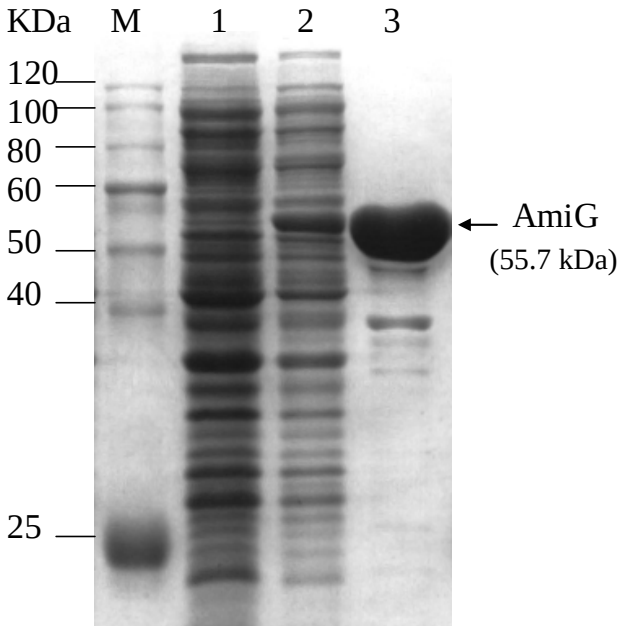
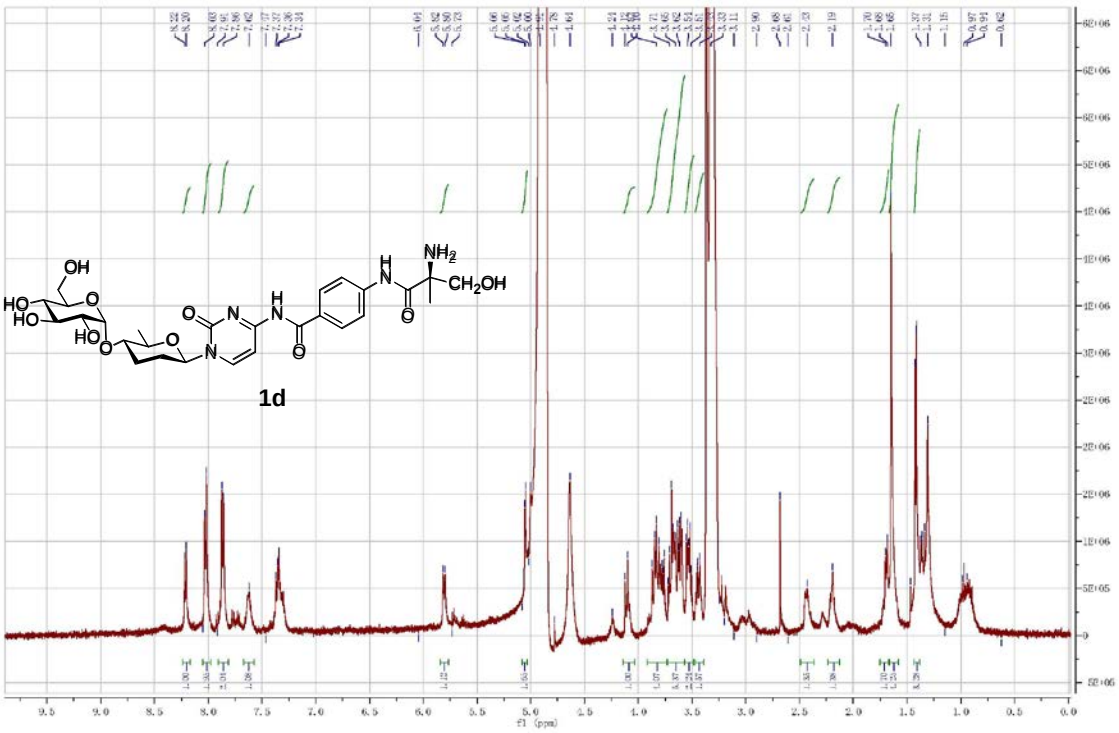


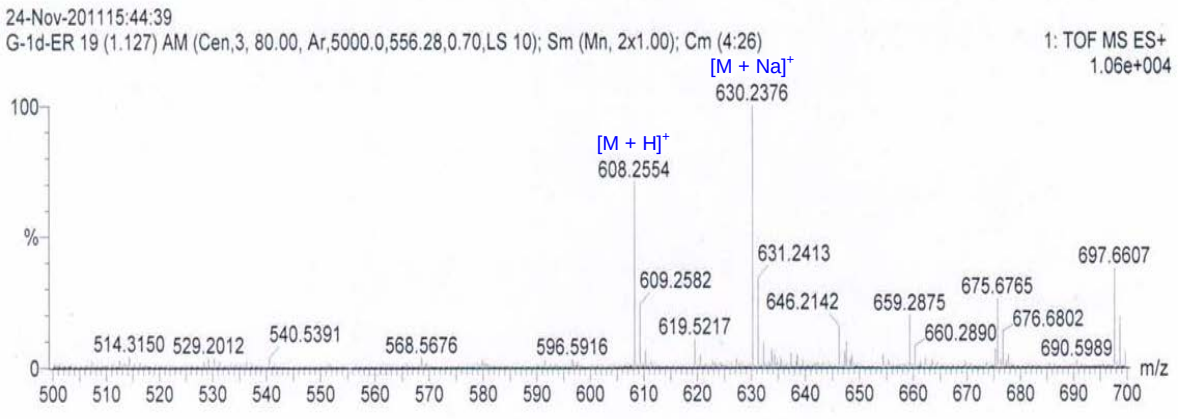
Figure S11. 12% SDS-PAGE analysis of expression and purification of AmiG. Lane M, protein marker; lane 1, supernatants of *E.coli* BL21/pET28a; lane 2, supernatants of *E.coli* BL21/pCSG3247; lane 3, purified N-$(His)_6$-tagged recombinant AmiG with a calculated MW of 55.7 kDa.

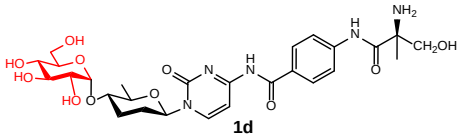
Figure S12. Spectral data of compound **1d**.

(A) The ^{1}H NMR spectrum of **1d** in methanol-*d4*.



(B) The HRESIMS spectrum of **1d**.





Chemical Formula: $C_{27}H_{37}N_5O_{11}$
Calculated for $[M + H]^+$: 608.2568
Calculated for $[M + Na]^+$: 630.2387

