Synthetic-biology-based discovery of fungal macrolide from *Macrophomina phaseolina*

Yohei Morishita,[a] Terutaka Sonohara, [a] Tohru Taniguchi,[b] Kiyohiro Adachi,[c,d] Makoto Fujita,[d,e] Teigo Asai[a,f*]

a. Department of Life Sciences, Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1 komaba, meguro-ku, Tokyo 153-8902, Japan.

b. Faculty of Advanced Life Science, Frontier Research Center for Post-Genome Science and Technology, Hokkaido University, Kita 21 Nishi 11, Sapporo 001-0021, Japan

c. Instrument Center, Institute for Molecular Science, National Institutes of Natural Sciences, , 5-1 Higashiyama, Myodaiji, Okazaki, 444-8787 , Japa

d. Department of Applied Chemistry, Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan

e. Division of Advanced Molecular Science, Institute for Molecular Science, National Institutes of Natural Sciences, 5-1 Higashiyama, Myodaiji, Okazaki, 444-8787 , Japan

f. Graduate School of Pharmaceutical Sciences, Tohoku University, Aoba-yama, Aoba-ku, Sendai 980-8578, Japan.

Table of contents

Experimental equipment

Analytical TLC were performed on silica gel 60 F254 (Merck). Column chromatography was carried out on silica gel 60 (70–230 and 40-50 mesh). The 500 MHz NMR spectra were recorded on a Bruker AVANCE III 500 spectrometer (^{1}H NMR, 500 MHz; ^{13}C NMR, 125 MHz). Chemical shifts for 1H and ^{13}C NMR are given in parts per million (δ) relative to tetramethylsilane (δ_H 0.00) and residual solvent signals (δ_C 77.0) for $CDCl_3$ as internal standards. Mass spectra were measured on Exactive Orbitrap Mass Spectrometer (Thermo Fischer Scientific). IR was measured on a JASCO FVS-6000 spectrometer and BioTools ChiralIR-2X spectrometer. VCD spectrum measurement was performed on a BioTools ChiralIR-2X spectrometer. UV spectra were recorded on a JASCO-V-730 spectrophotometer. HPLC analysis was performed on a JASCO AS-1555-10 Intelligent Sampler, JASCO PU-4180 RHPLC Pump, JASCO MD-4017 Photo Diode Array Detector (JASCO) and a nano quantity analyte detector (NQADTM, OSAKA SODA), which equipped with COSMOSIL Packed Column 5C18-MS-II (ϕ4.6 mm×150 mm) (Nacalai tesque).

Fungal material

Macrophomina phaseolina NBRC 7317 was obtained from Biological Resource Center, NITE (NBRC). The fungus was cultured in a PDB liquid medium for 3 days and the cultured mycelium was ground to a fine powder in liquid N_2. To the mycelial powder in 1.5 mL tube, 100 µL TE buffer and lysis solution (SDS 1g, 2.5 mL Tris-HCl buffer (1 M), EDTA•2Na•2H_2O 186 mg, NaCl 292 mg, 50 mL mQ) were added, gently inverted and incubated 5 min on ice. After centrifugation at 15,000 rpm, 4ºC, for 20 min, 100 µL of the supernatants were transferred to a new 1.5 mL tube, followed by phenol extraction using Phenol:Chloroform: Isoamyl Alcohol 25:24:1 (pH 5.2, Nacalai tesque) and ethanol precipitation. After centrifugation again, TE buffer were added to the pellet of DNA. The DNA solution was used as a template for cloning of *mpmlA* and *mpmlB*.

Heterologous host strains

Asperillus oryzae NSAR1 (*niaD*$^-$, *sC*$^-$, Δ*argB*, *adeA*$^-$)[1] was used as the host for fungal expression[S1].

Escherichia coli Dh5a (TAKARA) was used for the cloning of *mpmlA* and *mpmlB*[S2].

Culture medium

・Culture medium for *M. phaseolina* NBRC 7317

- PDB agar medium: Potato-Dextrose Broth (DIFCO) 7.2 g and agarose (Nacalai tesque) 4.5 g were suspended in 300 mL diluted water and autoclaved. 20 mL portions were dispensed into sterile petri dishes and cooled at room temperature.

- PDB liquid medium: Potato-Dextrose Broth (DIFCO) 1.44 g were suspended in 60 mL distilled water, autoclaved and cooled at room temperature.

・Culture medium for *Aspergillus oryzae* NSAR1

- PDB agar medium (adenine 0.05%, arginine 0.1%): Potato-Dextrose Broth (DIFCO) 7.2 g and agarose (Nacalai tesque) 4.5 g, adenine (TCI) 0.15 g, L-arginine (Wako) were suspended in 300 mL diluted water and autoclaved. 20 mL portions were dispensed into sterile petri dishes and cooled at room temperature.

・Culture medium for AO-*mpmlAB*

- CD agar medium (NaCl 0.8 M, 0.01% adenine, 0.15% L-methionine): Czapek-Dox Broth (Difco) 10.5 g, NaCl (Nacalai tesque) 14 g, agarose (Nacalai tesque) 4.5 g, adenine 30 mg and L-methionine 450 mg (Nacalai tesque) were suspended in 300 mL diluted water and autoclaved. 20 mL portions were dispensed into sterile petri dishes and cooled at room temperature.

- CPS liquid medium (0.015% adenine): Czapek-Dox Broth (Difco) 1.05 (2.63) g, peptone* 0.3 (0.75) g, Soluble Starch (Nacalai tesque) 0.6 (1.5) g, Maltose Monohydrate (Nacalai tesque) 0.3 (0.75) g and adenine 9 (22.5) mg in 60 (150) mL water in 200 (500 mL) Erlenmeyer flask and autoclaved. *peptone: mixture of 0.2 (0.5) g of soy peptone, casein peptone and meat peptone (Nacalai tesque).

Construction of the *mpmlA* and *mpmlB* co-expression system

The genes were amplified by PrimeSTAR® MAX DNA Polymerase (TAKARA) with primers in Table S1. *E. coli* DH5α were used for cloning, following standard recombinant DNA techniques. Fungal expression plasmid pUARA2 possessing the α-amylase promoter (*amyB*) of *A. oryzae* and auxotrophic marker *argB* of *A. nidulans* respectively was used. The *mpmlA* was divided into two equal-sized fragments; each fragment was amplified with the primers *mpmlA*_IFpUKpnI-FW and *mpmlA*_R1, or *mpmlA*_F1 and *mpmlA*_IFpUKpnI-RV. The *mpmlB* gene was amplified with the primers *mpmlB*_IFpUNotI-FW and *mpmlB*_IFpUNotI-RV. The PCR products were purified. A genomic DNA of *M. phaseolina* NBRC 7317 was used as a template. The resultant fragment *mpmlA* and *mpmlB* were subcloned into pUARA2 which had been digested with Asp718 and NotI to yield pUARA2-*mpmlAB* (Table S2). *A. oryzae* NSAR1 was transformed with pUARA2-*mpmlAB* to construct AO-*mpmlAB*.

Cultivation of the transformants in CPS liquid medium and HPLC analysis

The transformants were cultivated on a selection agar plate at 30°C and its mycelia was inoculated in CPS medium and incubated at 30ºC for 5 days. The 3 mL of culture medium at 3 and 5-7 days were extracted with 2 mL EtOAc in a 5 mL tube and the 800 μL x 2 of EtOAc layer were evaporated in a 1.5 mL tube. The extract was resolved in 100 μL MeOH, centrifuged at 13,500 rpm for 15 min, and 10 μL of the supernatant was injected into HPLC. Flow rate; 1 mL/min, Solvent gradient system: acetonitrile and water with 0.01% TFA (0-2 min: 20:80, 2-12 min: 20:80 to

100: 0, 12-24 min: 100:0). Absorbance was monitored by a nano quantity analyte detector (NQAD^{TM,} OSAKA SODA).

