Supporting information

Synthesis of (R)-mellein by a Partially Reducing Iterative Polyketide Synthase

*Authors: Huihua Sun, Chun Loong Ho, Feiqing Ding, Ishin Soehano, Xuewei Liu, Zhao-Xun Liang**

School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551, Fax: (+) 65 6791 3856, E-mail: zxliang@ntu.edu.sg

Table of Content:

Experimental details:

Strains, plasmids, and chemicals

Coenzyme A, acetyl-CoA, malonyl-CoA, NADPH, ATP, malonic acid and other chemicals were purchased from Sigma-Aldrich and stored at -20°C. ^{13}C-labeled malonic acid (1, 2, 3-^{13}C, 99%) was purchased from Cambridge Isotope.

TOP10 and BL21(DE3) *E. coli* strains (Novagen) were used for cloning and protein expression respectively. The expression vectors pET-28b(+) and pCDF-2 Ek/LIC were obtained from Novagen. The protein sequences for SACE5532, MatB and NcsB were retrieved from Genbank under the accession numbers of YP_001107644, NP_626687 and AAM77986. The genes that encode SACE5532 and KR_{NcsB} were obtained from GenScript Corporation (NJ, USA) with the DNA sequences optimized for protein expression in *E. coli*. The genes were provides as pUC57-based plasmids with XhoI and NdeI restriction sites added to the C-terminus and N-terminus to facilitate cloning. The plasmid that contains the Sfp-encoding gene was obtained from Christopher Walsh's lab at Harvard Medical School.

Cloning and site-directed mutagenesis

The plasmid pUC57- SACE5532 was digested with NdeI and XhoI restriction enzymes. The 5.2 kb DNA fragment was gel-purified and ligated into pET-28b(+) to give pET28-SACE5532. The gene encoding Sfp was from Christopher Walsh's lab at Harvard Medical School and cloned into pCDF-2 plasmid to give pCDF-Sfp. Likewise, the plasmid pUC57-MatB was digested with NdeI and XhoI restriction enzymes to yield the 1.5-kb *matB* gene fragment, which was gel-purified and ligated into pET-28b(+) to give pET28-MatB. The plasmid harbouring the *matB* gene was sequenced and transformed into *E. coli* strain BL21(DE3) for protein expression.

To generate the SACE5532-C184A mutant, the PCR reaction was performed using pET28-SACE5532 as the template with the following primers: C184A (forward) 5'-CCTGACCATTGATACTGCTGCCGCGGGCAGC -3' / (reverse) 5'-GCTGCCCGCGGCAGCAGTATCAATGGTCAGG -3'. Successful reaction mixtures consists of 100 ng template DNA, 300 nM primer, 300 mM dNTPs, , 1 × KAPAHiFi™ Fidelity Buffer, and 0.5 unit of KAPAHiFi™ DNA Polymerase in a final volume of 25 µl. The PCR program was as following: initial denaturing at 96°C for 5 minutes, followed by 18 cycles at 96°C for 50 seconds, 54°C for 50 seconds, and 68°C for 6 minutes, and completed by an additional 7 minutes at 68°C. Upon completion, 1 µl (10 U) of DpnI was added directly to the PCR mixture and digested at 37 °C for 2 hours. An aliquot (5 µl) of the mixture was used for the transformation of *E. coli* Top10 competent cells. The cells were plated on LB supplemented with 50 µg/ml kanamycin. The mutation was confirmed by DNA sequencing of the pET28- SACE5532-C184A plasmid.