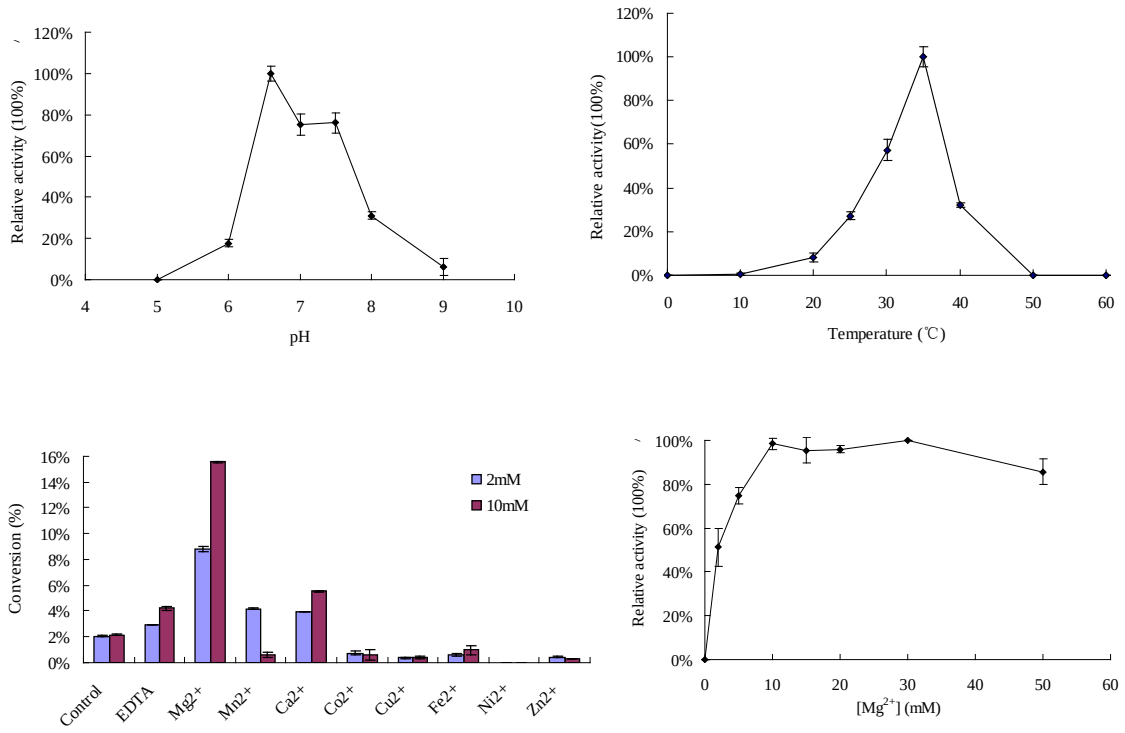


Figure S13. Optimization of AmiG reaction conditions. (**A**) Optimizing pH values: AmiG assays were carried out at 28°C for 0.5 h in buffers of pH ranging from 5.0 to 9.0 (0.1 M HAc-NaAc for pH 5.0-6.6, 50 mM MOPS buffer for pH 6.6-8.0, and 50 mM Tris-HCl buffer for pH 8.0-9.0); each 50 μL reaction mixture contained 100 μM **1**, 1 mM TDP-D-glucose and 1 μM purified AmiG, with supplementation of 2 mM $MgCl_2$. (**B**) Optimizing temperatures: AmiG assays were performed for 0.5 h in a total of 50 μL reaction mixture containing 100 μM **1**, 1 mM TDP-D-glucose, 1 μM AmiG, and 2 mM $MgCl_2$ in 50 mM MOPS buffer (pH 6.5) with temperatures ranging from 10-50°C. (**C**) Optimizing divalent cations: AmiG assays were performed at 35°C for 0.5 h in a total of 50 μL reaction mixture containing 100 μM **1**, 1 mM TDP-D-glucose, 1 μM purified AmiG in 50 mM MOPS buffer (pH 6.5) with 2 mM or 10 mM different divalent cations. Control assays were made as containing no cations or containing EDTA (2 mM or 10 mM). (**D**) Optimizing Mg^{2+} concentrations: AmiG assays were performed at 35°C for 0.5 h in a total of 50 μL reaction mixture containing 100 μM **1**, 1 mM TDP-D-glucose, 1 μM purified AmiG in 50 mM MOPS buffer (pH 6.5) with $MgCl_2$ ranging from 0-50 mM.

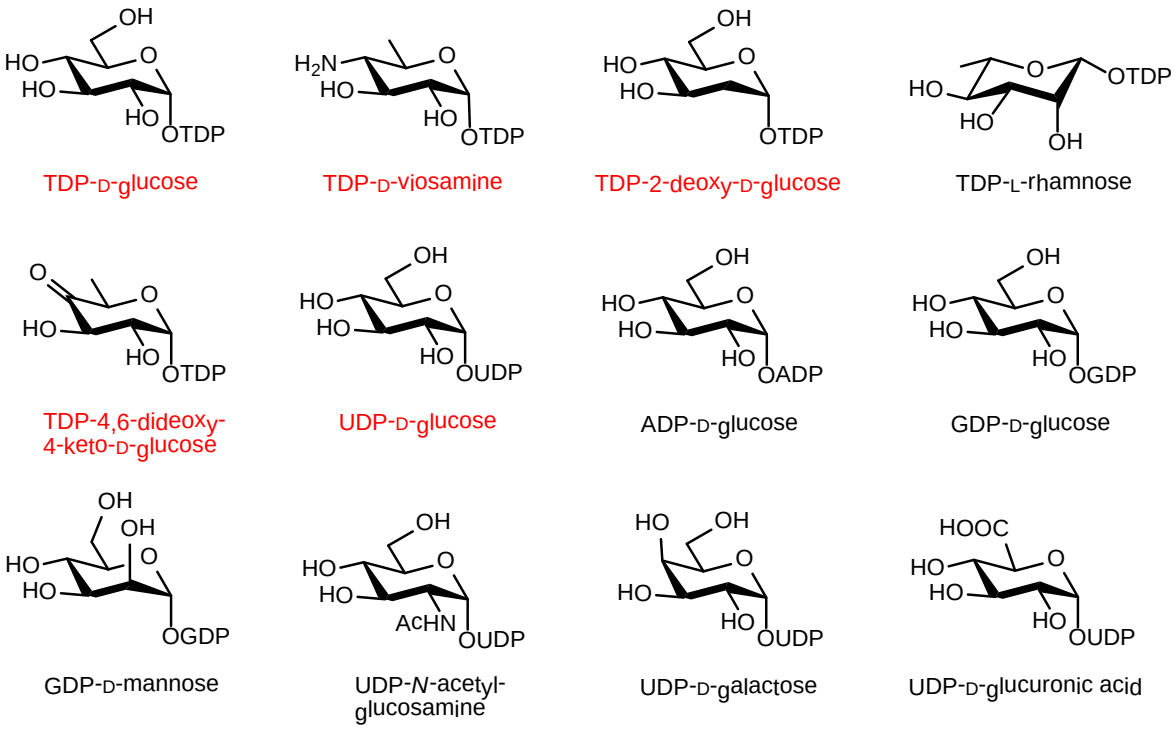
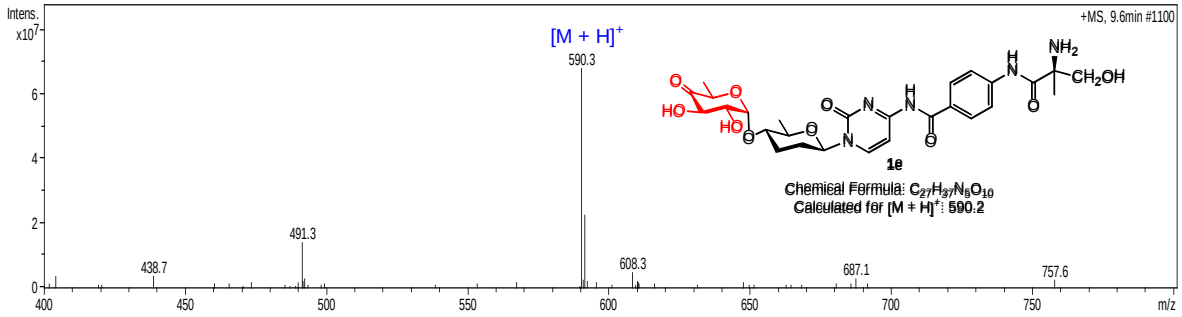


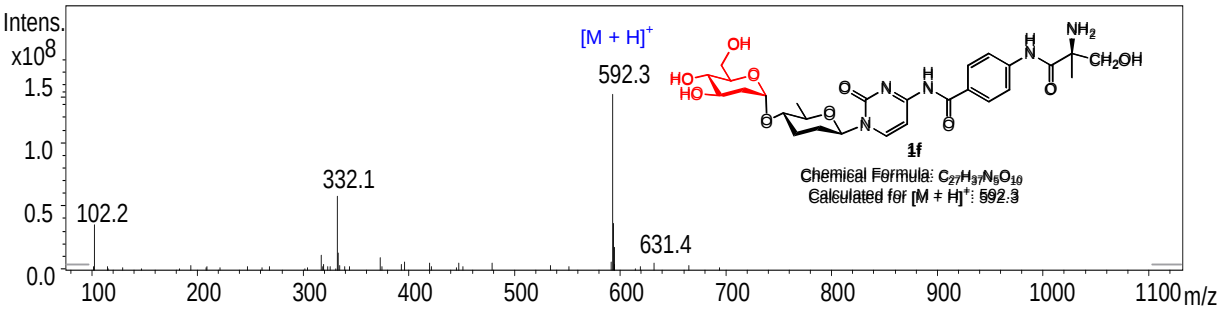
Figure S14. Structures of twelve NDP-sugars examined as an in vitro AmiG donor substrate. The red color highlights the AmiG donor substrates. TDP, UDP, GDP, CDP, ADP, TDP-D-glucose, UDP-D-glucose, ADP-D-glucose, GDP-D-glucose, GDP-mannose, UDP-D-*N*-acetylglucosamine, UDP-D-galactose, and UDP-D-glucuronic acid were purchase from Sigma Aldrich (USA), TDP-D-viosamine, TDP-4,6-dideoxy-4-keto-D-glucose, TDP-2-deoxy-D-glucose; and TDP-L-rhamnose were purchased from Carbosynth China Ltd (Suzhou, China).

Figure S15. LC-MS analyses of representative AmiG reactions.

