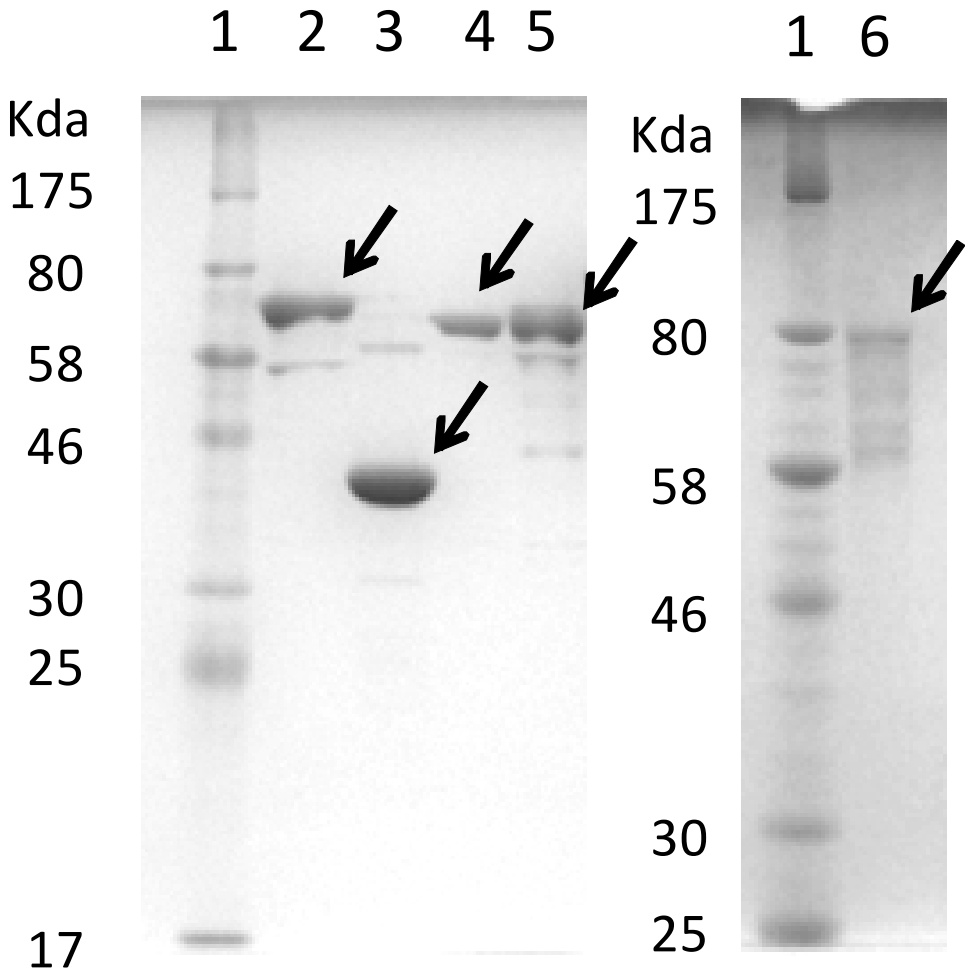
Supporting information for “Diversity of the early step of the futalosine pathway”

Chisato Arakawa[1], Masahiro Kuratsu[2], Kazuo Furihata[3], Tomoshige Hiratsuka[1], Nobuya Itoh[1], Haruo Seto[4], and Tohru Dairi[1,5]

Biotechnology Research Center, Toyama Prefectural University, Toyama 939-0398, Japan,[1] Kyowa Hakko Bio Co. Ltd., Chiyoda-ku, Tokyo 100-8185, Japan [2] Division of Agriculture and Agricultural Life Science, The University of Tokyo, Bunkyo-ku, 113-8657, Japan,[3] Faculty of Applied Bio-science, Tokyo University of Agriculture, Setagaya-ku, 156-8502, Japan,[4] Graduate School of Engineering, Hokkaido University, Hokkaido 060-8628, Japan. [5]

Contents:

Supporting Fig. 1.



Electrophoresis of the overproduced and purified recombinant enzymes. Molecular mass marker (lane 1), Acel_0106 (lane 2), Acel_0264 (lane 3), HP0089 (lane 4), SCO4327 (lane 5), and SCO5662 (lane 6) were analyzed on an SDS-PAGE (12.5%). Proteins were stained with Coomassie brilliant blue R-250.

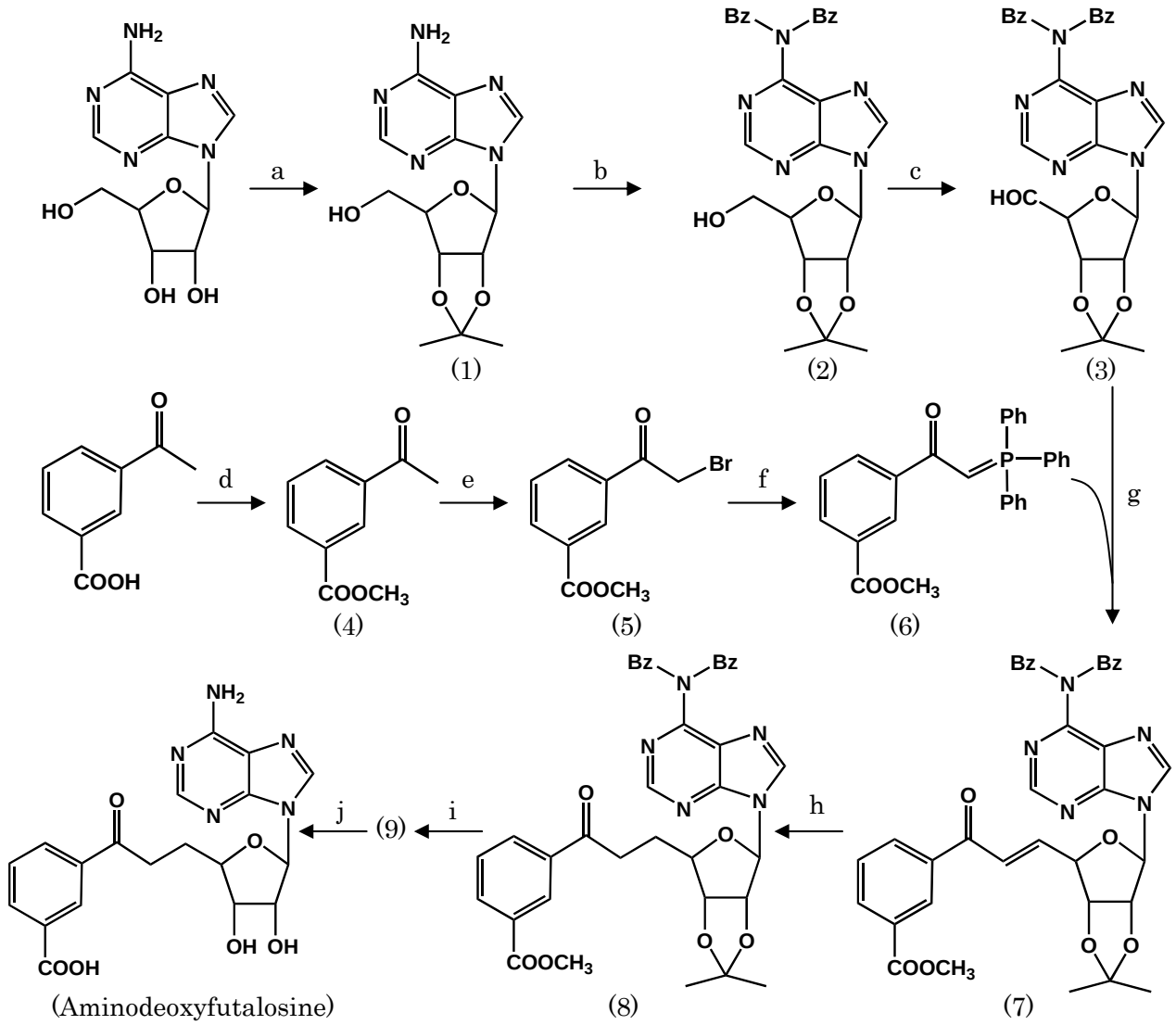
To overproduce Acel_0106, Acel_0264, HP0089, SCO4327 and SCO5662 as *N*-terminal maltose binding protein (MBP)-fused proteins or *N*-terminal His-tagged proteins, the corresponding DNAs were amplified by PCR with the following 5' and 3' primers: for Acel_0106, 5'-TTTGAATTCAGCGTCAAGCGGCTCATCATCACTGC-3' and 5'-TTTAAGCTTTCATGGGCACCCTCCGTCGATCTC-3'; for Acel_0264,