To clone the stand-alone KR domain of SACE5532, the PCR reaction was performed using pUC57-SACE5532 as a template with the following primers: SACE5532-KR (forward) 5'-TTTCATATGCGCGATCTGGCGTATGAAATCATTTGG -3' / (reverse) 5'-TTTCTCGAGGCCGGTATCGCCGCTCGCGGTCAGTTC -3'. Successful reaction mixtures consisted of 52 ng template DNA, 300 nM primer, 300 mM dNTPs, 1× KAPAHiFi™ Fidelity Buffer, and 1 unit of KAPAHiFi™ DNA Polymerase in a final volume of 50 µl. The PCR program was as follows: initial denaturing at 95°C for 5 minutes; 6 cycles at 95°C for

30 seconds, 45°C for 30 seconds, and 72°C for 40 seconds; 30 cycles at 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 40 seconds; and an additional 7 minutes at 72°C. The 1.5-kb PCR product was gel-purified and digested with NdeI and XhoI restriction enzymes. The digested SACE5532-KR fragment was gel-purified and ligated into pET-28b(+) to give pET28-SACE5532KR. The plasmid harbouring the SACE5532-KR was sequenced and transformed into the *E. coli* strain BL21(DE3) for protein expression.

To clone the stand-alone KR domain of NcsB, the PCR reaction was performed using pUC57-NcsB-TH-KR as a template with the following primers: NcsB-KR (forward) 5'-TTTCATATGAGCGAACTGGTTCACGAAATCGTCTGG -3' / (reverse) 5'-TTTCTCGAGGCCATCCGTTTCGCCAGACACCGGCAG -3'. Successful reaction mixtures consisted of 100 ng of template DNA, 300 nM each primer, 300 mM dNTPs, , 1× KAPAHiFi™ Fidelity Buffer, and 1.0 U of KAPAHiFi™ DNA Polymerase in a final volume of 50 µl. The PCR program was as follows: initial denaturing at 95°C for 5 minutes; 6 cycles at 95°C for 30 seconds, 45°C for 30 seconds, and 72°C for 40 seconds; 30 cycles at 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 40 seconds; and an additional 7 minutes at 72°C. The 1.5-kb PCR product was gel-purified and digested with NdeI and XhoI restriction enzymes. The digested SACE5532-KR fragment was gel-purified and ligated into the identical sites of pET-28b(+) to give pET28-NcsB-KR. Then the plasmid harbouring the gene fragment was sequenced and transformed into *E. coli* strain BL21(DE3) for protein expression.

Protein expression and purification

Co-expression of SACE5532 and Sfp - pET28-SACE5532 and pCDF-Sfp were co-transformed into *E. coli* BL21(DE3) competent cells. The cells were plated on LB medium supplemented with 50 µg/ml kanamycin and 50 µg/ml streptomycin. The colonies were screened by PCR to confirm the presence of both pET28-SACE5532 and pCDF-Sfp plasmids. A single colony was used to inoculate 20 ml of LB medium containing both kanamycin (50 µg/ml) and streptomycin (50 µg/ml). The culture was kept in a shaking incubator overnight at 37°C at 200 rpm. A 5 ml aliquot was transferred to 500 ml of LB medium that is supplemented with 10% glycerol, kanamycin (50 µg/ml) and streptomycin (50 µg/ml). The cells were kept at 37°C at 200 rpm till the OD_{600} reached ~0.5 (ca. 4 hours) before the culture was cooled down to 16°C and induced with 0.8 mM IPTG. After incubation at 16°C for an additional 20 hours at 130 rpm, cells were harvested by centrifuging at 8,000 rpm. The cell pellet was re-suspended in the lysis buffer [50 mM NaH_2PO_4 (pH 8.0), 300 mM NaCl, 20 mM imidazole, 5 mM β-ME and 10% (v/v) glycerol] and lysed by sonication. After centrifugation at 20,000 rpm for 30 minutes at 4°C, the supernatant was filtered by 0.45 µm membrane and loaded onto a HiTrap™ Ni^{2+}-NTA column (GE Healthcare). The column was then washed with the lysis buffer and the washing buffer that contains 40 mM imidazole before the elution with the elution buffer that contains 500 mM imidazole. The co-eluted SACE5532 and Sfp were separated and purified by size-exclusion chromatography using a HiLoad™ 16/60 Superdex™ 200 column (GE Healthcare). Proteins were desalted and exchanged into the Tris buffer [50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM DTT and 10% (v/v) glycerol]. The protein was concentrated, flash frozen in liquid nitrogen and stored in -80°C freezer.