Table S1. Primers used for cloning of *mpmlA* and *mpmlB*.

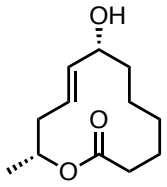
Primer name	DNA sequence 5' to 3'
*mpmlA*_IFpUKpnI-FW	CCGGAATTCGAGCTCGAACATGCTTGTCAGGGAAGTTG
*mpmlA*_IFpUKpnI-RV	ACTACAGATCCCCGGAATCACCCAACACCTCCACG
*mpmlA*_F1	GCTCCTCTCAAGCATACCTC
*mpmlA*_R1	GAGGTATGCTTGAGAGGAGC
*mpmlB*_IFpUNotI-FW	TTTGAGCTAGCGGCCAGCATGGCTGGCAGGCAAGAG
*mpmlB*_IFpUNotI-RV	GTCACTAGTGCGGCCCGTGGAAGCGAAATGGAAGC

Table S2. Expression plasmid vector constructed in this study.

Plasmid name	Original vector	Gene 1 (KpnI site)	Gene 1 (NotI site)
pUARA2_*mpmlAB*	pUARA2	*mpmlA*	*mpmlB*

Isolation of phaseolide A (1)

AO-*mpmlAB* was cultivated in CPS (+ 0.5% maltose) medium (3.6 L; 150 mL x 24) at 30ºC for 6.5 days. The culture broths were extracted with EtOAc twice, and the extracts (514.2 mg) were obtained. A portion of the extract (22.0 mg) was subjected to silica gel column chromatography eluted with n-hexane-EtOAc (2/1-3/2) to give **1** (11.8 mg). We calculated the titer of **1** (132 mg/L) from HPLC analysis.



Phaseolide A (**1**)

Phaseolide A (**1**): Colorless amorphous, Chemical formula $C_{12}H_{20}O_3$, HRESIMS: *m/z* 235.1304 $[M+Na]^+$ (235.1310 calcd. for $C_{12}H_{20}O_3Na$), $[\alpha]_D^{22}$ +25.6 (c 4.9, MeOH), IR (KBr) ν_{max} (cm^{-1}) 3402, 2936, 2910, 1727, 1447, 1363, 1225, 1218, 1154, 1077, 1075, 1010, 972.

The planal structure was established by the detailed NMR spectral analysis. The COSY correlation from a terminal methyl proton at δ_H 1.26 (d, 6.3, 3H-12) to methylene protons at δ_H 1.78 (m, 2H-6), through an oxymethine proton at δ_H 5.18 (m, H-11), methylene protons at δ_H 2.37 (m, Ha-10) and 2.18 (m, Hb-10), olefin protons at δ_H 5.51 (ddd, 15.1, 10.7, 3.4, H-9) and 5.42 (ddd, 15.1, 9.2, 1.6, H-8) and an oxymethine protone at δ_H 4.10 (m, H-7) indicated the partial structure at C-6-C-12. The another COSY correlations from methylene protons at δ_H 2.35 (m, Ha-2) and 2.25 (m, Hb-2) to a methylene proton at δ_H 1.48 (m, 2H-5), through methylene protons at δ_H 1.80 (m, Ha-3), 1.50 (m, Hb-3) and 1.25 (m, 2H-4) established the other part of the structure at C-2-C-5. The HMBC correlations at H-2 and H-11/C-1 showed the connectivity of C-1 and C-11 through the ester linkage.

Table S3. ^{13}C (125 MHz) and ^{1}H (500 MHz) NMR data for phaseolide A (1)[a,b].

Position	^{13}C	^{1}H (multi, *J* in Hz)
1	173.3	
2	32.8	2.35 (1H, m)
		2.25 (1H, m)
3	24.2	1.80 (1H, m)
		1.50 (1H, m)
4	23.8	1.25 (2H, m)
5	23.2	1.48 (2H, m)
6	34.2	1.78 (1H, m)
		1.51 (1H, m)
7	73.1	4.10 (1H, m)
8	135.5	5.42 (1H, ddd, 15.1, 9.2, 1.6)
9	130.0	5.51 (1H, ddd, 15.1, 10.7, 3.4)
10	40.6	2.37 (1H, m)
		2.18 (1H, dd, 13.9, 10.7)
11	68.3	5.18 (1H, m)
12	20.6	1.26 (3H, d, 6.3)

[*a*] Assignments were based on ^{1}H-^{1}H COSY, HSQC and HMBC experiments.
[*b*] Recorded in $CDCl_3$.

IR and VCD measurement condition

Instrument: BioTools ChiralIR-2X spectrometer, c = 0.18 M, $CDCl_3$ solution, measured at ambient temperature, l = 85 μm, IR: 16 scans, VCD: 4000 scans (resolution 8 cm^{-1}). All spectral data were corrected by a solvent spectrum obtained under the same experimental condition, and presented as Δε and ε (both in M^{-1} cm^{-1}).

IR and VCD calculation condition

Preliminary MMFF conformational search was carried out by CONFLEX 7 software. The obtained geometries were further optimized by DFT/B3LYP/6-311G(d,p)/PCM(chloroform) calculation on a Gaussian16 software. The IR and VCD spectra of the resultant conformers were calculated at the same level of theory. Each spectrum was simulated with Lorentzian lineshapes of 6 cm^{-1} width. The calculated frequencies ν were scaled with were scaled with the factor of 0.985.

Preparation of guest (1)-absorbed crystals

The crystalline sponge (CS)[S3–5] $[(ZnCl_2)_3(tpt)_2 \bullet \boldsymbol{x}(CHCl_3)]_n$ (**2**, tpt = 2,4,6-tris(4-pyridyl)-1,3,5-triazine) was prepared according to the reported procedure[S6] and the solvent was replaced with *t*-butyl methyl ether (MTBE) for 3 days at 50 ºC. A screw-top microvial (Osaka Chemical, cat. no. 11090620), a screw cap with a septum seal (Osaka Chemical, cat. no. 53951-09FB), and a syringe needle (TERUMO, cat. no. NN-2116R) were used for guest inclusion into the porous crystals.

The guest-absorbed crystal **1•2** was prepared by immersing a single crystal of **2** in 50 μL MTBE containing 5 μg of **1**. The solvent evaporation was conducted at 50°C for 15 hours with a needle.

The CS analysis of 1

Single crystal diffraction data were collected on a XtaLAB Synergy-S (Rigaku Oxford Diffraction) diffractometer equipped with a micro-focus Cu Kα radiation source (λ = 1.5418 Å), a hybrid pixel array detector (HyPix), and a low temperature system using cold nitrogen stream (100 K). CrysAlisPro performed data reduction, integration, and scaling. Empirical and numerical absorption corrections were applied in this process. Using Olex2[S7], the crystal structure was solved by using SHELXT (ver. 2018/3[S8]) and refined using SHELXL (ver. 2018/3[S9]) programs. All non-hydrogen atoms were refined anisotropically. Hydrogen atoms were grown using the appropriate crystallographic treatments and refined isotropically using the riding model. A minimum number of restraints and constraints were applied in our structural analysis. Solvent molecules were found in sponges through difference electron density mapping and refined accordingly by using suitable crystallographic treatments. An "Alert A" notification was found in the validation program CheckCIF, derived from the large void space in the crystalline sponge. Appropriate comments are described in the CIF using the validation response form (vrf).

In the asymmetric unit of this crystal structure, seven guest molecules were observed (Fig. 4). The occupancies of Guest 1–7 were 54%, 50%, 38%, 58%, 30%, 28% and 25%, respectively. SAME commands, by using Guest 2 as a template, were applied to Guests 3–7. SIMU commands were used for each guest molecules. Solvent molecules could not be modeled due to the heavy disorder and low occupancy. The absolute configuration of the guest was elucidated to be (7*R*,11*R*)-configuration.