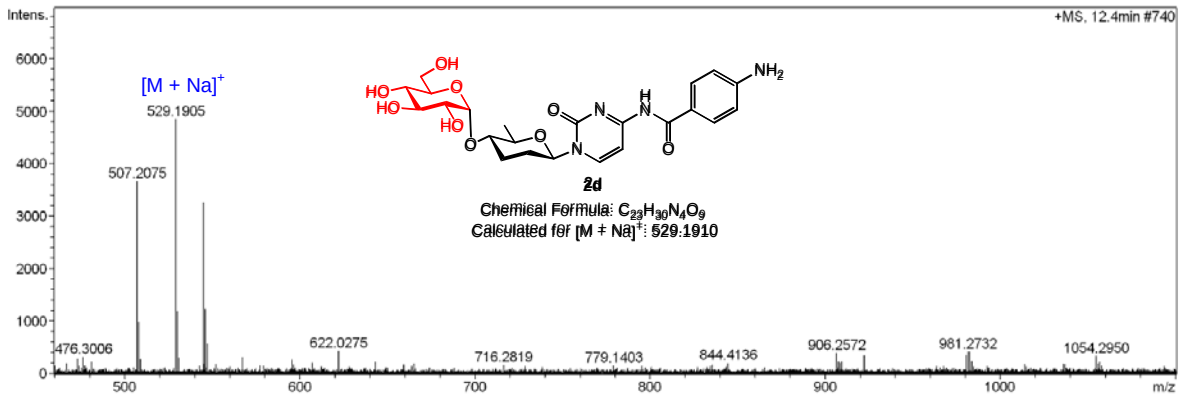
(**A**) AmiG assays with **1** and TDP-4,6-dideoxy-4-keto-D-glcuose to produce **1e**.



(**B**) AmiG assays with **1** and TDP-2-deoxy-D-glcuose to produce **1f**.



(**C**) AmiG assays with **2** and TDP-D-glucose to produce **2d**.



(**D**) AmiG assays with **2** and TDP-4,6-dideoxy-4-keto-D-glcuose to produce **2e**.

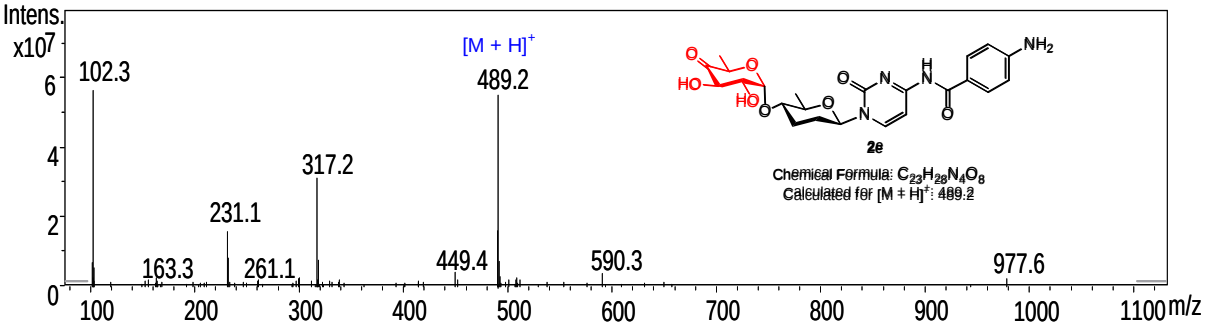
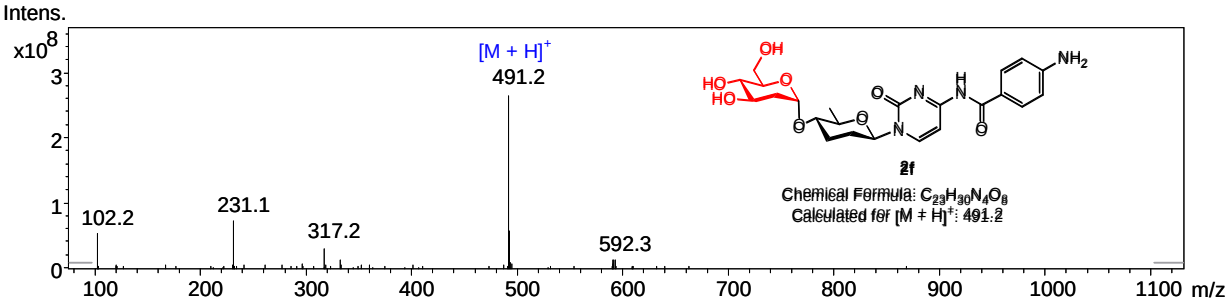
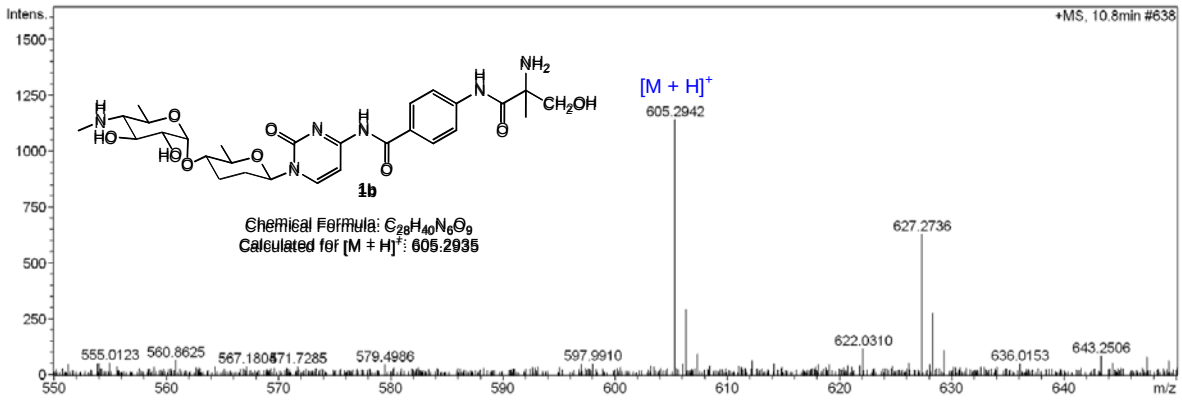


Figure S15. LC-MS analyses of representative AmiG reactions. (continued)

(**E**) AmiG assays with **2** and TDP-2-deoxy-D-glcuose to produce **2f**.



(**F**) AmiG aglycon exchange assays with **1** and **2b** in the presence of TDP to produce **1b**.