Expression and purification of Sfp - pCDF-Sfp was transformed into *E. coli* BL21(DE3) competent cells. The cells were plated on LB medium supplemented with 50 µg/ml streptomycin. A single colony was used to inoculate 20 ml of LB medium supplemented with

50 µg/ml streptomycin. The culture was kept in a shaking incubator overnight at 37°C at 200 rpm. A 5 ml aliquot was transferred to 500 ml of LB medium supplemented with 50 µg/ml streptomycin. The culture was kept at 37°C at 200 rpm till the OD_{600} reached ~0.6 (~3 hours) before it was cooled down to 16°C and induced with 0.2 mM IPTG. After incubation at 16°C for an additional ~20 hours at 130 rpm, cells were harvested. The cell pellet was re-suspended in the lysis buffer [50 mM NaH_2PO_4 (pH 8.0), 300 mM NaCl, 20 mM imidazole, 5 mM β-ME and 10% (v/v) glycerol] and lysed by sonication. After centrifugation at 20,000 rpm for 30 minutes at 4°C, the supernatant was filtered by 0.45 µm membrane and loaded onto a HiTrap™ Ni^{2+}-NTA column. The column was washed by the lysis buffer and the washing buffer containing 40 mM imidazole before the proteins was eluted out with the elution buffer containing 500 mM imidazole. The eluted protein was further purified by gel filtration using a HiLoad™ 16/60 Superdex™ 200 column. Proteins were desalted and exchanged into the Tris buffer [50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 mM DTT and 10% (v/v) glycerol]. The protein was concentrated, flash frozen in liquid nitrogen, and stored in -80°C freezer.

Expression and purification of MatB, $KR_{SACE5532}$ and KR_{NcsB} - The expression and purification of MatB, SACE5532-KR and NcsB-KR were similar to the procedure described above for Sfp, except that cells were grown in LB media supplemented with 50 µg/ml kanamycin.

In vitro enzymatic assay of the PKS activity

For the enzymatic assays of the PKS activity of SACE5532, a typical reaction mixture contains 3 µl of $MgCl_2$ (1 mM), 8 µl of CoA (50 mM), 70 µl of SACE5532 (11.3 mg/ml), 20 µl of Sfp (27.9 mg/ml) and 93 µl of reaction buffer (50 mM Tris (pH 8.5), 150 mM NaCl and 1 mM DTT). After incubation at 30°C for 20 minutes, the reaction mixture was supplemented with 5 µl of NADPH (10 mM), 1µl of acetyl-CoA (100 mM) and 2 µl of malonyl-CoA (100 mM) and incubated at 30°C. The reaction was quenched with 5 µl of 6 mM HCl and the solution was vortexed to precipitate the enzyme. Then the mixture was spun at 14,800 rpm for 10 min and the supernatant was subjected to HPLC analysis. HPLC analysis was performed with an analytical eclipse XDB C18 column (4.6 × 150 mm) using an Agilent 1200 HPLC system. A full gradient was employed from 100% buffer A (HPLC grade water with 0.045% TFA) to 40% buffer A + 60% Buffer B (100% acetonitrile with 0.045% TFA) at 1 ml/min in 60 minutes.

For the kinetic analysis of enzymatic reaction, the reaction was followed by monitoring the NADPH absorbance (340 nm) continuously by using a UV-Vis spectrophotometer. The reaction was also followed by monitoring the formation of the product by HPLC. The reaction was initiated (200 µl solution) under the same conditions as above. Aliquots (20 µl) of the solution were taken from the reaction mixture at 0, 15, 30, 45, 60, 75, 90 and 120 minute. The aliquots were quenched with 5 µl of 6 mM HCl and vortexed to precipitate the enzymes. Then the mixture was spun at 14,800 rpm for 10 minutes and the supernatant was loaded for HPLC analysis.