Table S4. Crystal data and structure refinement for 1•2.

Identification code	Phaseolide A (**1**)
Empirical formula	$C_{106.12}H_{104.87}Cl_{12}N_{24}O_{8.53}Zn_6$
Formula weight	2670.54
Temperature/K	100.00(10)
Crystal system	monoclinic
Space group	*C*2
a/Å	34.1049(3)
b/Å	14.43110(10)
c/Å	31.2664(2)
α/°	90
β/°	102.5540(10)
γ/°	90
Volume/Å^3	15020.5(2)
Z	4
ρ_{calc} g/cm^3	1.181
μ/mm^{-1}	3.435
F(000)	5447.0
Crystal size/mm^3	0.266 × 0.113 × 0.068
Radiation	CuKα (λ = 1.54184)
2Θ range for data collection/°	5.792 to 153.274
Index ranges	$-42 \le h \le 42$, $-18 \le k \le 18$, $-39 \le l \le 39$
Reflections collected	142970
Independent reflections	30777 [R_{int} = 0.0302, R_{sigma} = 0.0227]
Data/restraints/parameters	30777/1198/2118
Goodness-of-fit on F^2	1.044
Final *R* indexes [I>=2σ (I)]	R_1 = 0.0572, wR_2 = 0.1681
Final *R* indexes [all data]	R_1 = 0.0641, wR_2 = 0.1757
Largest diff. peak/hole / e Å^{-3}	0.69/-0.52
Flack parameter	0.054(8)

References

S1 F. J. Jin, J. Maruyama, P. R. Juvvadi, M. Arioka, K. Kitamoto, *Biosci., Biotechnol., Biochem.* 2004, **68**, 656–662.

S2 J. Sambrook, E. F. Fritsch and T. Maniatis, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press), 1989.

S3 Y. Inokuma, S. Yoshioka, J. Ariyoshi, T. Arai, Y. Hitora, K. Takada, S. Matsunaga, K. Rissanen, M. Fujita, *Nature* 2013, **495**, 461–466.; Corrigendum: Y. Inokuma, S. Yoshioka, J. Ariyoshi, T. Arai, Y. Hitora, K. Takada, S. Matsunaga, K. Rissanen, M. Fujita, *Nature* 2013, **501**, 262.

S4 Y. Inokuma, S. Yoshioka, J. Ariyoshi, T. Arai, M. Fujita, *Nat. Protoc.* 2014, **9**, 246–252.

S5 M. Hoshino, A. Khutia, H.-Z. Xing, Y. Inokuma, M. Fujita, *IUCrJ.* 2016, **3**, 139–151.

S6 F. Sakurai, A. Khutia, T. Kikuchi, M. Fujita, *Chem. Eur. J.* 2017, **23**, 15035–15040.

S7 O. V. Dolomanov, L. J. Bourhis, R. J. Gildea, J. A. K. Howard, H. Puschmann, *J. Appl. Cryst.* 2009, **42**, 339–341.

S8 G. M. Sheldrick, *Acta Crystallogr. Sect. A* 2015, **71**, 3–8.

S9 G. M. Sheldrick, *Acta Crystallogr. Sect. C* 2015, **71**, 3–8.