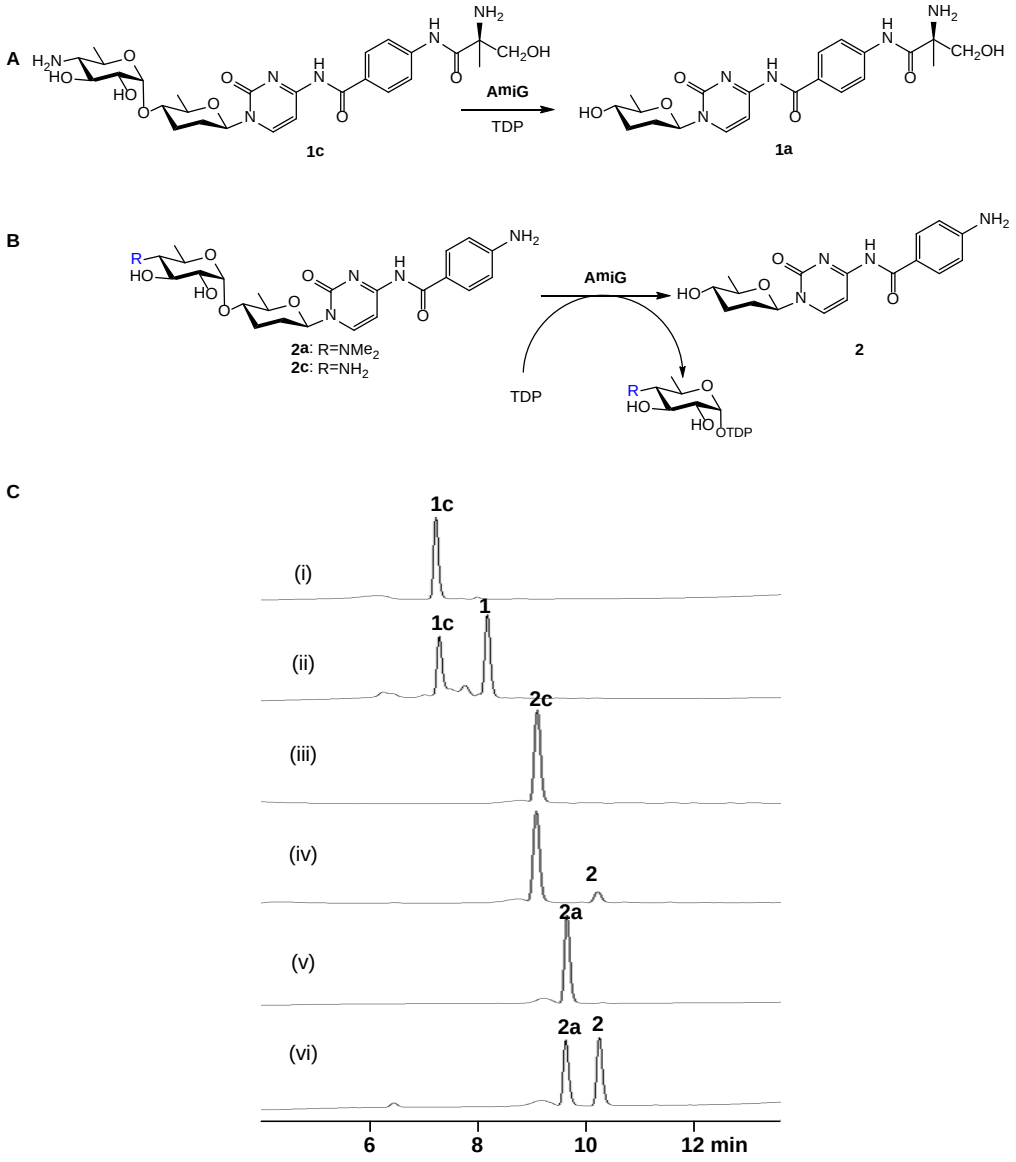
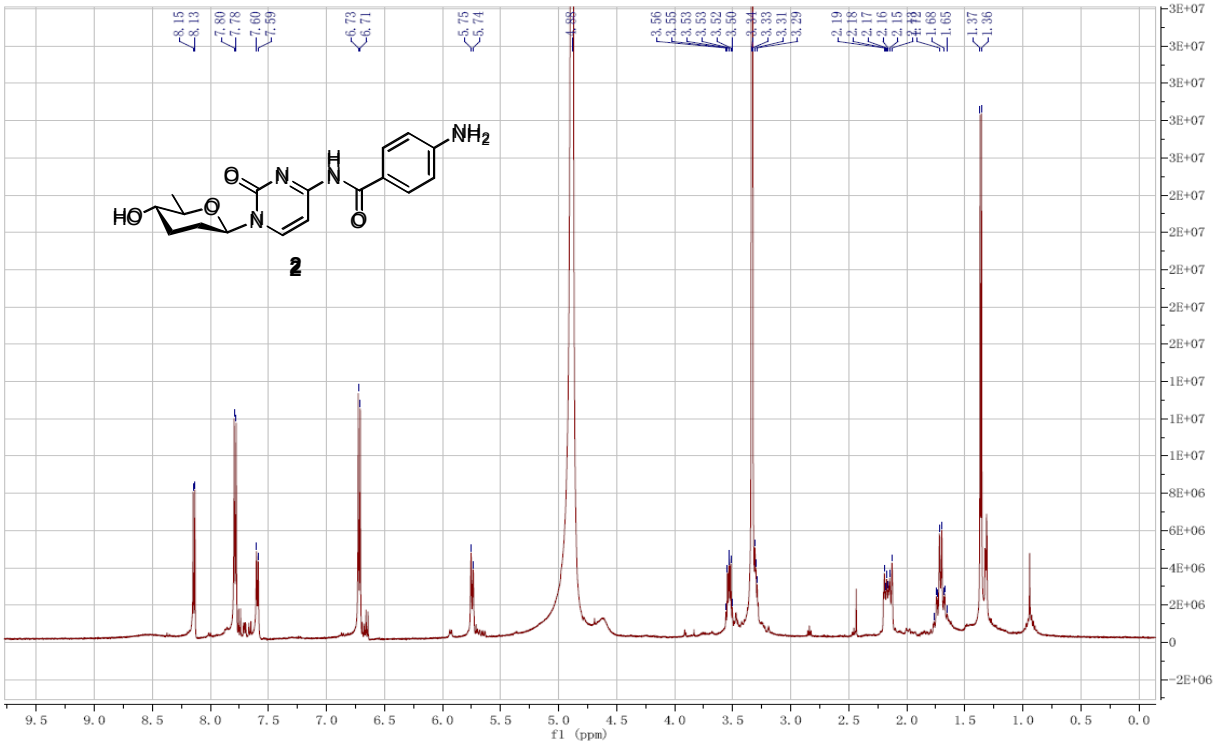


Figure S16. Schemes and HPLC analyses of AmiG-catalyzed reverse reactions. (**A**) A scheme for AmiG reverse catalysis on compound **1c**. (**B**) A scheme for AmiG reverse catalysis on compounds **2a** and **2c**. (**C**) HPLC analyses of AmiG-catalyzed reverse reactions. (i) **1c** + TDP + AmiG (cooked); (ii) **1c** + TDP + AmiG; (iii) **2c** + TDP + AmiG (cooked); (ii) **2c** + TDP + AmiG; (i) **2a** + TDP + AmiG (cooked); (ii) **2a** + TDP + AmiG. AmiG assays were performed at 35°C for 6 h in a total of 50 μL reaction mixture containing 100 μM substrate (**1c**, **2c** or **2a**), 2 mM TDP, 3.3 μM AmiG in 50 mM MOPS buffer (pH 6.5) with 10 mM $MgCl_2$.

Figure S17. Spectral data of de-amosaminyl-plicacetin (**2**).

(A) The ^{1}H NMR Spectrum of de-amosaminyl-plicacetin (**2**) in methanol-*d4*.



(B) The ^{13}C NMR spectrum of de-amosaminyl plicacetin (**2**) in methanol-*d4*.

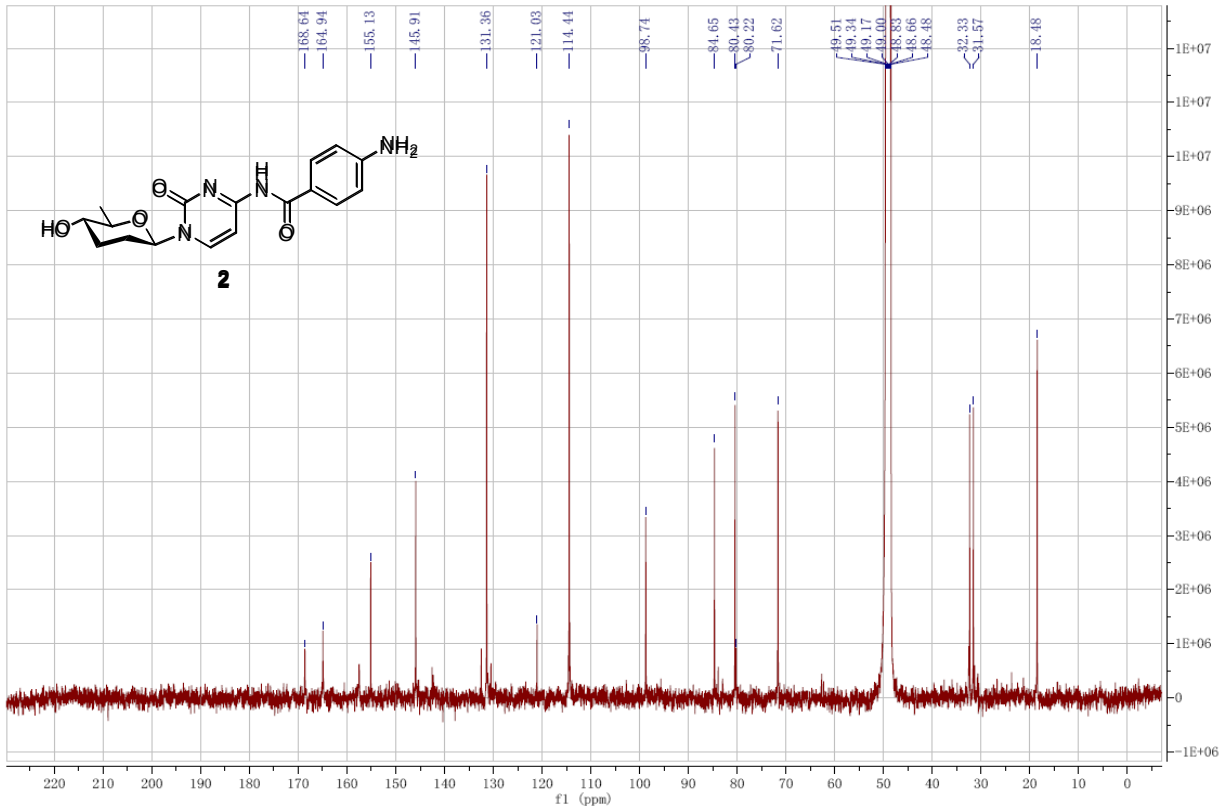
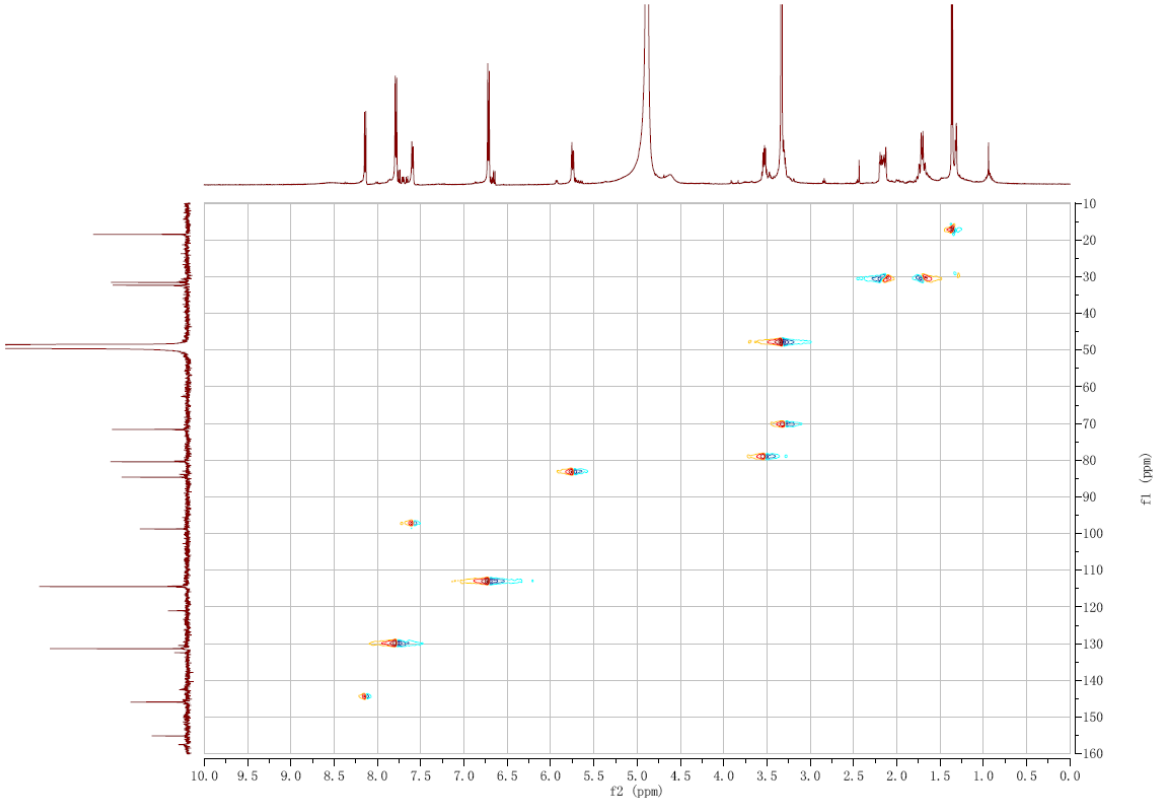