5'-CCCTGGCCATGGACACCCCACGATCCCGTCTCGGTTGAGG-3' and 5'-CCCCCAAGCTTTCACAGCGACTCTCCCGCCGCTGTGGTCG-3'; for HP0089, 5'-GGGGGATCCGTGCAAAAAATTGGCATTTTAGGGGC-3' and 5'-AAACTGCAGCTAAAGCTCATCCACCATGCTTTTTA-3'; for SCO4327, 5'-TTTGAATTCCACCTCCTCGTGGCCACCGCGGTCT-3' and 5'-TTTCTGCAGTCAGCGCTCATGCGGTTTCCAACTCTCAAGGACG-3'; and for SCO5662, 5'-TTTGGATCCACCGAGCACCTCGTCGACCCCGACGT-3' and 5'-AAAAAGCTTTCAGGAGGCGAGCCAGGCGGCGGTGTA-3'. To facilitate their insertion into expression vectors, additional restriction sites (underlined) were incorporated into the primers. After sequence confirmation, the amplified DNA fragments were inserted into pQE30 (Acel_0264) and pMAL-c2X (Acel_0106, HP0089, SCO4327 and SCO56626). The former plasmid and the latter plasmids were introduced into *E. coli* M15/pREP4 and *E. coli* TB1, respectively. Expression and purification conditions for the recombinant enzyme were essentially the same as those described in the manufacturer's protocols. Purity of the recombinant enzymes was analyzed by an SDS-PAGE on 12.5% gels. Protein concentration was determined by a protein-dye standard assay (Bio-Rad) using bovine serum albumin as a standard.

Assay for recombinant enzymes. A reaction mixture (50 μl) containing 50 mM citrate buffer (pH 6.0), 0.4% (w/v) 2-mercaptoethanol, 10 mM futalosine (or aminodeoxyfutalosine) and 5 to 50 μg of the purified recombinant enzyme was incubated at 30 °C for 30 min. The reaction product was subjected to HPLC analysis. The analytical conditions were as follows: Merck Mightisil RP-18GP Aqua column (250 × 4.6 mm); temperature, 30 °C; detection, 230 nm; mobile phase, 20 mM potassium phosphate:acetonitrile = 100:0 at 0 min, and a linear gradient to 0:40 for an additional 40 min; flow rate, 0.8 ml/min.

Supporting Fig. 2.

Synthesis of aminodeoxyfutalosine



(a) 2,2-dimethoxypropane, acetone, *p*-toluenesulfonic acid, rt, O.N.; (b) (1)trimethylsilyl chloride, pyridine, rt, 30min; (2) benzoyl chloride, rt, 3h; (c) N'N'-dicyclohexylcarbodiimide, P_2O_5, DMSO, rt, 24h; (d) MeOH, H_2SO_4, rt, O.N.; (e) $CuBr_2$, EtOAc, reflex, 3h; (f) triphenylphosphine, benzene, rt, O.N. (g) pyridine, rt, 24h (h) H_2/Pd/C, rt, 2 days; (i) MeOH, K_2CO_3, rt, O.N. (j) 90% aq. TFA, rt, 3h

(a) Preparation of 2',3'-*O*-isopropylidene adenosine (1)

To a suspension of adenosine (1 g, 3.7 mmol) and 2**,**2-dimethoxypropane (1.5 g, 14.8 mmol) in acetone (70 mL) was added *p*-toluenesulfonic acid (1.4 g, 7.4 mmol) and stirred overnight at room temperature. After neutralization with saturated $NaHCO_3$, the reaction mixture was evaporated and the residual aqueous layer was extracted with

chloroform. Compound **(1)** was obtained as white needle by crystallization from methanol (yield, 79%). **(1)**: $C_{13}H_{17}N_5O_4$; HR-TOF-MS (*m/z*), calcd. $(M+H)^+$ 308.1353, obsd. $(M+H)^+$ 308.1333; ^{1}H-NMR (400 MHz, $CDCl_3$): δ 8.29 (s, 1H), 7.81 (s, 1H), 6.53 (d, *J* = 11.5 Hz, 1H), 5.93 (br s, 2H), 5.83 (d, *J* = 4.9 Hz, 1H), 5.19 (t, *J* = 5.3 Hz, 1H), 5.09 (dd, *J* = 5.3 Hz, 1.0 Hz, 1H), 4.52 (m, 1H), 3.95 (m, 2H), 3.77(m, 1H), 1.62 (s, 3H) , 1.35 (s, 3H).