DNA sequence of *apmlA*

ATGCTTGTCAGGGAAGTTGTCGATTCAGGGTCCTCGCGGGCCCCGTCGTCAACGAGGCAGTATGTGCAGGAGCCGGTCGCTGT
GGTGGGCATGGCCTGCCGCCTGCCCGGCGACAGCAACACACCGCACAAGTTGTGGGAGTTCCTGATGAAGGGCGGCGTGGCAG
ACACAAACCCGCCATCCACCCGCTTCAACCTCAAGGGCCATTACGACGGGTCGGACAAGCCGACCACCTTGAGAAATCCCGGC
GGCATGTTCCTGGAGAACGTCGACCCGGCCGACTTCGATGCCGGATTCTTCAACATCTCGCGCACCGACGCCATATCCATGGA
CCCACAGCAACGTCAGCTGTTGGAGGTAGTTTATGAGGGCCTTGAGAGCGCAGGGATTCCAATCGAGACGCTTTCCGGAACGC
AATACGGCTGCTTTGTCGGCTCGTACGAAGTTGGTACGTGATAGTCGTCTTGGGCATCCGGCCCGGTGTGGGACAGTGGTAAC
TGACCGGTGATGGGGGTAGACTATGCATGGATGCATGCACGTGATCCTCAGAACCGGCCACCATCGTGCGCAATAGGAGTCGG
GCGGTCGATAATGAGCAACAGGATAAGCCACTTCCTAAATATCAAAGGCCCAAGGTGCGTGCTCATCCATCCATTTAAGATTC
AGACGGCAGCGCCACGTCATGCGCAACATGGTTTCGCGCCATCGCGGACCGTTGTGTCGGGCGCACGTTTCTGGAGGACCCAA
CAACCTGGATCAGCTCTCACTGACTCGTTATGGATCCGTGTGGACAGTTTGACTCTAGATACGGCCTGTTCCGGAAGTCTGGT
AGGTGTCGATGTGGCCTGCCGGTACCTGCAAACGGGCGAGATCTCTGGCGCTATTGTCGCGGGGGCCAACCTGTACATAAAGT
AGGTTTTTTCCCCTTCTTTGGAGTTATTCCCAGCCGAAAGGGGAGAAATTCCCCGATCGGTAACCCGCCGGAGCCAATAGCTG
ATGAGAATGAATCGTCTAGTCCTGACCACAGTATCGAGAGTAAGCAATCCCCAACAGAAACCCCCGTCTATGCTGTAATTTTC
GGGGAAACTGACGGCGATCCTGTTAGACGGGGCGCTGAGAAACGCGCACTCGAAATCTGGCAAATGCCACACATTCGACTCCA
AGGCTGACGGCTACATCAAGGCGGAAGCCATCAACGCTGTTATCCTCAAGCGTCTGGGTGATGCACTCAGGGACGGGGATCCT
ATCAGGGCAGTCATTCGCGGCACTGCCAACAACCATAACGGTAGGACGCCAGGAATCGCCAGTCCGAGCTCTGAAGATCAAGC
ACAGGCGATCCGCGCTGCCTACAGAAACGCTGGCATCGATGATTTTGGCCTGACCAGCTATCTGGAATGCCACGGAACGGGGA
CGAATGCGGGCGATCCGATTGAGGTCAAGGGTATTGCCAGCGTCTTCGCAGATGGCAGAGCAAGGGACAAGCCCCTAGTTATA
GGCTCGGTAAGGGCGGCTCTGTCTTTCGCGCGGGCTCTTTTCGAGTCAACTAGTCGACTGACAGAGCGGTGTGCAGTTAAAAA
GCAATATTGGCCACTCCGAGCCGGCGGCGGGGATATCAGGTCTGCTTAAGGCTGTTCTCACGATCGAACACGGCACCATCCCG
GGGAACCCGACGTTCGAGAAGCCCAACAACAATAGTGAGGTTTTCCCGGTATCGCTTGTCCTGCACATAGCGCTAACTTTGCC
ACAGTCGATTTTGAGGGCCTGAAGGTCAAGCCAATAAGAGCAGCGATCCCTTTTCCTAACGCTCCCTTCCGCCGTGCCAGCGT
AAACTCTTTCGGATTTGGCGGTTCAAATTGCCACGCCATCGTGGACGAACCTAAGGCGTTGATCAAGAACTTTGTACCAGCAC
ACACGTCGTCTTTCAAGTCTTTTGACGACGACGACGATGATGACGATTTTCTCAGCGATGATGAGGCAGCTCCAAGGCCGTAT
CTGCTCCTCTTTTCGGCCAACGATGAGAGTTCGCTCAGGTCTTACACAAAGGAGATGAGTAAGCACCTGAGCAACCTCTCCGT
CAAGGCAAACGTCCGCGACCTCGCCTACACGCTTTCCGAACGAAGAAGTCATCATTTCAACCGCGCATTTGTCATCTCTCGGA
AACCGGTGCTCGATGAGAATGCCCTGGTCTTCGGCAAGAAGGAGCCAAACATTCCGCGAATTGGCTTCGTCTTCACCGGACAG
GGTGCTCAATGGCCTGGTATGGCAAAAGAGCTGTTTGAAACATTTCCAGGCGCCGCTCAGGTGGTAAAAGAGCTCGACACGGT
GCTACGTGCACTGCCGACGCCCCCTGACTGGTCATTGTTTGAGGAGCTGACTGCTCCTCGCTCGGCTGAGCATCTGCGCCAGC
CTGAATTCTCCCAGCCTCTCGTTACTGCGCTCCAGCTTGCCCTTCTCTCAGTACTGAGCTCATGGGGTGTCAAGCCACAGAGC
GTCGTTGGCCACTCCTCGGGGGAAATAGCCGCTGCATATGCCGCTGGATTTCTCACGCAGGCTGATGCAATTGTTGTTGCTTA
CCACAGAGGTTTGGCAGCGAAGCACCGCCAGCAGGACAACGAAACGCCTCTTGGCATGCTTGCTGTCGGTATCGGCGCGGACG
CAATTGGCCCGTATCTCGGGGGTACGGAGACCGTCCAGATCGCCTGCTACAACAGTCCTAAGTCTCTCACACTCTCCGGCGTG
GTCAGCAAACTGGAGCAGGTCAAGTCCCGCCTCACTAAGGATGGCCACTTTGCTCGCATGCTCCAGGTCAACCTCGCCTACCA
CTCAACCTTCATGGACGAAATCGGCAAGGAGTACGAGAGTCTGCTTCGGAGGGACTTCCACGGCCAATCTTGCAAACCAGGCC
CCTTCGGCAGCAACGTGACGATGTTCTCGTCCGTATACGGACACGCCCTTGCGCCCAGCCAGACGACAGATGTCAGCTACTGG
AAGAGCAACATGGTCTCGCCCGTGCGCTTCAACGACGCCTGCCGTGACATGATCGCAGGCAAGGACGGCGCCGACTTCATTAT
CGAGATCGGCCCCGCGGGTGCGCTCGCCGGCCCTGTCGGCCAGATTAAAGACACCATTGACAGCGGCTCCGCCGTTCAGTATG
TTACGGCCCTCAAGCGCGGCGTCGACTCGGTCGACGCCACACTCGACGTCGCCGGCAAGCTGTATCTCAAAGGCGTGCCACTC
AACATGGCCCGCGTCAATGCCGACACGGAGCTGCCCGCGCGCCCGCTCGTCATCACCGACTTGCCTAATTATGCGTGGAACCA
TTCGGTCAAATACTGGTTCGAGTGCAGCGAGAGTAGGGAGTGGCGCCATAAGCAGTTTCCCTGCCATGACCTGCTCGGCTCAA
AGGTCCTCGGCACTCCCTGGCGCAGCCCCTCGTTCAAGCGCATCCTCAAGATCAACGAACTGCCGTGGCTGCGCGACCATAAG
GTAGGTTCGGATATCGTCATGCCGGCGGCAGGCTTCGTAGCTATGGCTGTCGAGGCTATGTACCAGACGCATGTCTCACTGAA
GCCGGAAGAGGGAGTGAGCGCGGTGGGTGAGCTGTGCTATCGGCTGAGGAATATCAGATTCGATAAGGCGCTCGTTGTGGAGG
AGGATGTGCCGGCCGAGTATGTCTTGTCACTGAATCCGCTGGCGGGGACGAAGGATTCGTGGTATGAGTTTAAAATCTCTTCT
ACGAAGGAAGGTGTGATCATGGACCATGCTGGCGGTATGATTCGTGTTGAGGATGCAGCCAAGACGGGTAAGTTTTCTGGTCT
TTTGTAGTCCACAGACCTCGTACTGACAGCCTTTCAGTGGCAACCAAGGAGGAGATTGCTCCTCTCAAGCATACCTCGCCCGG
CCATCTATGGTACAAGATCTCTTCTGAGGTCGGCTATGCCTTCGGCCCGGCTTTCCAGAAGCAGGTCTTGGTTGAAGCCATCG
CCGGGCAGCGCAAGAGCCGCTCCCATATCTCTTTCGAAGAACCTCCCTCCGCCTACCCGCAATCCTCGTACCCCATCCATCCC
GCCGTCCTCGACGGCTGCTTCCAGTCCACCAGCCCGTCCATCTGGGCCGGCAGTCGCAGCACCATGAACGCCGTCCTCATCCC
CGCTATCATCGACGATCTTGTCATCAACAGCTCCGTCCCGCAACACGGCTCGCCGCCACTCGCCCTCGGCGTGGCCGTCGCGC
ACGCCGAGTACTCGGGCCGCGGCCGCCCGGACATGGCCAAGAACTACAGGTCCAGCACCACCGTCTTCCACCCGCACACGGGC
CGGCTGGTCGTGCAGATGACCAACCTGTCGTGCCATCAGCTCGACGTCGGCGTGGACCTGTTCGCCAAACACACGTTCATGCG
GTCCGTCTTCCGCCCGGACATCTCGCTGCTGTGCACCCAGGCGCAGGGCGCGATGCTGCAGCAGCTGGCTGGTGGCGTGCAGG
ACGTGCTCGATCTCGCCGCGCACAGGAAGCCGGCGCTGAAGGTGCTCGAGGTCAATCTCGTGCCGGGCGACACGTCGTGTTTG
TGGTTCGAGGGCAGGGACCGTGCGAGTAGGCTGGCTTATGGTGCGTATAGGTTTTTGGCGTCGGATGCGAGAGATCTGGTGGC
TGTGCAGAATGAACATGGCTCGGAGAGGGGTGCAGGCTTCGGTCTGCTGGATGCTACTAAGGCGGATTTGGGTGCAGGAGATA
AGGAGTTTGACTTCATTGTTGTCAAGATGCGGACGGCGTTGAGGGAGGCGGCGGCGCCTGTATTGAAGGCTGCTGCTGGTTTG
CTTGCCGAGGCTGGGCGTGTGCTTGTTGTCGAGCAGGGCGAGCAGACTTCTTCGGGGAGCAGTGAGGATATGGTCATTGTCGA
GAAGGAGCACGTTGGCGCGGACGCGAAGCTACTGGCTACCATCAAGAGCAGTGGTTTCGGGTCGCATGTTCGTCTCACTTGTG
AGGATGAGGTCAAAAGCGCGTATCTCTGTGCGCCTGCCGTTGAGTTTGCCGCTACTGCTAAGCCGCAGGAACTGTACGTTGTG
CGTCTCGACAAGACTACTCGTCTTCAAGATGGCATCAGGACAGCACTCGCAAAGGCTGGCTGGCGTATCAAGGAGCGAGCGTT
GCCATCCGTCAACGTCCCGGCCAAGAGCACTGTCCTGATCCTCGACGAGCTCTACACCTCCGTCCTCCCCACCGTCACCCCCC
AGCAGTGGCAGACCATCAAGGACATCACCACCCGCGGCTGCAAGCTCCTCTGGGTCACTCAAGGCTCGCACCTGCACGTCACC