Figure S17. Spectral data of de-amosaminyl-plicacetin (**2**). (continued)

(C) The HSQC spectrum of de-amosaminyl-plicacetin (**2**).



(D) The COSY spectrum of de-amosaminyl-plicacetin (**2**).

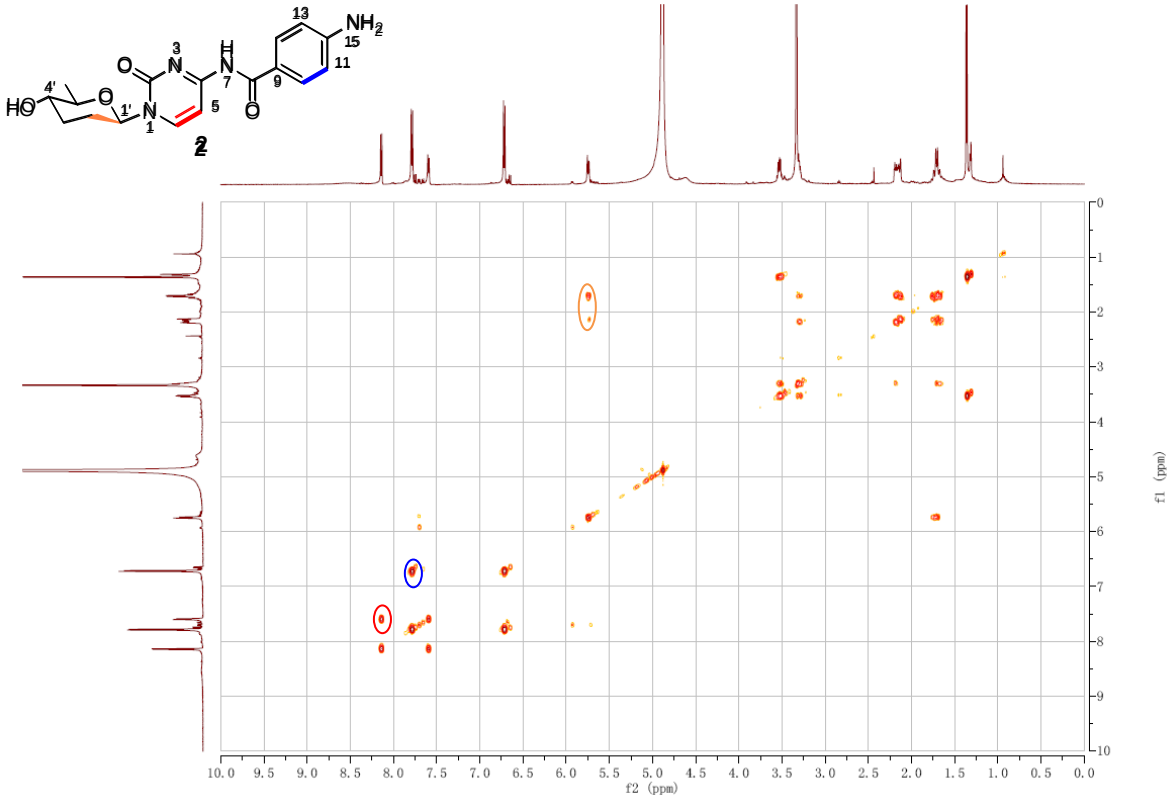
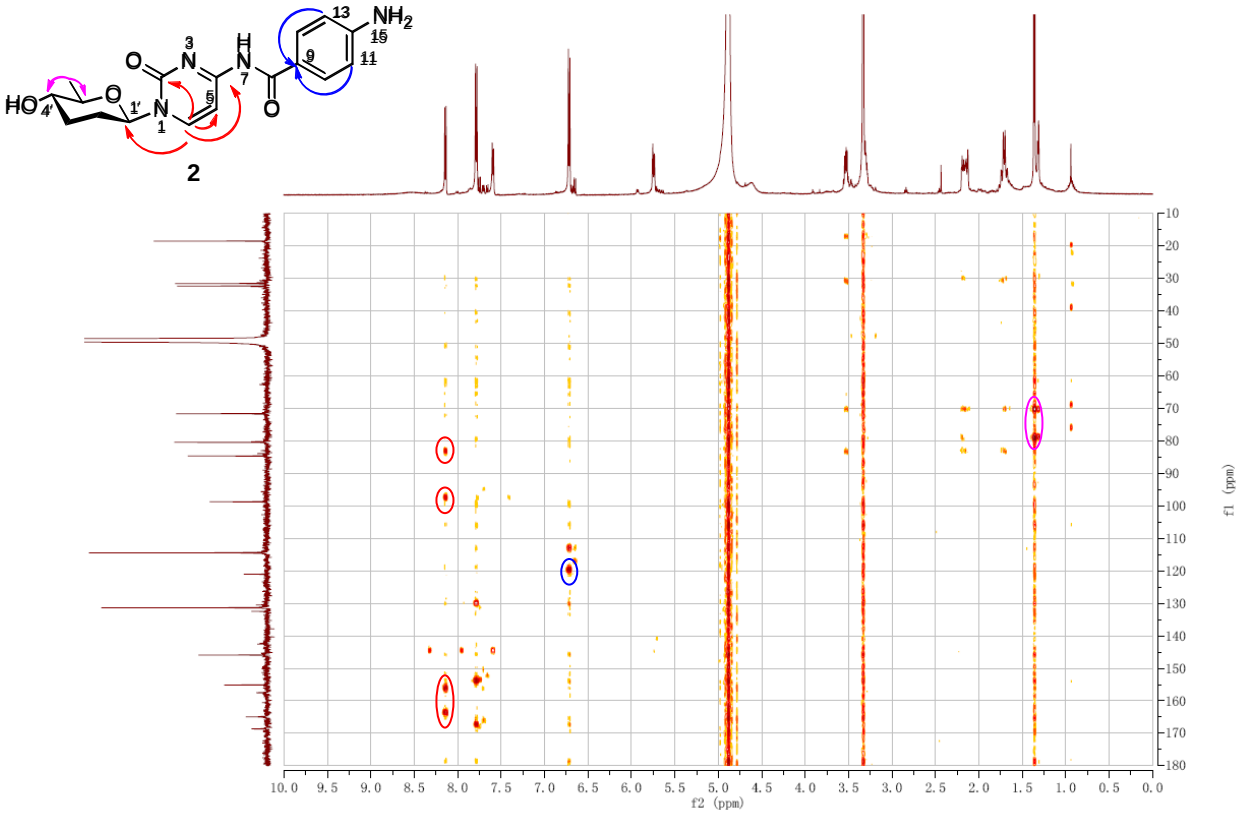
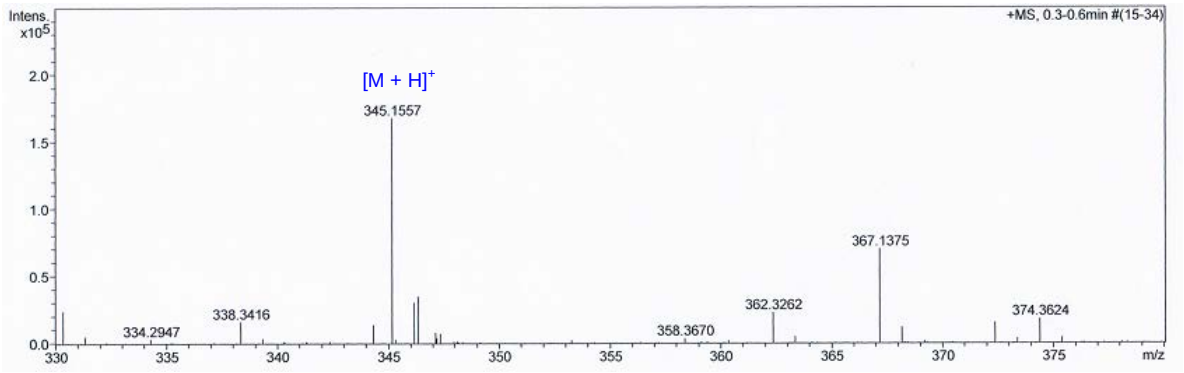