In vivo production of the enzymatic product

To find out whether mellein can be generated by the *E. coli* cells that co-expressing SACE5532 and Sfp, SACE5532 and Sfp were co-expressed in the BL21(DE3) *E. coli* cells under similar conditions described above for protein expression, except that the cells were supplemented with 10 mM $MgCl_2$ upon induction. After overnight induction, the cells were

centrifuged at 8,000 rpm for 10 min. The supernatant was extracted twice with equal volume of ethyl acetate. The combined organic extract was dried over anhydrous $MgSO_4$, concentrated in vacuum, and re-dissolved in methanol for HPLC analysis.

Large scale enzymatic reaction for the preparation of enzymatic product

For the large scale reaction, malonyl-CoA was synthesized *in situ* by using the malonyl-CoA synthase MatB. The MatB reaction was carried out in 60 ml of 100 mM HEPES buffer (pH 8.5) that contains 20 mM malonic acid, 10 mM MgCl2, 5 mM ATP, 1 mM Coenzyme A and 30 mg MatB. The reaction was incubated at 23°C overnight and analyzed by HPLC to ensure completion of the reaction. The PKS reaction was set up by mixing the reaction mixture from above with 400 µl of 100 mM acetyl-CoA, 1.5 ml of 10 mM NADPH and 3 ml of SACE5532 and Sfp protein mixture (~ 60 mg). After incubation of the reaction in 30°C waterbath for two hours, additional 1.5 ml of 10 mM NADPH was added to the reaction mixture. The reaction was incubated in 30°C water bath for another 2 hours. The final reaction mixture was extracted twice with ethyl acetate (2 × 60 ml). The combined organic extract was dried over anhydrous magnesium sulfate and concentrated in vacuum. The compound was purified by preparative thin layer chromatography (TLC) with the solvent system of hexane: ethyl acetate (5:1 by volume). The pure fractions collected from preparative TLC were collected for structure characterization by LC-MS and NMR spectroscopy.

The preparation of ^{13}C-labeled enzymatic product followed essentially the same protocol described above except that ^{13}C-labeled malonic acid was used to replace the normal malonic acid substrate. TLC purification of the ^{13}C-labelled product was performed following the same procedure as described above.

Liquid chromatography coupled Mass spectrometry (LC-MS)

High-resolution LC-MS was performed by using a Michrom Rp18 column (0.1 × 50 mm). The gradient employed in the analysis was from 99% buffer A (HPLC grade water with 0.1% FA) + 1% buffer B (100% acetonitrile with 0.1% FA) to 40% buffer A + 60% buffer B in 30 minutes. The flow rate was set at 1 µl/min and the UV detector was set at 314 nm. The ionization energy was set with Nanospray Ionization source for the Finnigan LTQ Orbitrap mass spectrometer (Thermo Electron). The results were analyzed with the software Xcalibur for the deduction of the plausible molecular formulas based on the observed *m/z*.

NMR spectroscopy

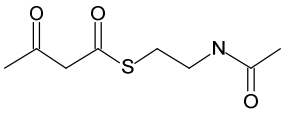
One-dimensional ^1H, ^{13}C NMR and two-dimensional NOESY spectra were collected on a Bruker 400 MHz NMR spectrometer (Bruker DPX 400) with $CDCl_3$ as the solvent and TMS as the internal reference. The colorless powder form of the enzymatic product obtained from the large scale *in vitro* reactions was used for NMR analysis. The chemical shifts and coupling constants for the ^1H NMR spectrum [$CDCl_3$, 400 MHz, Supporting Figure S5] are: δ 4.73 (1H, m, H-3), δ 2.93 (2H, d, J = 7.2 Hz, H-4), δ 6.69 (1H, J = 7.2 Hz, H-5), δ 7.41 (1H, t, J = 7.8 Hz, H-6), δ 6.89 (1H, d, J = 8.4 Hz, H-7), δ 1.53 (3H, d, J = 6.4 Hz, H-11), δ 11.03 (1H, s, OH-C8). The chemical shifts for the ^{13}C NMR spectrum [$CDCl_3$, 400 MHz, Supporting Figure S6] are: δ 169.94 (C-1), δ 76.09 (C-3), δ 34.63 (C-4), δ 117.88 (C-5), δ 136.13 (C-6), δ 116.27 (C-7), δ 162.23 (C-8), δ 108.32 (C-9), δ 139.38 (C-10), δ 20.76 (C-11). 2D ^1H, ^1H-NOESY NMR [$CDCl_3$, 400 MHz, Supporting Figure S7]: (H-7, H-6), (H-6, H-5),