AACCCGGAGCGCGCCATGGCCAACGGCCTCTTCCGCACCATCAGCGCCGAAGACCCCAGCGCCCACTGCGTCACCCTCGACGT
CGAGCACGCCTCCTCTCCCGCCACGCCCTCCGCCATCGTCGACGTCCTCGCCACCCTCCGCAACCCCAAGCCCAAATCCGCCT
CCGACTCCGAGTTCTTCCTCCGGGCCGGCCTGCTGCACGTCAACCGTATTGTCCCCGACGGCCCCGTCAACGCCGCCAAGGAC
GGCAGCTCCGCCGCCCCCGTCCGCACCCGCCTCCACGACGCCCGCAGTATCGTCCGCCTGCGCGCCGAGCGCCTCGGCACCTT
TGACGCCCTGCAATGGAGCGCCGTTGCGCCCCCAGACGCACCCCCGCCCCCCGGACACGTCGAGATCGAGCTGTTCGCCGCGG
GGCTCAACTTCAAGGACATTGCCGTCACGATGGGCATCGTGCCCGAGAACGAGCACCTGCTGGGTCTCGAGGGCGCGGGCGTC
GTGCGGCGCGTCGGCGCGGGCGTGTCGGCATTCGCTGTCGGGGATCGCGTCGTAACGATCGAGAAGGGGCTGCTGGCGAATAG
GATACATGCGCCGGAGAAGCGCGTGGTCAAGATCCCGGACGTGATGTCGTTCGAGGAGGCCGCGACGATGCCGGTGGTGTTTG
CGACGGCGGTGTATGCGGTGCTGGAGGTGGGGAGGTTGACAAGCGGGCAGAGTTTGTTGGTGCATTCGGCAGCCGGGGGTGTC
GGGATTGCGTGTTTGCAGTTGGCGAGGTGGATCGGTGCGGAGGTGAGTGCTCTCCCCCCTTCCTTTTTTTTTTTTTTTCATT
TTCGTTCTTTTCTTGCTCGTCTGGATGTGTGGATGAAGGCTGACTCTTGTTTGCTGTTAAGGTCTATATTACCGTCAGCAACG
AGGCGAAGCGTAAGTTCCTTAACGAGACGTTCGGCATCCCGTTCGAGCGCATGTTCTCGTCGCGTGATGCCAGCTTTGCCCCT
GCGATCATGGAGGCGACGAGTGGGAAGGGTATTGATCTGATCATCAACTCCCTGACTGGTGATCTGCTGGATGCGTCGTGGCG
TATCTGCGCAGATGGCGGCACCATGGTGGAGATCGGTAAGAGAGATATTGTTGAAAGGAACTTCCTTTCCATGGAGCCGTTTG
ATAGAAACTGCTCATTCAGGGCAGTCGATATCTCGTACAGCAAGCATATTACAGACGACGTGTATCACAGGTAAGACCCTTGG
GTTCCGAGAGAAGGGCTGGACAGTGCTGACGATATATAGGCTCTTGACGACGGTATTTGATCTGGTCAAGAAAGGTGCCATTG
GCCCTATCAAACCGATCACCACTTTCTCTTTCGCCCAGATCCCGGAGGCTTTCGCGTACATGCGCAGCGGTAGACATATCGGT
AAGATCGTCATCAGTGACGGTCCCAATGGCAATGCCGAGGTCTCCGTGCGTCCCTACAAACGCGACAATCAGCTCCGTGGGGA
CGTAGCCTACTTGATCGCTGGGGGTCTGAAGGGCCTCTGCGGTAGCCTTGCCATTCACATGGCCCGCCGCGGCGCCAAGCACA
TTATCTCCATGTCTCGCAGCGGCTGCTCCGACAGCAGGTCTCAGACCGTGGTCACCAACTGCAACTCGCTCGGCTGCAAGGTA
TACGAAGCTAGGGGCGACGTTGCCAACCTCGCAGACGTCAAGAAGGCCTTCGCCGCCGCCCCGGGTGTCCCCGTTGGCGGCAT
TATCCAGGGCTCCATGGTCCTCCGCGACAAGCCGTACGAAGCCATGACGGTCAAGGAGTACCACGACTGTATCGCCTGCAAGG
TTGCCGGGACATGGAACCTGCACAAAGCCTCGCTCGCCCTCACGCAGCCGCTCGACTTCTTCACCTTACTCTCTTCCATCTCC
GGTGTCGTCGGCAACAAGGGCCAGGCGAACTACGCCGCCGCGAACACCTTCCTTGACGCCTTCGCCGCCTACCGCCGCGCCCA
GGGGCTTGCCGCCAACAGCACCGACCTCGGCGTCATCCACGACGTCGGCTACGTGGCCGAGCAAGGCGGCATGCAAACGCACC
TCGGCGGCAGCCAGTGGACCCCCCTCGCCGAAAACGCCCTCCACCTCGTCCTCGACACCAGCTTCGCGCAGCAGGATCCCGTC
GCACCCATCAACGCCGACTCCGCCGCCCAGCTCGTCACGGGCATCGCCCTCCCCCAGAATCCTGCCGACAGCGACCTCGCCCG
CGACGCCCGCTTCGCCGCCCTCTTCCTCCCGGATAACGCGGGCGGCGGGGCTGGCGGTGCGGGCGGCAACAGGAACAGCGCCG
ACGCCGCGGTGCAGGCTTTCCTGGCGCTGCACGCGAGCGGAGCGGCTCAGGAGGCGCTGCTGCCGGTCGTGCTGGAGGTGCTG
AATGCGCAGCTGACGAGGACGTTGAGGCTGCCCGAGGCACTGGAGCCGGGGAAGCCGCTGAGCGTGTATGGGCTGGACAGCTT
GAGCGCGGTCGAGATCCGGAATTGGATCAGGATGGAACTTGGGGCCGAGGTGACGACGTTAGAGATCGTGGGGGCGGCGAGCT
TGGTGGCGTTGGGGGAGAAGGTCGTTGCGAAGCTGCCGGGGAAGGAGTAA

Amino acid sequence of ApmlA (predicted by using 2ndfind; http://biosyn.nih.go.jp/2ndfind/)