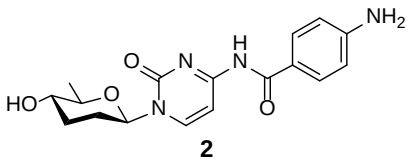
Figure S17. Spectral data of de-amosaminyl-plicacetin (**2**). (continued)

(E) The HMBC spectrum of de-amosaminyl-plicacetin (**2**).



(F) The HRESIMS spectrum of de-amosaminyl-plicacetin (**2**)



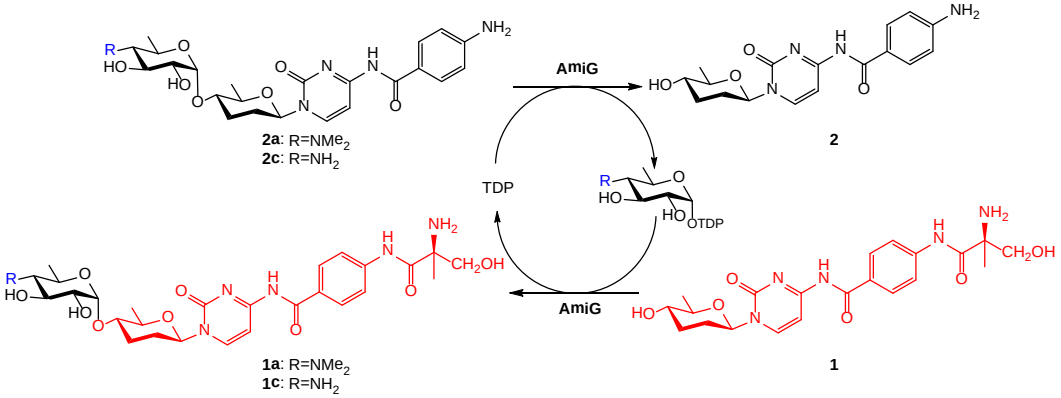


2

Chemical Formula: $C_{17}H_{20}N_4O_4$

Calculated for $[M + H]^+$: 345.1563

A



B

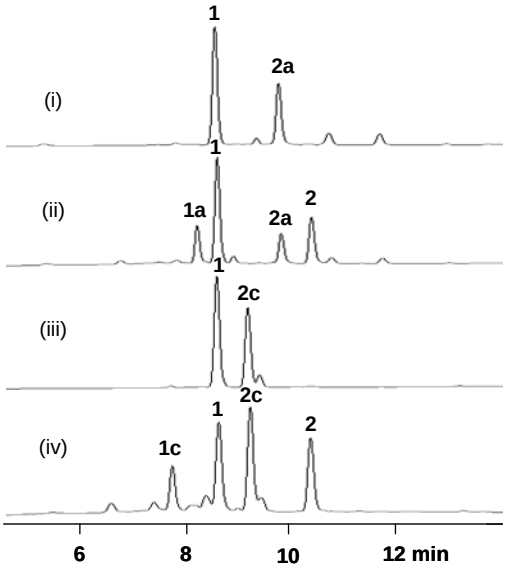


Figure S18. Scheme and HPLC analyses of AmiG catalyzed aglycon exchange reactions. (**A**) A Scheme for AmiG catalyzed aglycon exchange reactions. (**B**) HPLC analyses of AmiG-catalyzed aglycon exchange reactions. (i) **1** + **2a** + TDP + AmiG (cooked); (ii) **1** + **2a** + TDP + AmiG; (iii) **1** + **2c** + TDP + AmiG (cooked); (ii) **1** + **2c** + TDP + AmiG. AmiG assays were performed at 35°C for 6 h in a total of 50 μL reaction mixture containing 80 μM sugar donor substrate **2a** (or **2c**) substrate, 100 μM aglycon substrate **1**, 2 mM TDP, 3.3 μM AmiG in 50 mM MOPS buffer (pH 6.5) with 10 mM $MgCl_2$.

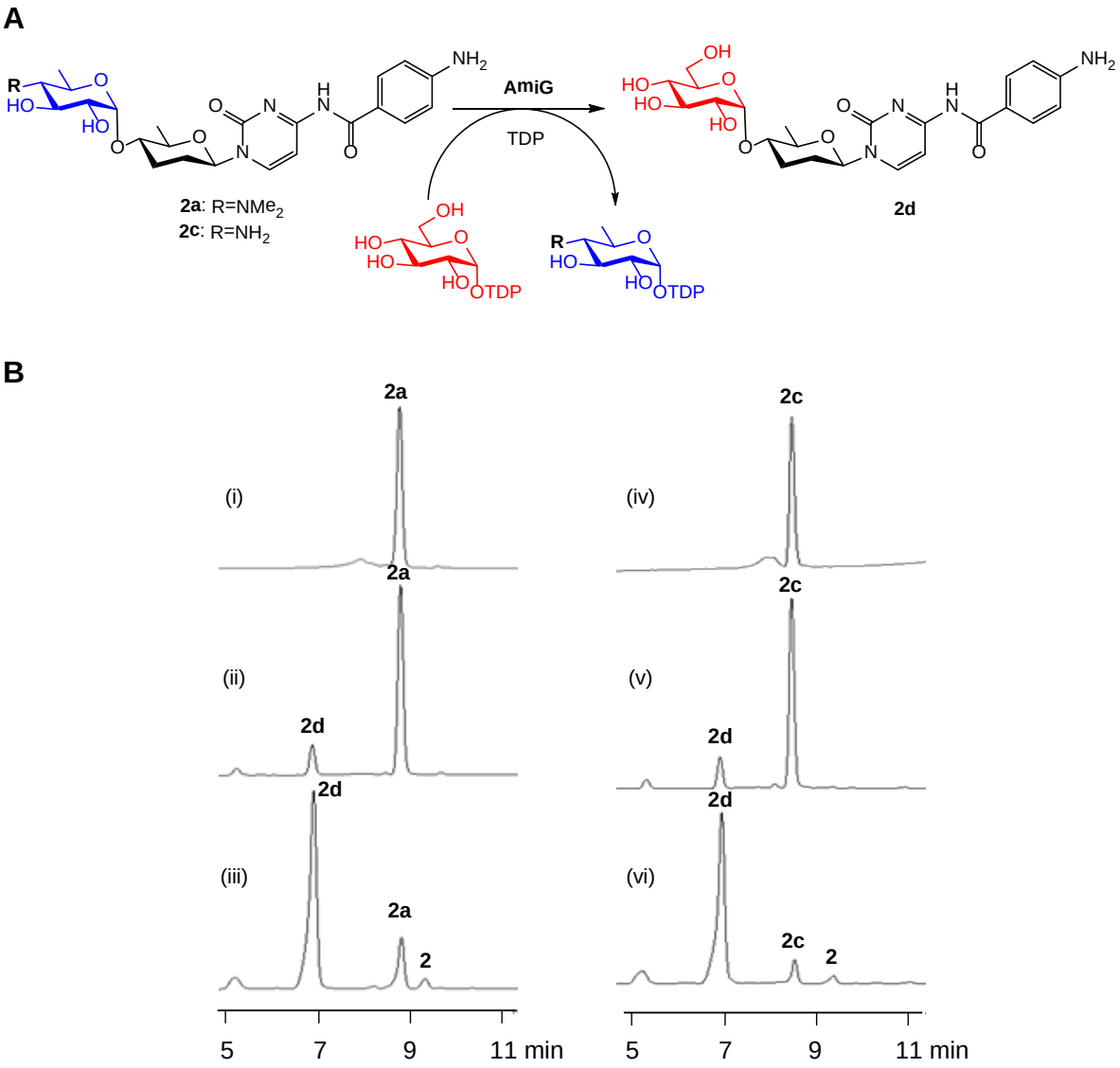


Figure S19. Scheme and HPLC analyses of AmiG catalyzed sugar exchange reactions. (**A**) A Scheme for AmiG catalyzed sugar exchange reaction using **2a** (or **2c**) and TDP-D-glucose as co-substrates. (**B**) HPLC analyses of AmiG-catalyzed reverse reactions. (i) **2a** + TDP-D-glucose + AmiG (cooked); (ii) **2a** + TDP-D-glucose + AmiG; (iii) **2a** + TDP-D-glucose + TDP + AmiG; (iv) **2c** + TDP-D-glucose + AmiG (cooked); (v) **2c** + TDP-D-glucose + AmiG; (vi) **2c** + TDP-D-glucose + TDP + AmiG; AmiG assays were performed at 35°C for 6 h in a total of 50 μL reaction mixture containing 100 μM **2a** (or **2c**), 5 mM TDP-D-glucose, 3.3 μM AmiG in 50 mM MOPS buffer (pH 6.5) with 10 mM $MgCl_2$ in the absence (ii and v) or in the presence of 1 mM TDP (iii and vi).