(b) Preparation of 2',3'-*O*,*O*-isopropylidene-*N*6,*N*6-dibenzoyladenosine (2)

TMS-Cl (5 equiv) was added to a pyridine solution of **(1)** and stirred at room temperature for 30 min, followed by the addition of benzoyl chloride (2.6 equiv) and stirred for additional 3.5 hr. After addition of water (half volume of that of benzoyl chloride), the organic layer was separated and the solvent was removed by evaporation. To the residuess was added ice-cold 2N sulfuric acid and sodium hydrogen carbonate. The product was extracted with chloroform, washed with water and then subjected to silica gel chromatography ($CHCl_3$:MeOH=50:1) to give **(2)** (yield, 75%). **(2)**: $C_{27}H_{25}N_5O_6$; HR-TOF-MS (*m/z*), calcd. $(M+H)^+$ 516.1878, obsd. $(M+H)^+$ 516.1891; ^{1}H-NMR (400 MHz, CD_3Cl): δ 8.50 (s, 1H), 8.06 (s, 1H), 7.79 (d, *J* = 7.2 Hz, 4H), 7.43 (t, *J* = 7.2 Hz, 2H), 7.30 (m, 4H), 5.87 (d, *J* = 4.8 Hz, 1H), 5.19 (m, 1H), 5.04 (t, *J* = 3.0 Hz), 4.48 (m, 1H), 3.80 (m, 2H), 3.73(m, 1H), 1.58 (s, 3H) , 1.32 (s, 3H).

(c) Preparation of (3)

To a solution of 1 g of freeze-dried compound **(2)** in dryed 15 ml of DMSO was added 1 mM of anhydrous phosphate (1.5 mL) and freeze-dried DCC (3.4 g, 16.25 mmol), and stirrred at room temperature for 24 hr. Since the formed aldehyde **(3)** was unstable, the product was, without purification, used for synthesis of **(7)** after confirmation of the formation of **(3)** by TLC.

(d) Preparation of methyl 3-acetylbenzoate (4)

To a solution of 3-acetylbenzoic acid (5 g, 30.5 mmol) in 70 mL of methanol was added 0.3 mL of conc. sulfuric acid, and refluxed overnight. The reaction mixture was neutralized with saturated $NaHCO_3$ and concentrated *in vacuo*. The residual aqueous layer was diluted with water and extracted with ethyl acetate three times. The combined organic layer was concentrated *in vacuo* to afford white needles (yield, 93.2%). **(4)**: $C_{10}H_{10}O_3$; HR-TOF-MS (*m/z*), calcd. $(M+H)^{++}$ 179.0703, observed $(M+H)^+$ 179.0725; ^{1}H-NMR (400 MHz, $CDCl_3$): δ 8.56 (m, 1H, Ar2H), 8.20 and 8.13 (ddd, 2H, Ar4 and Ar6H), 7.53 (ddd, 1H, Ar5H), 3.93 (s, 3H, COOMe), 2.63 (s, 3H, COMe). ^{13}C-NMR (100 MHz, $CDCl_3$): δ 197.4(s), 167.5(s),166.5(s), 137.5(s), 134.1(d), 132.5(d), 130.9(s),

129.8(d), 129.1(d), 52.6(q), 27.0(q); IR (KBr) = 1700 cm^{-1}.

(e) Preparation of methyl 3-bromoacetylbenzoate (5)

An ethyl acetate solution of **(4)** was refluxed in the presence of copper (II) bromide for 3 hr. After the reaction, precipitates were removed by filtration. The organic layer was washed with water and concentrated *in vacuo* to afford white solid of **(5)** (yield, 90.4%). **(5)**: $C_{10}H_9B_rO_3$; HR-TOF-MS (*m/z*), calcd. $(M+H)^+$ 256.9808, obsd. $(M+H)^+$ 256.9799; ^{1}H-NMR (400 MHz, $CDCl_3$): δ 8.61 (m, 1H, Ar2H), 8.28 and 8.19 (2ddd, 2H, Ar4 and Ar6H), 7.60 (ddd, 1H, Ar5H), 4.49 (s, 2H, CH_2Br), 3.97 (s, 3H, COOMe). ^{13}C-NMR (100 MHz, $CDCl_3$): δ 190.5(s), 165.9(s), 135.0(s), 134.6(d), 132.9(d), 130.8(s), 129.9(d), 129.1(d), 52.5(q), 30.7(q); IR (KBr) = 1700 cm^{-1}.

(f) Preparation of methyl 3-(triphenylphosphoranylidene)acetyl benzoate (6)

Triphenylphosphine (1 g, 3.82 mmol; ca. 1 eq) and **(5)** (1 g, 3.89 mmol) were dissolved in 5 mL of dry benzene and the solution was stirred overnight at room temperature under argon atmosphere. The reaction mixture was filtrated and the white-purple precipitates were dissolved in methanol/water (1:1) and neutralized with 0.5 N NaOH. After stirring at room temperature for 10 min, the mixture was extracted with ethyl acetate. After evaporation, **(6)** was obtained as a yellow powder (yield, 89.5%). **(6)**: $C_{28}H_{23}O_3P$; HR-TOF-MS (*m/z*), calcd. $(M+H)^+$ 439.1458, obsd. $(M+H)^+$ 439.1469; ^{1}H-NMR (400 MHz, $CDCl_3$): δ 9.01-7.50 (m, 19H, 4Ar), 6.53 (d, *J* = 12 Hz, 1H, CH), 3.90 (s, 3H, COOMe). $C_{28}H_{24}O_3P$ by TOF-MS 439.1462 $[M+H]^+$ (calcd. 439.1463).