(H-5, H-4), (H-4, H-3), (H-4, H-11), (H-11, H-3). The ^{1}H NMR and ^{13}C NMR are also in excellent agreement with those obtained for a synthesized mellein.

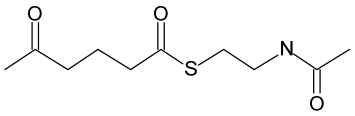
Measurement of the specific rotation by polarimeter

The specific rotation $[\alpha]_D^{25}$ of the enzymatic product was measured by using a polarimeter as -123^o (c 0.43, $CHCl_3$, 25 ^oC), which is similar to the specific ration measured for the synthetic (R)-(-) -mellein ($[\alpha]_D^{22} = -102^o$ (c 0.53, $CHCl_3$), Islam et al, Tetrahedron 63 (2007) 1074-1079).

Synthesis of N-acetylcysteamine (SNAC) thioesters.



Acetoacetyl-SNAC (15). The diketide analog acetoacetyl-SNAC was prepared according to the reported procedure (Gilbertt, I et al, *Bioorg. Med. Chem. Lett.*, **1995**, *5*, 1587-1590). Analytical thin layer chromatography (TLC) was performed using pre-coated silica gel plate. Visualization was achieved by UV light (254 nm) and/or $KMnO_4$ stain. Flash chromatography was performed using silica gel and a gradient solvent system (EtOAc:hexane as eluent). NMR spectra were recorded at room temperature on Bruker DPX 400 spectrometers with $CDCl_3$ as the solvent and TMS as the internal reference. ^{1}H NMR spectrum [$CDCl_3$, 400 MHz]: δ 5.95 (1H, s, NH), δ 2.27 (3H, s, H-1), δ 3.71 (2H, m, H-3), δ 3.10 (2H, m, H-5), δ 3.46 (2H, m, H-6), δ 1.97 (3H, s, H-8). ^{13}C NMR spectrum [$CDCl_3$, 400 MHz]: 199.86, 192.29, 170.47, 58.03, 39.17, 30.29, 29.25, 23.17.



5-oxohexanoyl-SNAC (16). The 4-acetylbutyric acid (65 mg, 0.5 mmol) was dissolved in 5 ml of DMF at 0°C and then treated with diphenylphosphoryl azide (163 µl, 0.75 mmol) and triethylamine (129 µl, 1 mmol) for 2 h with stirring. N-acetylcysteamine (HSNAC, 64 µl, 0.6 mmol) was added to the solution. The mixture was stirred at room temperature for an additional 3 h. The reaction was quenched with the addition of 25 ml of H_2O and extracted twice with ethyl acetate. The organic layer was dried, and the title compound was purified with silica gel chromatograph to give 40 mg of solid powder (34% yield): ^{1}H NMR ($CDCl_3$, 400 MHz): δ 6.03 (br, 1H), 3.40 (q, J = 6.0 Hz, 2H), 3.00 (t, J = 6.4 Hz, 2H), 2.58 (t, J = 7.2 Hz, 2H), 2.49 (t, J = 6.4 Hz, 2H), 2.12 (s, 3H), 1.95 (s, 3H), 1.91 (t, J = 7.2 Hz, 2H); ^{13}C NMR ($CDCl_3$, 100 MHz): δ 207.8, 199.4, 170.4, 42.9, 42.1, 39.5, 29.9, 28.6, 23.2, 19.4.