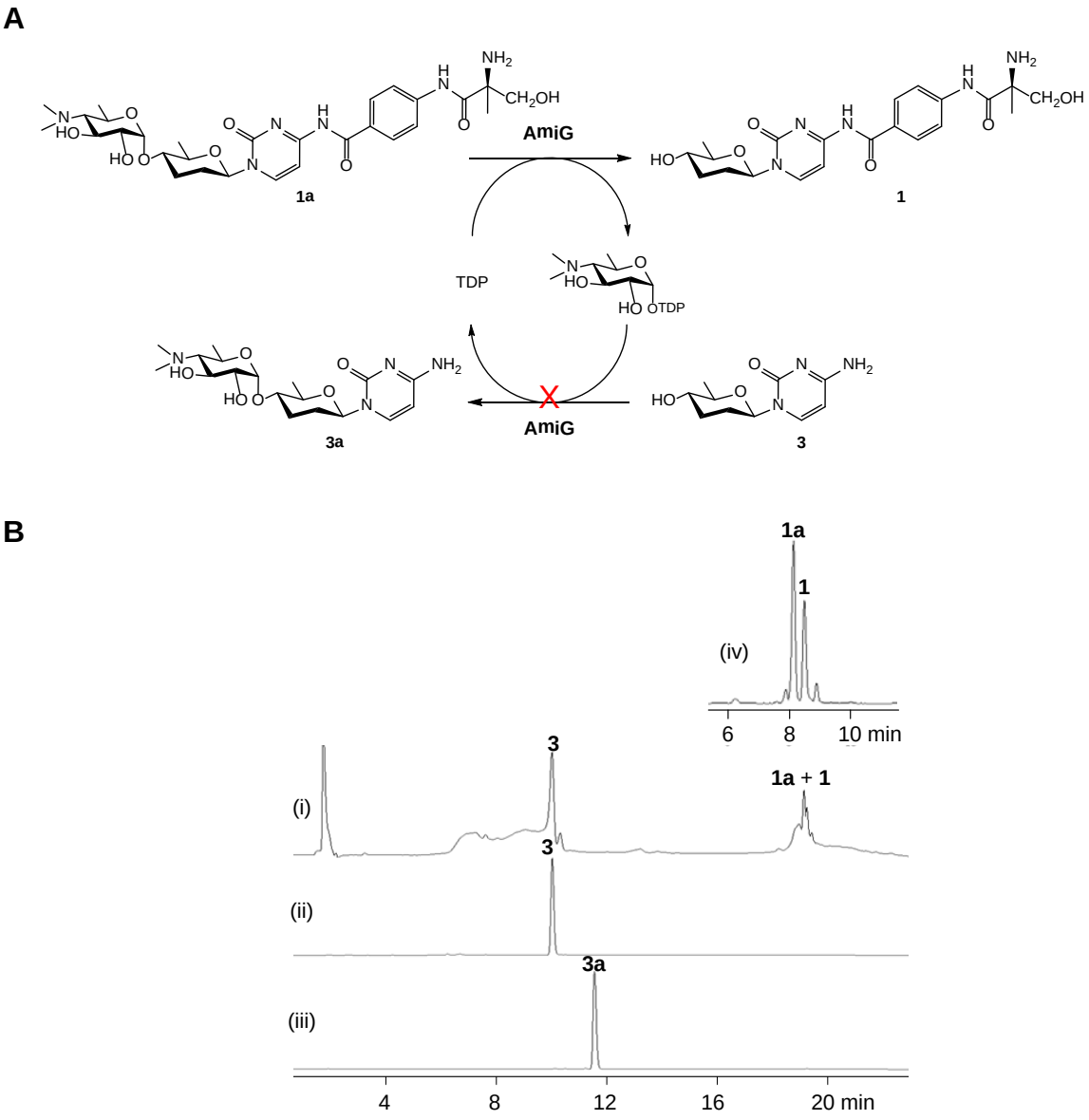


Figure S20. AmiG catalyzed an aglycon exchange reaction using **1a** and **3**. (**A**) A Scheme for a designed aglycon exchange reaction using **3**, **1a**, and TDP as co-substrates. The red cross denotes that the *in situ* generated TDP-D-amosamine from AmiG reverse catalysis on **1a** could not be transferred to an alternative aglycon **3**. (**B**) HPLC analyses of AmiG-catalyzed reactions. (i) and (iv) **1a** + **3** + TDP + AmiG, (ii) **3** standard; (iii) **3a** standard. AmiG assays were performed at 35°C for 6 h in a total of 50 μL reaction mixture containing 100 μM **3**, 100 μM **1a**, 2 mM TDP, 3.3 μM AmiG in 50 mM MOPS buffer (pH 6.5) with 10 mM $MgCl_2$. HPLC traces shown in (i), (ii) and (iii) were carried out with on a reversed phase column Luna C18 (Phenomenex, 150 × 4.6 mm, 5 μm) with UV detection at 254 nm under the following program: solvent system (solvent A, 0.15% TFA in water; solvent B, 90% CH_3CN in water); process: 0 -20% B (0 – 15 min); 20 - 80% B (15 -18 min); 80 - 0% B (18-19 min); 0% B (19 – 25 min), flow rate at 1 mL min^{-1}. HPLC trace shown in (iv) were carried out with a general analysis program: a reversed phase column Luna C18 (Phenomenex, 150 × 4.6 mm, 5 μm) with UV detection at 254 nm under the following program: solvent system (solvent A, 10% CH_3CN in water supplementing with 0.8% trifluoroacetic acid (TFA); solvent B, 90% CH_3CN in water); process: 0 -40% B (0 – 15 min); 40 - 80% B (15 -18 min); 80 - 0% B (18-19 min); 0% B (19 – 25 min), flow rate at 1 mL min^{-1}.

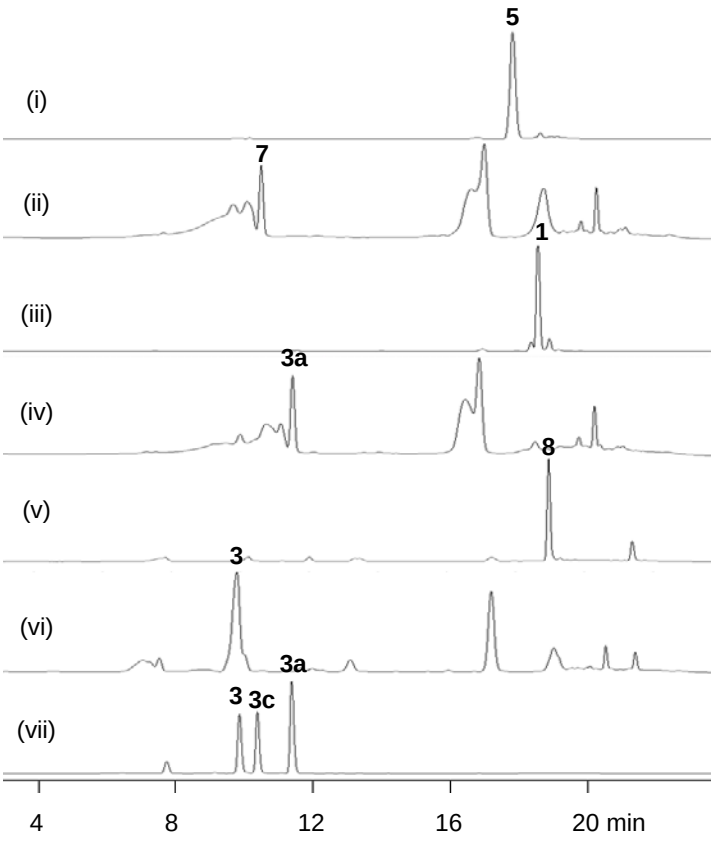


Figure S21. Labile degradation of amicetin analogues to afford **3c**, **3**, and **3a**.* (**i**) **5** standard; (**ii**) Conversion of **1c** to **3c** under heating; (**iii**) **1a** standard; (**iv**) Conversion of **1a** to **3a** under heating; (**v**) **1** standard; (**vi**) Conversion of **1** to **3** under heating; (**vii**) **3**, **3c**; **3a** standards. HPLC were carried out with on a reversed phase column Luna C18 (Phenomenex, 150 × 4.6 mm, 5 μm) with UV detection at 254 nm under the following program: solvent system (solvent A, 0.15% TFA in water; solvent B, 90% CH_3CN in water); process: 0 -20% B (0 – 15 min); 20 - 80% B (15 -18 min); 80 - 0% B (18-19 min); 0% B (19 – 25 min), flow rate at 1 mL min^{-1}.