(g) Preparation of (7)

Pyridine (1 mL) and **(6)** (850 mg, 1.63 mmol; 0.5 eq.) were added into the flask containing **(3)** and stirred overnight. The reaction mixture was filtered with celite and the filtrate was washed successively with ice-cold water, 1N HCl and saturated brine. The organic layer was dried over anhydrous Na_2SO_4, concentrated *in vacuo*, and subjected to silica gel chromatography (hexane:ethyl acetate=1:2) to give **(7)** (yield, 24.0%). **(7)**: $C_{37}H_{31}N_5O_8$; HR-TOF-MS (*m/z*), calcd. $(M+H)^+$ 674.2245, obsd. $(M+H)^+$ 674.2261; ^{1}H-NMR (400 MHz, $CDCl_3$) δ: 8.60 (1.0H, s), 8.47 (1.0H, t, J = 1.69 Hz), 8.20 (1.0H, dt, J = 7.77, 1.39 Hz), 8.15 (1.0H, s), 7.95 (1.1H, dt, J = 7.83, 1.46 Hz), 7.81 (5.0H, dd, J = 8.33, 1.19 Hz), 7.52-7.47 (1.2H, m), 7.45 (2.3H, ddd, J = 9.91, 5.06, 2.18 Hz), 7.32 (5.0H, t, J = 7.73 Hz), 7.10 (1.0H, dd, J = 15.42, 5.01 Hz), 7.01 (1.0H, dd, J = 15.46, 1.29 Hz), 6.21 (1.0H, d, J = 2.28 Hz), 5.50 (1.0H, dd, J = 6.34, 2.28 Hz), 5.15 (1.0H, dd, J = 6.34, 3.97 Hz), 4.94 (1.0H, t, J = 3.97 Hz), 3.91 (3.0H, s), 1.65 (3.2H, s),

1.40 (3.2H, s).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 188.7, 172.1, 166.1, 152.4, 152.2, 144.0, 143.5, 137.4, 133.9, 133.1, 132.7, 130.8, 130.0, 129.6, 129.4, 129.0, 128.8, 128.4, 127.9, 127.7, 125.8, 115.3, 90.6, 86.3, 84.0, 83.8, 52.4, 29.7, 27.2, 25.4.

(h) Preparation of (8)

To a solution of 1 g of **(7)** in 100 mL of ethanol was added PtO_2 (25 mg) and hydrogenated with a Parr hydrogenation apparatus under 4 atm. for 2 days. After filtration with celite, the solvent was evaporated to afford **(8)** (yield, 65%). **(8)**: $C_{37}H_{33}N_5O_8$; HR-TOF-MS (*m/z*), calcd. $(M+H)^+$ 676.2402, obsd. $(M+H)^+$ 676.24111; ^{1}H-NMR ($CDCl_3$) δ: 8.64 (0.9H, s), 8.54 (1.0H, t, J = 1.54 Hz), 8.20 (1.0H, dt, J = 7.73, 1.44 Hz), 8.12 (1.0H, s), 8.07 (0.9H, dt, J = 7.83, 1.51 Hz), 7.83 (3.9H, dd, J = 8.33, 1.29 Hz), 7.50 (1.0H, t, J = 7.78 Hz), 7.45 (2.2H, tt, J = 7.44, 1.42 Hz), 7.33 (4.2H, t, J = 7.73 Hz), 6.06 (1.0H, d, J = 2.68 Hz), 5.45 (1.0H, dd, J = 6.54, 2.68 Hz), 4.89 (1.0H, dd, J = 6.54, 4.06 Hz), 4.32-4.28 (1.0H, m), 3.92 (3.1H, s), 3.19-3.04 (2.0H, m), 2.28-2.10 (2.0H, m), 1.59 (3.1H, s), 1.37 (3.1H, s).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 197.9, 172.2, 166.2, 152.4, 152.1, 144.0, 136.9, 134.0, 134.0, 133.0, 132.1, 130.7, 130.1, 129.5, 129.1, 128.9, 128.7, 128.4, 128.0, 115.1, 90.4, 85.9, 83.9, 52.4, 34.4, 27.3, 27.2, 25.4.

(i) Preparation of (9), a debenzoyated (8)

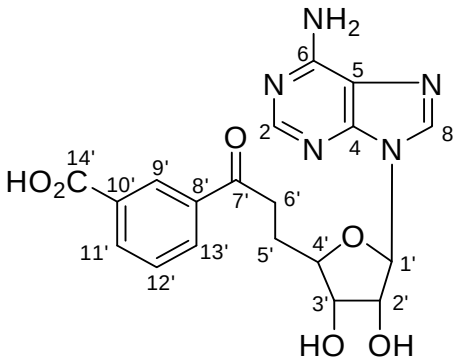
To a solution of 1 g of **(8)** in 90% methanol was added K_2CO_3 (815mg) and stirred overnight at room temperature. The reaction mixture was evaporated to dryness and subjected to Sephadex LH-20 column chromatography to afford **(9)** (yield, 38.0%). **(9)**: $C_{22}H_{23}N_5O_6$; HR-TOF-MS (*m/z*), calcd. $(M+H)^+$ 454.1721, obsd. $(M+H)^+$ 454.1702; ^{1}H-NMR (400MHz, CD_3OD) δ: 8.42 (1.0H, d, J = 1.19 Hz), 8.16 (1.0H, s), 8.09 (1.1H, s), 8.04 (1.0H, dd, J = 7.53, 1.19 Hz), 7.82 (1.0H, dd, J = 7.93, 1.19 Hz), 7.34 (1.0H, t, J = 7.73 Hz), 6.03 (1.0H, t, J = 1.19 Hz), 5.41-5.39 (1.0H, m), 4.87 (0.9H, dd, J = 6.34, 3.57 Hz), 4.18 (1.0H, dd, J = 10.31, 6.74 Hz), 3.04 (2.0H, td, J = 6.84, 2.64 Hz), 2.04-1.98 (2.0H, m), 1.47 (3.0H, s), 1.27 (3.1H, s).

^{13}C-NMR (100 MHz, $CDCl_3$) δ: 201.2, 173.3, 157.3, 154.0, 150.3, 141.8, 138.6, 137.8, 134.9, 130.9, 130.1, 129.3, 120.5, 115.8, 91.0, 87.2, 85.4, 85.2, 35.5, 28.8, 27.5, 25.6.

(j) Preparation of aminodeoxyfutalosine

(9) was dissolved in 10% TFA and stirred at room temperature for 4 hr. After concentration *in vacuo*, crystalline aminodeoxyfutalosine was collected by filtration

(yield, 43%). Data for aminodeoxyfutalosine is as following:



Structure of aminodeoxyfutalosine

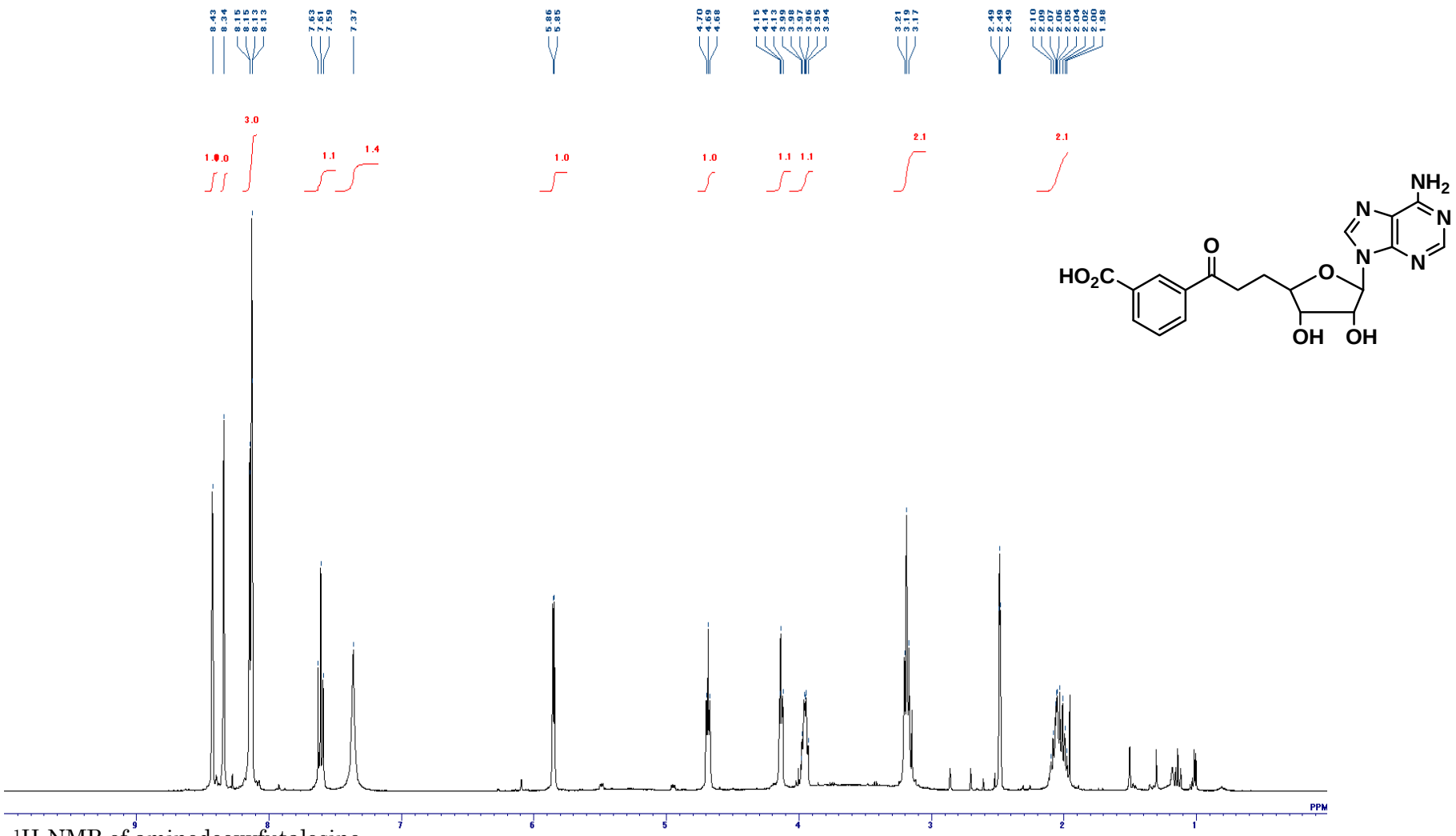
NMR data of aminodeoxyfutalosine

	Aminodeoxyfutalosine (DMSO-D_6)			
C		δ (^{13}C)	δ (^{1}H)	Multiplicity
2	CH	152.4	8.43	s
4	C	149.4		
5	C	119.3		
6	C	155.9		
8	CH	140.3	8.13	s
1'	CH	87.8	5.85	d (5.1)
2'	CH	83.1	4.69	t (5.1)
3'	CH	73.2	4.14	t (5.1)
4'	CH	73.2	3.96	td (7.3, 5.1)
5'	CH_2	27.3	2.08 2.01	ddt (13.4, 7.3, 7.1) ddt (13.4, 7.3, 7.1)
6'	CH_2	34.4	3.19	t (7.1)
7'	C=O	199.0		
8'	C	136.9*		
9'	CH	129.3	8.43	br t (1.6)
10'	C	131.3*		
11'	CH	133.6	8.14	d (7.7)
12'	CH	128.4	7.61	t (7.7)
13'	CH	132.1	8.14	d (7.7)
14'	COOH	166.8		
	NH_2		7.37	br s