MLVREVVDSGSSRAPSSTRQYVQEPVAVVGMACRLPGDSNTPHKLWEFLMKGGVADTNPPSTRFNLKGHYDGSDKPTTLRNPG
GMFLENVDPADFDAGFFNISRTDAISMDPQQRQLLEVVYEGLESAGIPIETLSGTQYGCFVGSYEVDYAWMHARDPQNRPPSC
AIGVGRSIMSNRISHFLNIKGPRCVLIHPFKIQTAAPRHAQHGFAPSRTVVSGARFWRTQQPGSALTDSLWIRVDSLTLDTAC
SGSLVGVDVACRYLQTGEISGAIVAGANLYINPDHSIENGALRNAHSKSGKCHTFDSKADGYIKAEAINAVILKRLGDALRDG
DPIRAVIRGTANNHNGRTPGIASPSSEDQAQAIRAAYRNAGIDDFGLTSYLECHGTGTNAGDPIEVKGIASVFADGRARDKPL
VIGSLKSNIGHSEPAAGISGLLKAVLTIEHGTIPGNPTFEKPNNNIDFEGLKVKPIRAAIPFPNAPFRRASVNSFGFGGSNCH
AIVDEPKALIKNFVPAHTSSFKSFDDDDDDDDFLSDDEAAPRPYLLLFSANDESSLRSYTKEMSKHLSNLSVKANVRDLAYTL
SERRSHHFNRAFVISRKPVLDENALVFGKKEPNIPRIGFVFTGQGAQWPGMAKELFETFPGAAQVVKELDTVLRALPTPPDWS
LFEELTAPRSAEHLRQPEFSQPLVTALQLALLSVLSSWGVKPQSVVGHSSGEIAAAYAAGFLTQADAIVVAYHRGLAAKHRQQ
DNETPLGMLAVGIGADAIGPYLGGTETVQIACYNSPKSLTLSGVVSKLEQVKSRLTKDGHFARMLQVNLAYHSTFMDEIGKEY
ESLLRRDFHGQSCKPGPFGSNVTMFSSVYGHALAPSQTTDVSYWKSNMVSPVRFNDACRDMIAGKDGADFIIEIGPAGALAGP
VGQIKDTIDSGSAVQYVTALKRGVDSVDATLDVAGKLYLKGVPLNMARVNADTELPARPLVITDLPNYAWNHSVKYWFECSES
REWRHKQFPCHDLLGSKVLGTPWRSPSFKRILKINELPWLRDHKVGSDIVMPAAGFVAMAVEAMYQTHVSLKPEEGVSAVGEL
CYRLRNIRFDKALVVEEDVPAEYVLSLNPLAGTKDSWYEFKISSTKEGVIMDHAGGMIRVEDAAKTVATKEEIAPLKHTSPGH
LWYKISSEVGYAFGPAFQKQVLVEAIAGQRKSRSHISFEEPPSAYPQSSYPIHPAVLDGCFQSTSPSIWAGSRSTMNAVLIPA
IIDDLVINSSVPQHGSPPLALGVAVAHAEYSGRGRPDMAKNYRSSTTVFHPHTGRLVVQMTNLSCHQLDVGVDLFAKHTFMRS
VFRPDISLLCTQAQGAMLQQLAGGVQDVLDLAAHRKPALKVLEVNLVPGDTSCLWFEGRDRASRLAYGAYRFLASDARDLVAV
QNEHGSERGAGFGLLDATKADLGAGDKEFDFIVVKMRTALREAAAPVLKAAAGLLAEAGRVLVVEQGEQTSSGSSEDMVIVEK
EHVGADAKLLATIKSSGFGSHVRLTCEDEVKSAYLCAPAVEFAATAKPQELYVVRLDKTTRLQDGIRTALAKAGWRIKERALP
SVNVPAKSTVLILDELYTSVLPTVTPQQWQTIKDITTRGCKLLWVTQGSHLHVTNPERAMANGLFRTISAEDPSAHCVTLDVE
HASSPATPSAIVDVLATLRNPKPKSASDSEFFLRAGLLHVNRIVPDGPVNAAKDGSSAAPVRTRLHDARSIVRLRAERLGTFD
ALQWSAVAPPDAPPPPGHVEIELFAAGLNFKDIAVTMGIVPENEHLLGLEGAGVVRRVGAGVSAFAVGDRVVTIEKGLLANRI
HAPEKRVVKIPDVMSFEEAATMPVVFATAVYAVLEVGRLTSGQSLLVHSAAGGVGIACLQLARWIGAEVYITVSNEAKRKFLN
ETFGIPFERMFSSRDASFAPAIMEATSGKGIDLIINSLTGDLLDASWRICADGGTMVEIGKRDIVERNFLSMEPFDRNCSFRA
VDISYSKHITDDVYHRLLTTVFDLVKKGAIGPIKPITTFSFAQIPEAFAYMRSGRHIGKIVISDGPNGNAEVSVRPYKRDNQL
RGDVAYLIAGGLKGLCGSLAIHMARRGAKHIISMSRSGCSDSRSQTVVTNCNSLGCKVYEARGDVANLADVKKAFAAAPGVPV
GGIIQGSMVLRDKPYEAMTVKEYHDCIACKVAGTWNLHKASLALTQPLDFFTLLSSISGVVGNKGQANYAAANTFLDAFAAYR
RAQGLAANSTDLGVIHDVGYVAEQGGMQTHLGGSQWTPLAENALHLVLDTSFAQQDPVAPINADSAAQLVTGIALPQNPADSD
LARDARFAALFLPDNAGGAGGAGGNRNSADAAVQAFLALHASGAAQEALLPVVLEVLNAQLTRTLRLPEALEPGKPLSVYGL
DSLSAVEIRNWIRMELGAEVTTLEIVGAASLVALGEKVVAKLPGKE