Enzymatic assay of the stand-alone KR domains

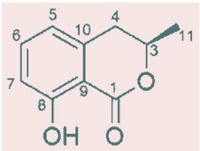
The *in vitro* assays of the ketoreductase (KR) domain activity were conducted by using a semi-micro quartz cuvette and a Shimazu UV-Vis 1700 spectrophotometer. When *trans*-1-decalone was used as the substrate, a typical enzymatic reaction contained 0.82 mg/ml

$KR_{SACE5532}$ or KR_{NcsB} protein, 0.25 mM NADPH and 10 mM *trans*-1-decalone in 100 mM Tris buffer (pH 8.0) in a total volume of 200 µl. The reaction was incubated at 37°C within the sample chamber through an external temperature controller. When S-Ethyl acetothioacetate was used as the substrate, a typical enzymatic reaction contains 0.02 mg/ml $KR_{SACE5532}$ or KR_{NcsB}, 0.25 mM NADPH and 3.7 mM S-Ethyl acetothioacetate in 100 mM HEPES buffer (pH 8.5) in a total volume of 200 µl. The reaction was incubated at 20°C within the sample chamber through an external temperature controller. When acetoacetyl-SNAC was used as the substrate, a typical enzymatic reaction contains 1.46 mg/ml $KR_{SACE5532}$ or KR_{NcsB}, 0.25 mM NADPH and 3.3 mM acetoacetyl-SNAC (**15**) or 5-oxohexanoyl-SNAC (**16**) in 100 mM HEPES buffer (pH 8.5) in a total volume of 200 µl. The reaction was incubated at 30°C within the sample chamber through an external temperature controller. The reaction progress was monitored continuously by recording the NADPH absorbance at 340 nm.

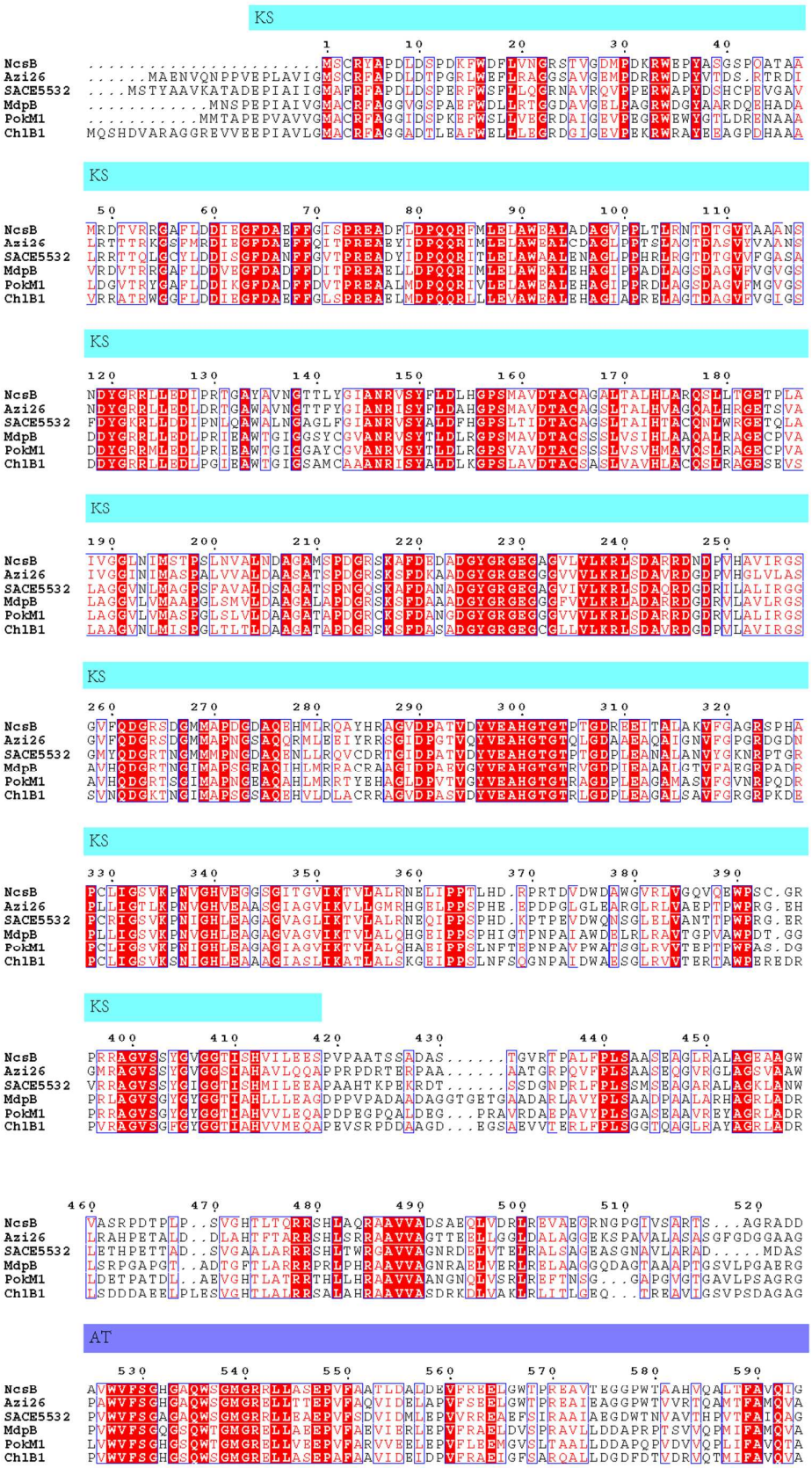
Table S1. SACE3352 and homologous bacterial iPKSs.