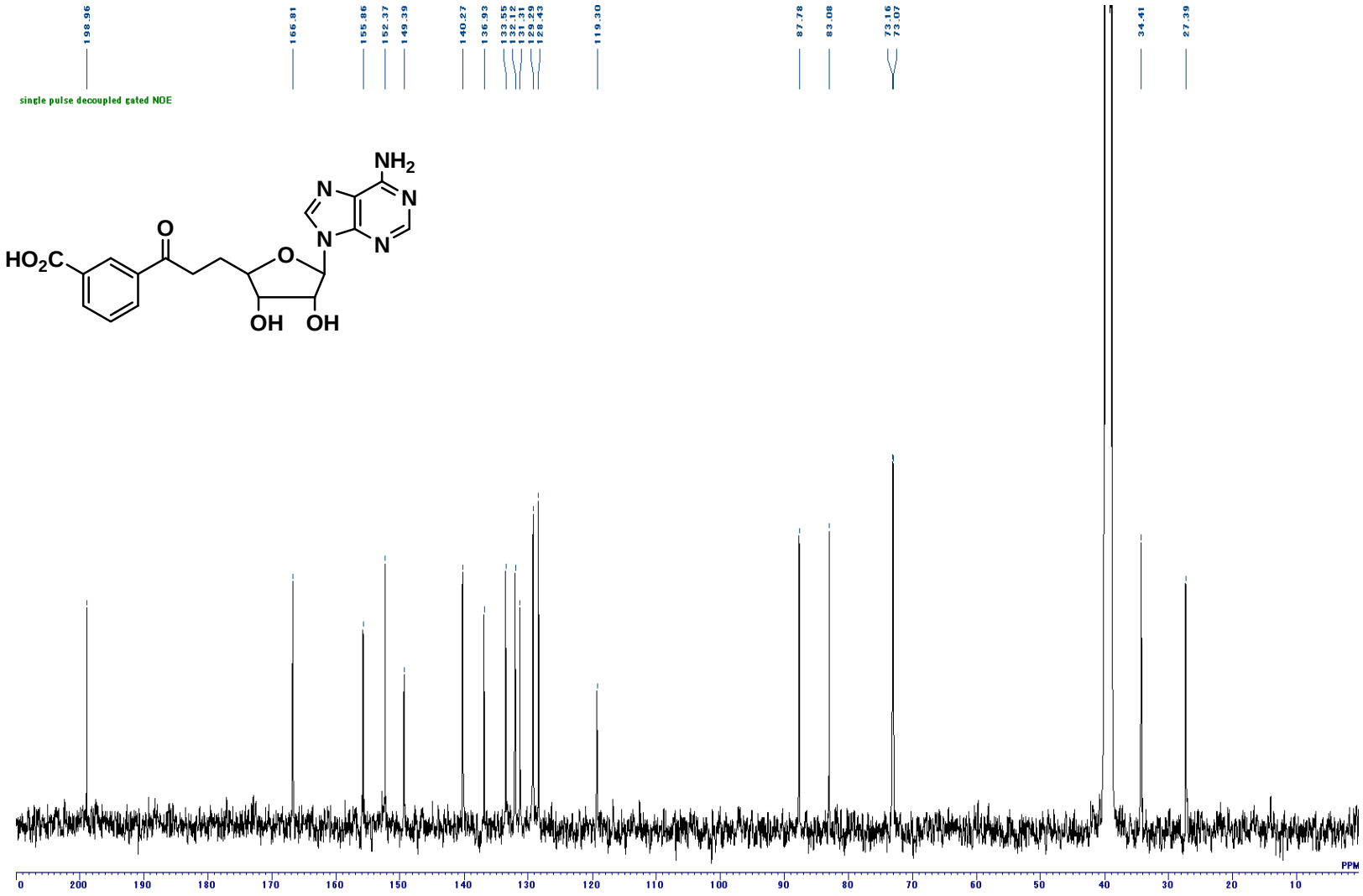
*Exchangeable

Aminodeoxyfutalosine

$C_{19}H_{19}N_5O_6$; HR-TOF-MS (*m/z*), calcd. $(M+H)^+$ 414.1408, obsd. $(M+H)^+$ 414.1397; IR ν KBr max: 3382, 1681, 1203 cm^{-1}; UV (MeOH) ν_{max} nm (log ε), 258 (3.9); mp, >200 °C (dec).



^{1}H-NMR of aminodeoxyfutalosine



^{13}C-NMR of aminodeoxyfutalosine

Supporting Fig. 3.

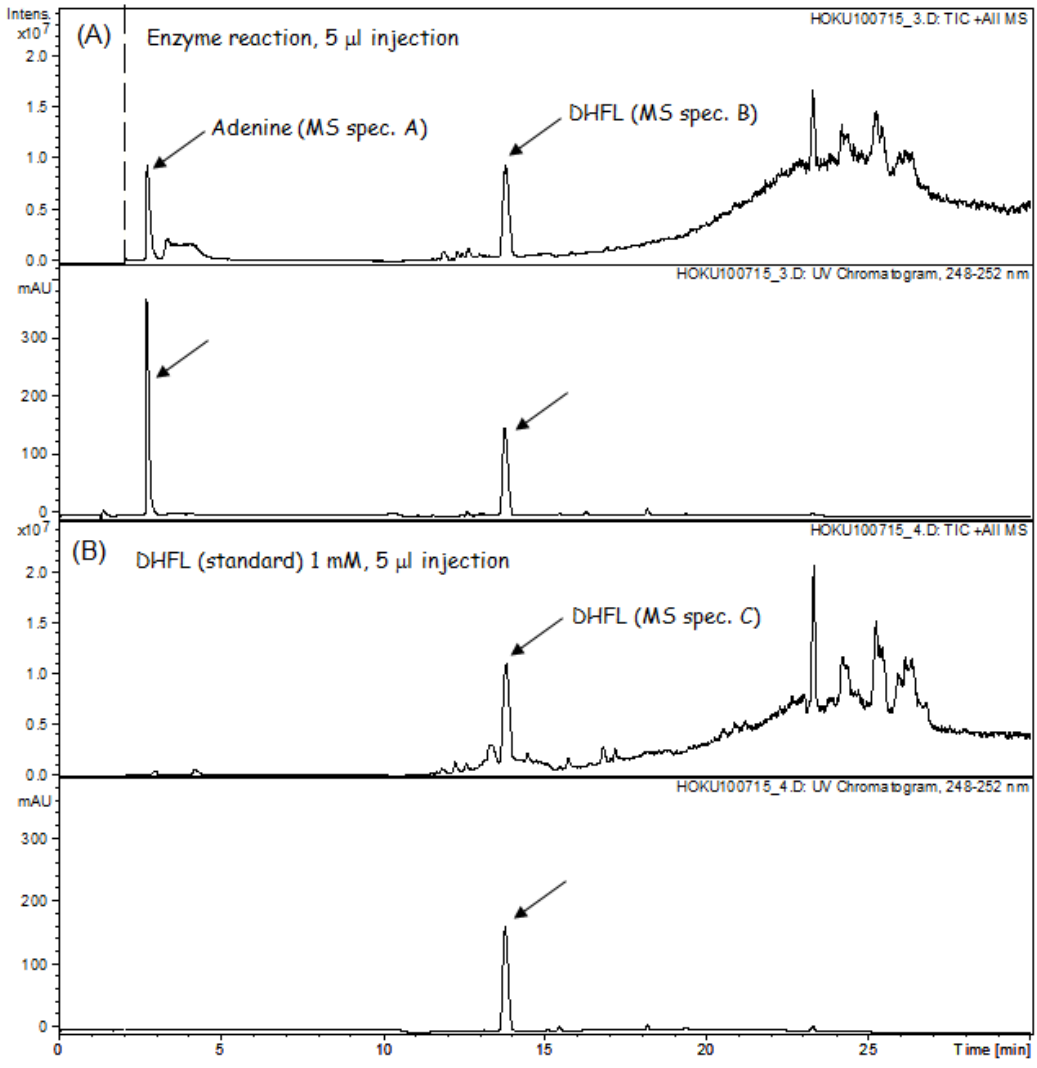
LC-MS analysis of the reaction products formed from AFL by recombinant HP0089. The reaction products (A) and the standard DHFL (B) were analysed.