* By mimicking the isolation conditions, 0.2 mg amicetin (**1a**) and its analogues (**1c** and **1**) were dissolved in 50μL methanol and subsequently put into 2.5 ml mixture solvent of n-butanol:water:methanol (1:1:0.5) with 2.5 μL anhydrous ethylene diamine. The mixtures were heated at 55°C for 6 h. and then the solvent was removed in vacuum. The residues were dissolved in methanol and subjected to HPLC-MS analysis.

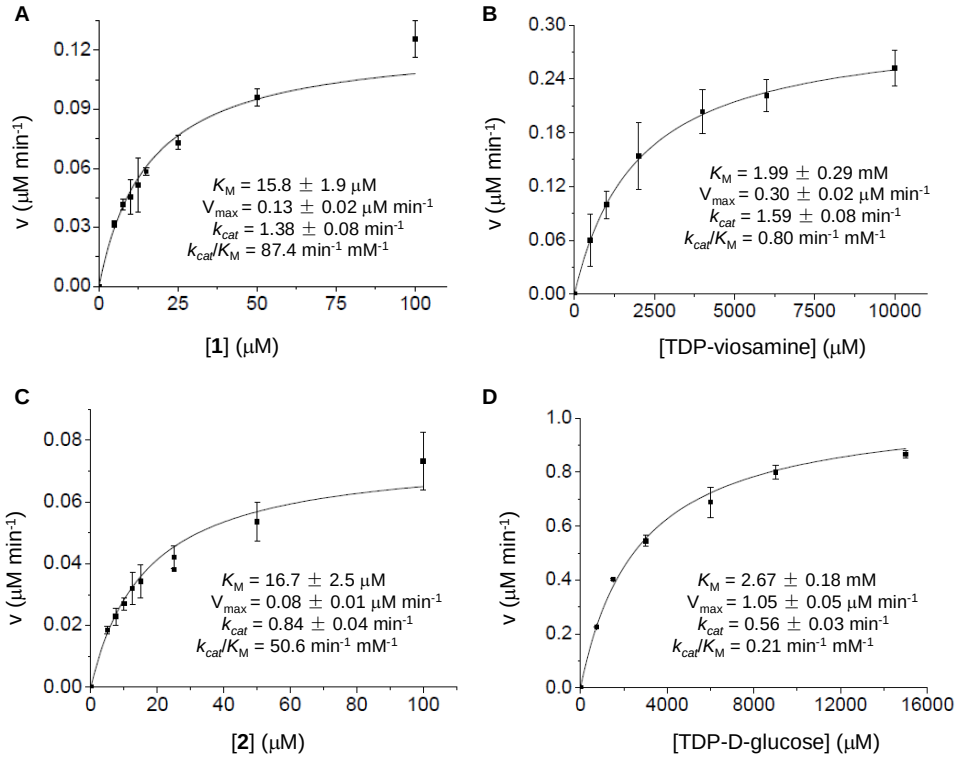


Figure S22. Steady state kinetic analysis of AmiG catalyzed reactions. (**A**) Determination of kinetic parameters for **1** with saturating TDP-D-viosamine (10 mM): **1** was set as a variable substrate in concentrations of 5, 7.5, 10, 12.5, 15, 25, 50, and 100 μM. Enzyme assay was performed in MOPS buffer (50 mM, pH 6.5) containing 0.095 μM AmiG and 10 mM $MgCl_2$ at 35°C for 20 min in triplicate. (B) Determination of kinetic parameters for TDP-D-viosamine with saturating **1** (100 μM): TDP-D-viosamine was set as a variable substrate in concentrations of 0.500, 1, 2, 4, 6, and 10 mM. Enzyme assay was performed in MOPS buffer (50 mM, pH 6.5) containing 0.19 μM AmiG and 10 mM $MgCl_2$ at 35°C for 10 min in triplicate. (C) Determination of kinetic parameters for **2** with saturating TDP-D-viosamine (10 mM): **2** was set as a variable substrate in concentrations of 5, 7.5, 10, 12.5, 15, 25, 50, and 100 μM. Enzyme assay was performed in MOPS buffer (50 mM, pH 6.5) containing 0.095 μM AmiG and 10 mM $MgCl_2$ at 35°C for 20 min in triplicate. (B) Determination of kinetic parameters for TDP-D-glucose with saturating **1** (100 μM): TDP-D-glucose was set as a variable substrate in concentrations of 0.75, 1.5, 3, 6, 9 and 15 mM. Enzyme assay was performed in MOPS buffer (50 mM, pH 6.5) containing 1.9 μM AmiG and 10 mM $MgCl_2$ at 35°C for 10 min in duplicate. K_M and V_{max} were calculated by nonlinear regression analysis using Origin 8 software.

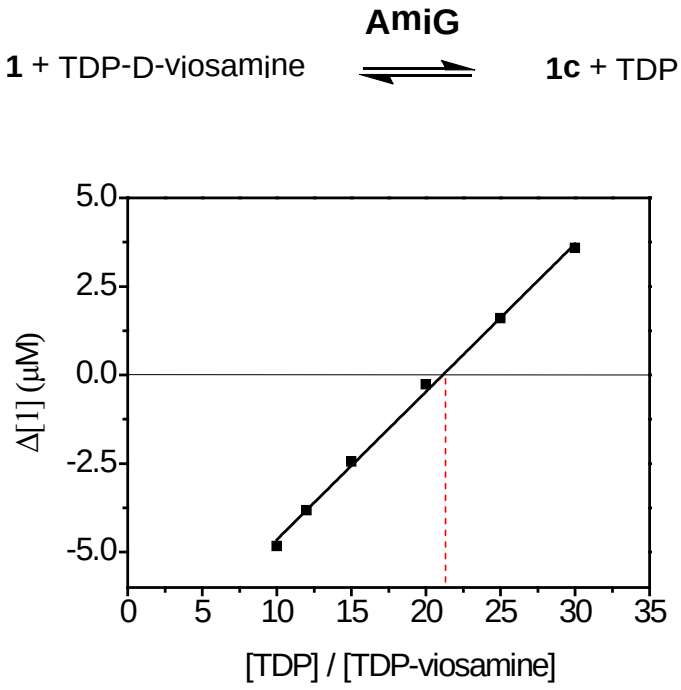


Figure S23. Determination of the equilibrium constant (K_{eq}) of AmiG reaction with **1** and TDP-D-viosamine. The equilibrium constant (K_{eq}) was measured by using **1** as glycosyl acceptor and **1c** as glycosyl donor. The K_{eq} was determinated via a series of reactions under saturation conditions in which the ratio of [TDP]/[TDP-viosamine] varied from 10 to 30 while the ratio of [**1**]/[**1c**] was fixed at 15/85. The change in [**1**] was measured by HPLC after incubatation in 50mM MOPS buffer (pH 6.5) at 35°C for 18 h. The value of the abscissa axis that corresponds to the 0-value intercept of the ordinate axis is the uncorrected K_{eq} (in this case is 21), which was corrected by multiplying by the initial [**1**]/[**1c**], as this ratio was not exactly 1. The K_{eq} was calculated using the equation K_{eq} = ([TDP]/[TDP-viosamine])×([**1**]/[**1c**]). Thus, K_{eq} = 21 × (85/15) = 120.

Supplemental references

(1) Gust, B.; Challis, G. L.; Fowler, K.; Kieser, T.; Chater, K. F. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1541.

(2) Bierman, M.; Logan, R.; O'Brien, K.; Seno, E. T.; Rao, R. N.; Schoner, B. E. *Gene* **1992**, *116*, 43.

(3) Zhang, G.; Zhang, H.; Li, S.; Xiao, J.; Zhu, Y.; Niu, S.; Ju, J.; Zhang, C. *Appl. Environ. Microbiol.* **2012**, *78*, 2393.

(4) Bradford, M. M. *Anal. Biochem.* **1976**, *72*, 248.