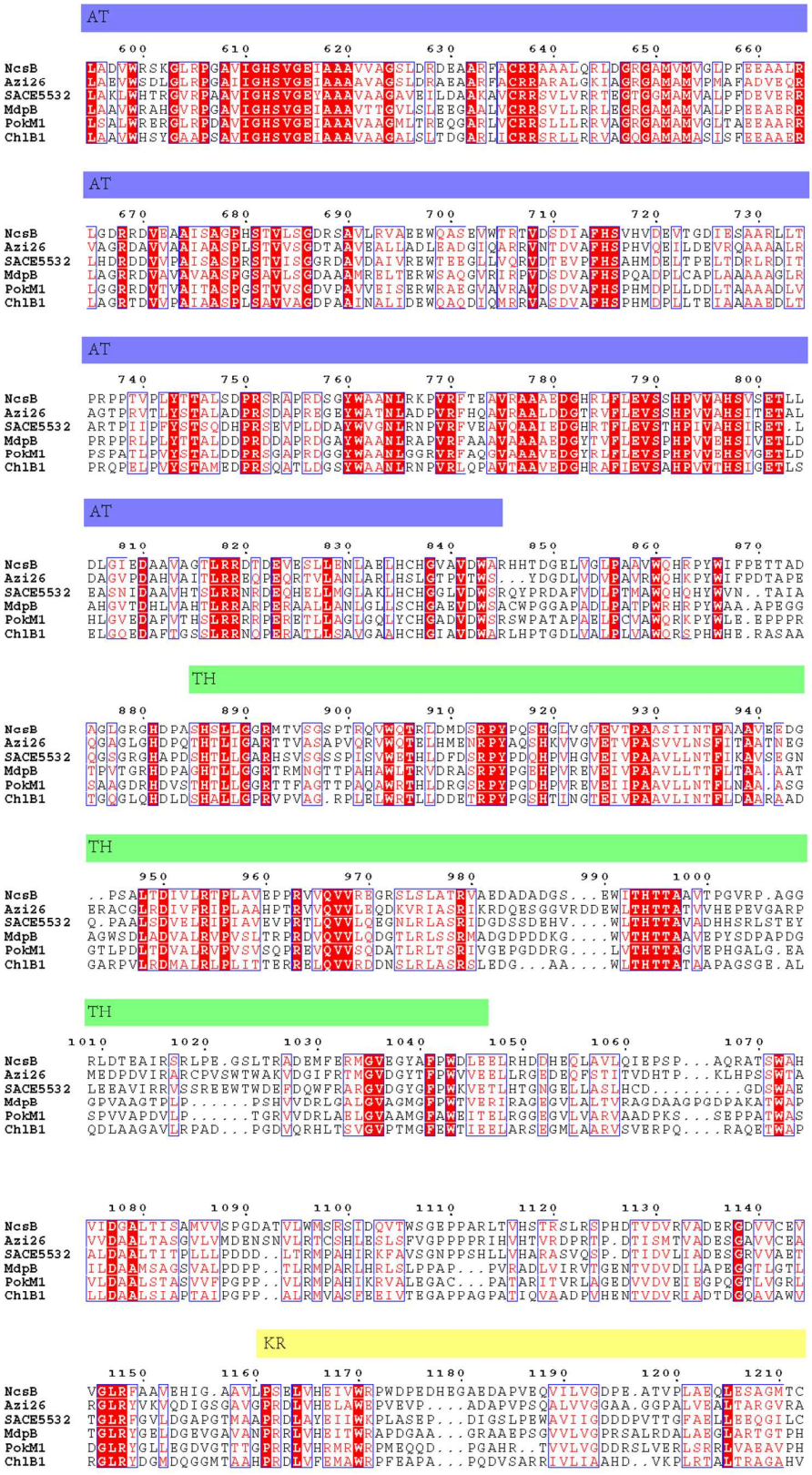
Gene	Ref	Access no.	Organism	Identity%/Homology%
NcsB	(*27, 28*)	AAM77986	*Streptomyces carzinostaticus subsp. neocarzinostaticus*	51/66
AziB	(*29*)	ABY83164	*Streptomyces sahachiroi*	48/64
PokM1	(*26*)	ACN64831	*Streptomyces diastatochromogenes*	49/63
ChlB1	(*23, 24*)	AAZ77673	*Streptomyces antibioticus*	46/61
MdpB	(*25*)	ABY66019	*Actinomadura madurae*	47/60
CalO5	(*44*)	AAM70355	*Micromonospora echinospora*	48/62
AviM	(*45, 46*)	AAK83194	*Streptomyces viridochromogenes*	48/61