Date; 10/7/15
Column; Sunniest, RP-Aqua (5μm, 150 X 2.0 mm), 30° C
Flow; 0.3 ml/min
Solvent; A, 0.1% HCOOH in H_2O
B, CH_3CN

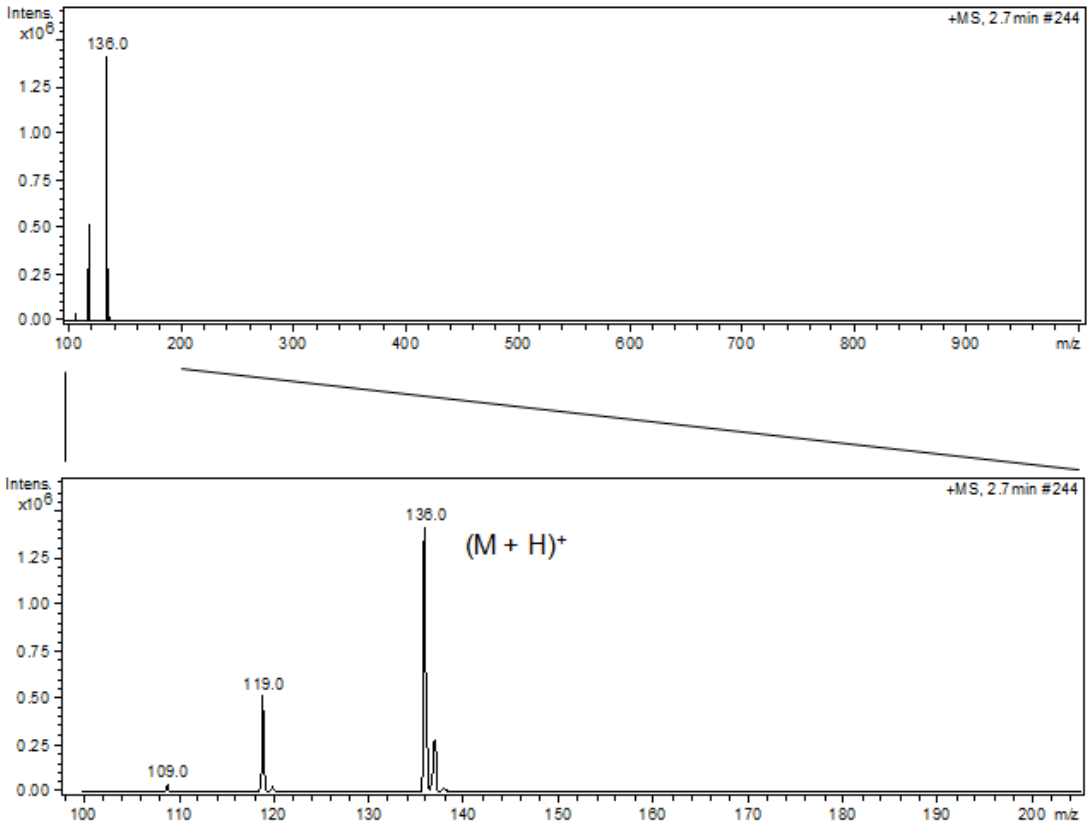
Mode; positive
Injection; 5 μl

Gradient; Method

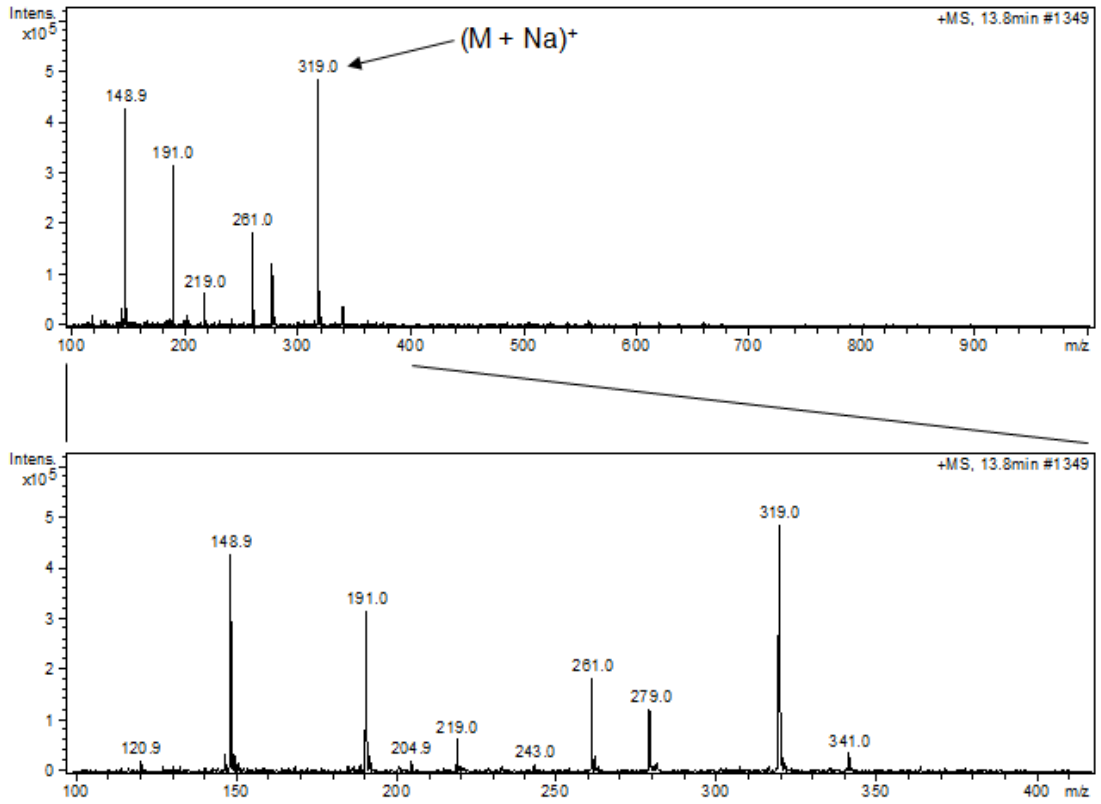
Time (min)	Sol. B (%)	Sol. A (%)
0	0	100
5	0	100
20	90	10
30	90	10