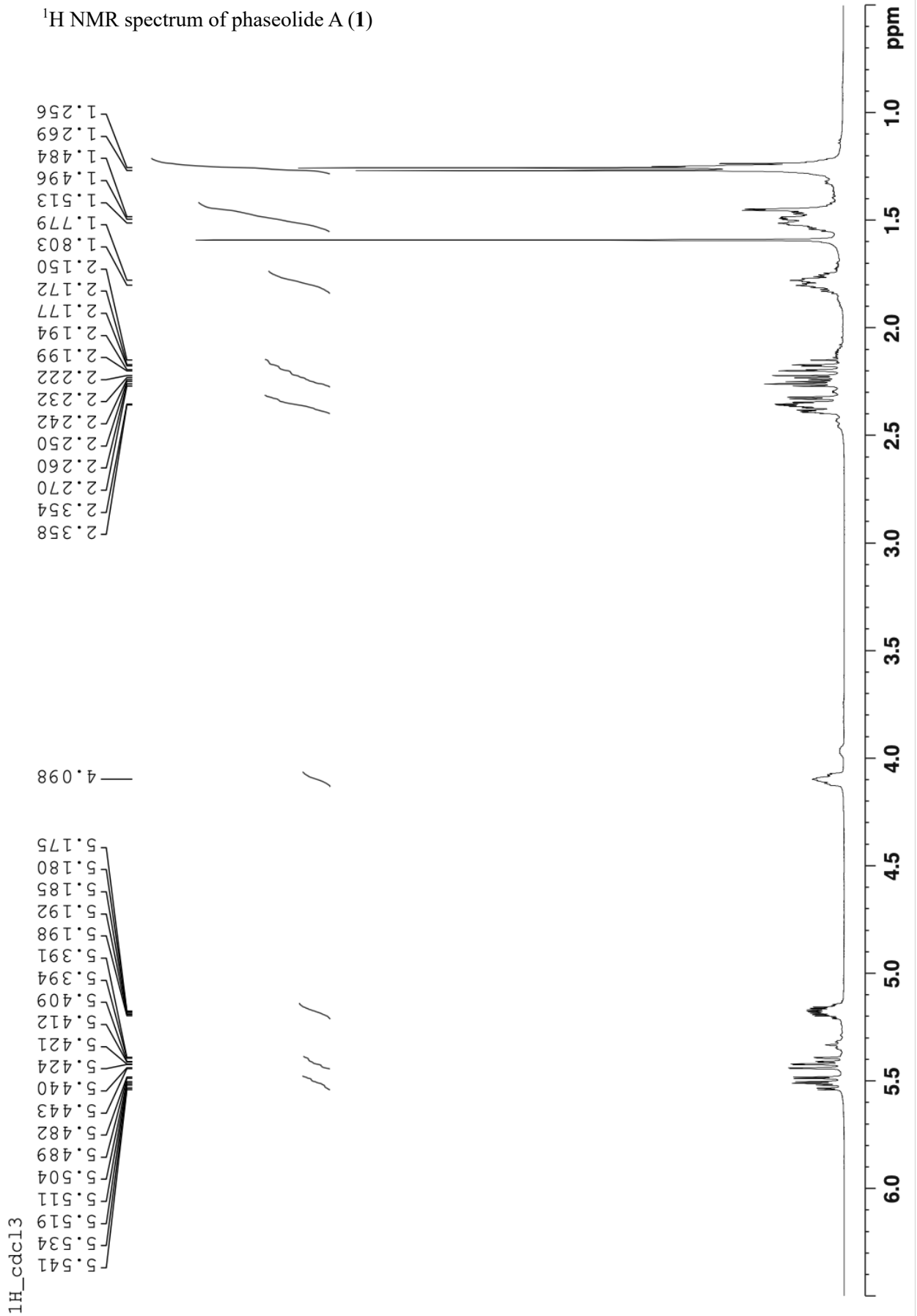
DNA sequence of *apmlB*

ATGGCTGGCAGGCAAGAGGTGGAATTCAAGGCGCTGGACGGCCTAACACTTCGGGGCTGGCTGTATCCAGCGAAAAGCCGAGG
ACCGGGCATTATCATGACACCAGGAGTAAGGAGCAGCCGGACGCCAGGATTGCGGCATGGCGGAAGCATCTGACAGAAGAGCA
GTTCAACATCCCGAAAGACATGATGCTGCCCGAGGTGGCAGAAGCATTCCAGCGCGCAGGCATCACGGCATTGGTGTACGATC
CGCGCAGCATAGGGGCTAGCGACGGCACACCGCGAAACGAGATCGACCCATGCAAGCAGGTCGAGGACTACTCGGATGCGTTG
ACATTCCTAGCCGGGCAAGTGAGCGTTGACAGCAACAAGGTCGGGGTCTGGGGCATGAGCTTTGGGGGCATGGTGGCGGCTTG
CGCGGCTGCGCTGGACAAGCGGATCAAGTTTGTCATCATGGTCTGTCCTCTGGTGCGCCTGTACGACCCGGCCAAGATGGTCA
AGCTGCTGGCGAAGGCGATGAAAGATCGCGAGGCGCAGGCAAAGGGCAACGCGCCGTACACGCTACCACTGCTGACCAAAGCG
GGCGACAACCCGGCAGGCATGAACGCGGCAGGCGGCGCCGAGGCGCTGGAGCTTCTGAACGGGCTGGCGATGCGTTCGCCAGG
ATATGAAAACCGTACGACGCTGCAGTCATACTACCGACTGGCCATGTGGCAGCCGCACGGGCTGCTGCCGATGATCGCTCCCA
CACCGGTGCTGATGGTGGTGCCTGAGAGAGACTCGGTGTCGCCGGCGGCCGAGCAGTCGGCGCTGTTTGACTCGCTGCAGGGC
GCCAAGCGGAAGCATGTCGCCGCTGGCAAGGGCCACTTCAACGTGCTCAGCGGGCTGGCCGAGGCCGGGGGAGAGGCAGACGT
ACAGGCGGTGCAGGAGCGGTTTGTGGGCGACGCACTTGCAGGGCGCGTGGTGGACGGGGAGGACGCCGGCGAGTGTTGA

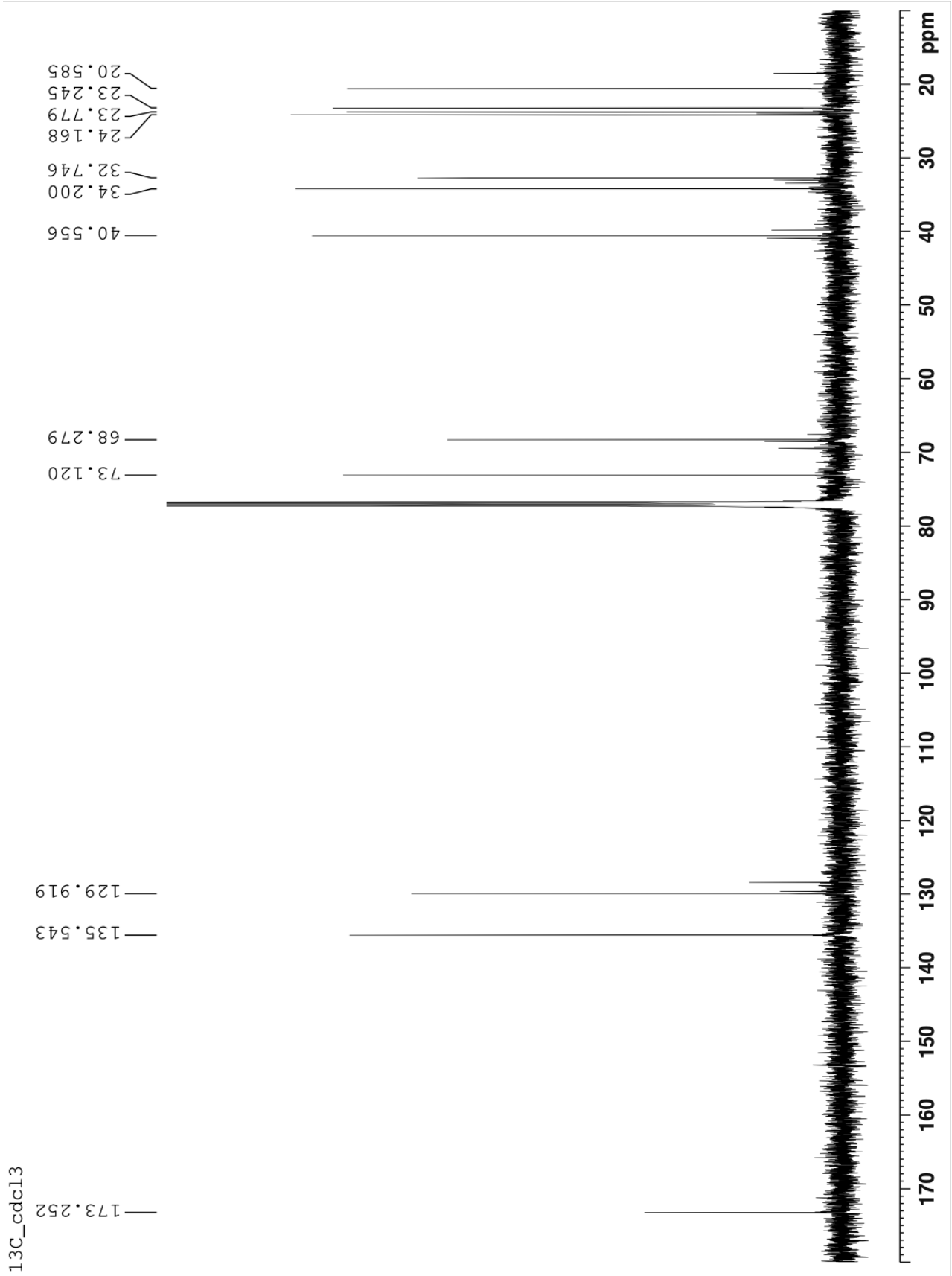
Amino acid sequence of ApmlB (predicted by using 2ndfind)

MAGRQEVEFKALDGLTLRGWLYPAKSRGPGIIMTPGFNIPKDMMLPEVAEAFQRAGITALVYDPRSIGASDGTPRNEIDPCKQVEDYS
DALTFLAGQVSVDSNKVGVWGMSFGGMVAACAAALDKRIKFVIMVCPLVRLYDPAKMVKLLAKAMKDREAQAKGNAPYTLPLLTKAGD
NPAGMNAAGGAEALELLNGLAMRSPGYENRTTLQSYYRLAMWQPHGLLPMIAPTPVLMVVPERDSVSPAAEQSALFDSLQGAKRKHVA
AGKGHFNVLSGLAEAGGEADVQAVQERFVGDALAGRVVDGEDAGEC