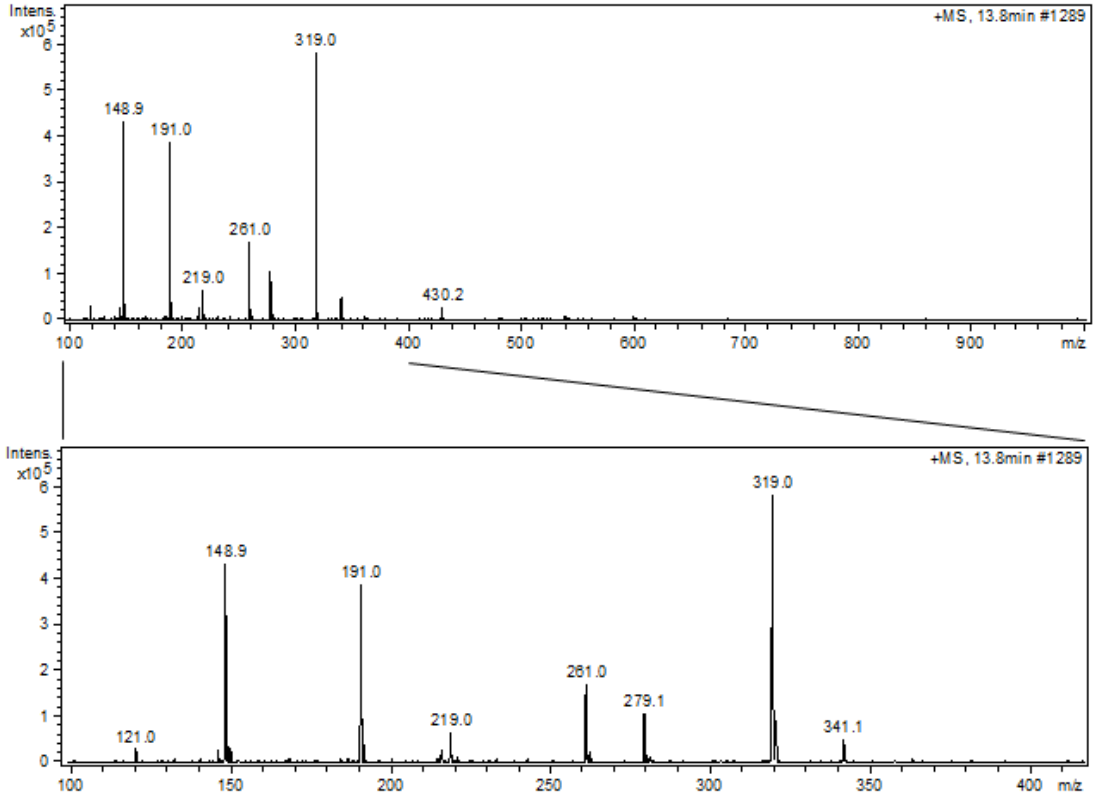
Adenine (MS spec. A)



DHFL (*MS spec. B*)



DHFL (*MS spec. C*)



Supporting Fig. 4.

(1)

```
Paenibacillus sp. JDR-2: Pjdr2_0345  ------------------------------MSENGKILVMTAVAVERD
Desulfotomaculum reducens: Dred_3235 --------------MELKLVEKVATPHGLMPGRAEMRVLVMTAVSAERD
Brevibacillus brevis: BBR47_09530    ----------------------------MILSQQYRSILVVTSVDAERD
Geobacillus sp. Y412MC10: GYMC10_1380------------MQEHTQSNSVDTIESYSSAPHSSKRVLIVTAVDAEKD
Bacillus halodurans: BH2690          MSGGKGLSLEKNMDFRYTRKGTTYIRRSWHISMFNQQKVLIATSVTAEQK
Bacillus pseudofirmus: BpOF4_09540   ------------------------------MSDEGRILIVVSVDAEKE
Streptomyces coelicolor: SCO4327     ---------------------------------------------MSRPG
Streptomyces avermitilis: SAV_3905   ----------------------------------MARAFTPAPHEVPLPG
Streptomyces scabiei: SCAB_50841     ----------------------------------MARALPGPATEVRLPA
Streptomyces griseus: SGR_3174       ------------------MRVLVVTAVPVERDAVTRAFGGAPEAVALPG
```

(2)

```
Paenibacillus sp. JDR-2: Pjdr2_0345  ------------------------------MSENGKILVMTAVAVERD
Desulfotomaculum reducens: Dred_3235 --------------MELKLVEKVATPHGLMPGRAEMRVLVMTAVSAERD
Brevibacillus brevis: BBR47_09530    ----------------------------MILSQQYRSILVVTSVDAERD
Geobacillus sp. Y412MC10: GYMC10_1380------------MQEHTQSNSVDTIESYSSAPHSSKRVLIVTAVDAEKD
Bacillus halodurans: BH2690          MSGGKGLSLEKNMDFRYTRKGTTYIRRSWHISMFNQQKVLIATSVTAEQK
Bacillus pseudofirmus: BpOF4_09540   ------------------------------MSDEGRILIVVSVDAEKE
Streptomyces coelicolor: SCO4327     ------------------MHLLVATAVSVERDAVARAFPAPGTEMSRPG
Streptomyces avermitilis: SAV_3905   ----------------------------------MARAFTPAPHEVPLPG
Streptomyces scabiei: SCAB_50841     ----------------------------------MARALPGPATEVRLPA
Streptomyces griseus: SGR_3174       ------------------MRVLVVTAVPVERDAVTRAFGGAPEAVALPG
```

(3)

```
GTGCACCTCCTCGTGGCCACCGCGGTCTCCGTCGAACGGGACGCGGTGGCTCGGGCGTTCCCCGCGCCGGGCACGGAG
GTGTCCCGCCCCGGGATCACCCTCCACCGGCTGCCGGACGGCTGGGACCTGCTGGCCGCCGGGGTGGGCCCGGCCCGC
GCCGCCGCCTCCACGGCCGCCGCCCTGACCGCGGCCGCCCTCGACGGCAGGCCCTACGACCTGGTCGTCTCGGCCGGT
ATCGGCGGCGGCTTCGCGCCCGAGGCGCCCGTCGGCTCCCTGGTCGTCGCCGACGCGATCACCGCCGCCGACCTGGGC
GCCGAGACCGCCGACGGCTTCCTGCCGGTCACCGACCTCGGCTTCGGCACCGTCACCCACCTGCCGCCGGCCCCTCTG
GTCCGCGCCGCCGCCGAGGCGACCGGGGCCCGGCCCGGCACCGTGCTGACCGGGTCCACCGTCACCGGCACCGCGGCC
CGCGCCGCGCTGCTGCGCGAGCGCCACCCGGGCGCCCTGGCCGAGGCGATGGAGGGCTTCGGCGTCGCCGAGGCCGCC
GCCGCGCACGGCGTTCCGGTGCTGGAACTGCGCGCGGTCTCCAATCCGGTCGGACCGCGCGACCGCGCCGCCTGGCGC
ATCGGCGAGGCGCTCGCGGCCCTGACGGACGCCGTCGGGAAGCTCGCACCCGTCCTTGAGAGTTGGAAACCGCATGAG
CGCTGA
```

Alignments of the original SCO4327 ORF (1) and the estimated SCO4327 ORF (2) with MqnB orthologs. Extended amino acid (2) and nucleotide (3) sequences were shown in red letters (3).