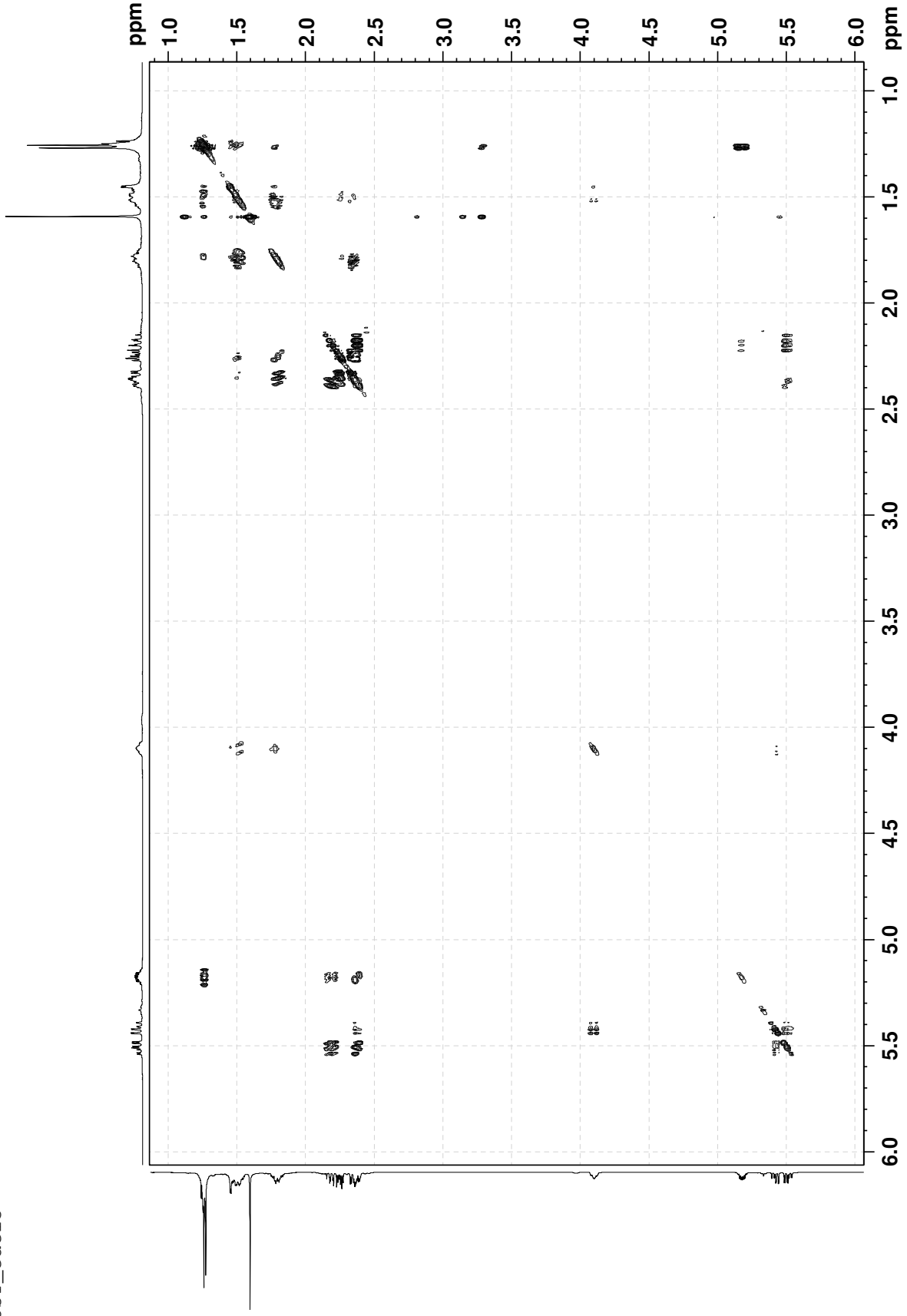
^{1}H NMR spectrum of phaseolide A (**1**)



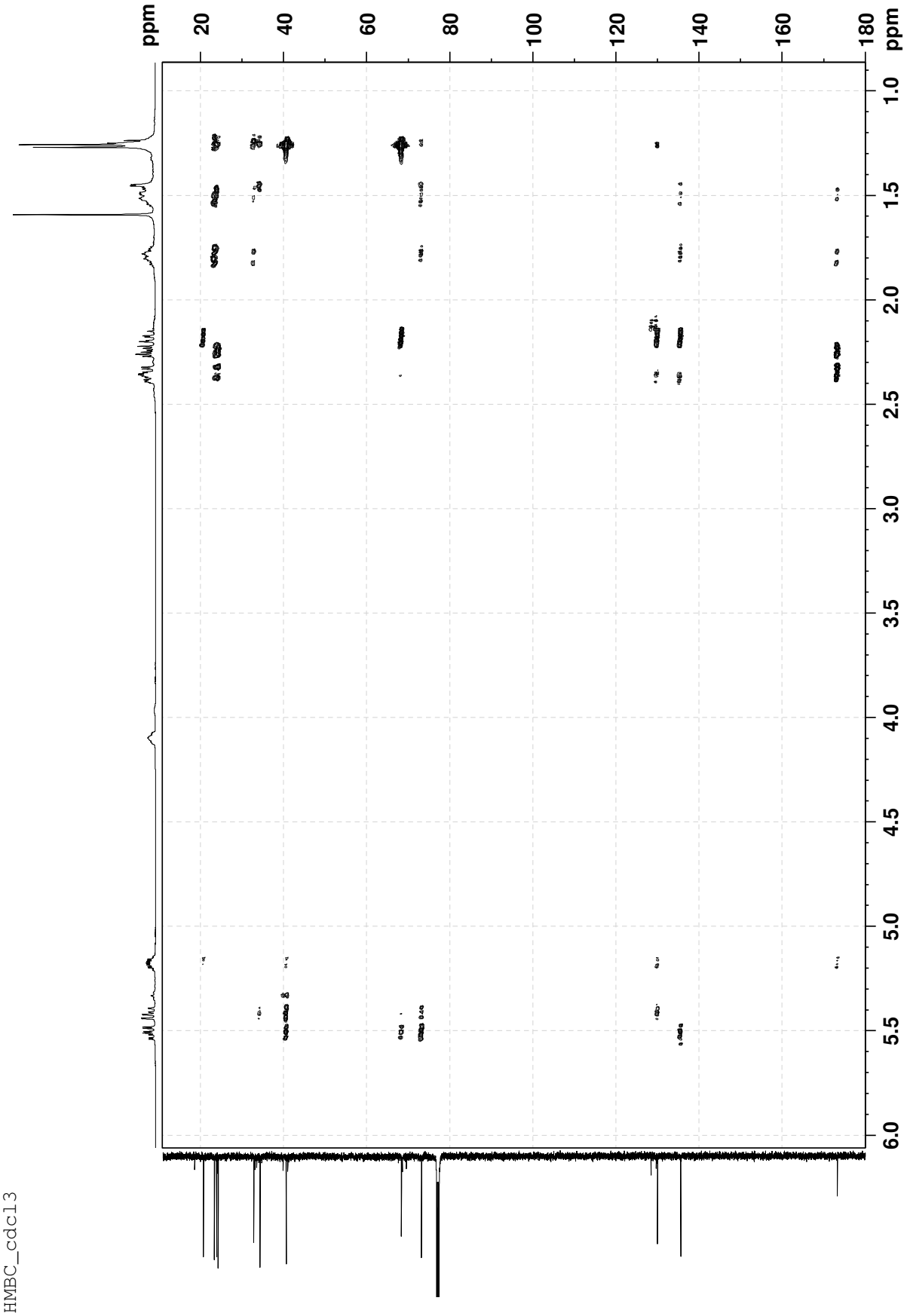
^{13}C NMR spectrum of phaseolide A (**1**)



^{1}H-^{1}H COSY spectrum of phaseolide A (**1**)



HMBC spectrum of phaseolide A (**1**)



HSQC spectrum of phaseolide A (**1**)