Table S2. NMR data for the SACE5532 produced mellein ($CDCl_3$ at 400 MHz).



Carbon No.	δ ^{1}H (ppm)	δ ^{13}C (ppm)
1		169.94
3	4.73 (1H, m)	76.09
4	2.93 (2H, d, *J* = 7.2 Hz)	34.63
5	6.69 (1H, d, *J* = 7.2 Hz)	117.88
6	7.41 (1H, t, *J* = 7.8 Hz)	136.13
7	6.89 (1H, d, *J* = 8.4 Hz)	116.27
8		162.23
9		108.32
10		139.38
11	1.53 (3H, d, *J* = 6.4 Hz)	20.76
8-OH	11.03 (1H, s)	


KS
KS
KS
KS
KS
KS
AT
NcsB
Azi26
SACE5532
MdpB
PokM1
ChlB1


AT
600 610 620 630 640 650 660
NcsB LADVWRSKGLRPGAVIGHSVGEIAAAVVAGSLDRDEAARFACRRAAALQRLDGRGAMVMVGLPFEEAALR
Azi26 LAEVWSDLGLRPGAIIGHSVGEIAAAAVAGSLDRAEAARFACRRARALGKIAGRGAMAMVPMAFADVEQR
SACE5532 LAKLWHTRGVRPAAVIGHSVGEYAAAVAAGAVELILDAAKAVCRRSVLVRRTEGTGGMAMVALPFDEVERR
MdpB LAAVWRAHGVRPGAVIGHSVGEIAAAVTTGVLSLEEGAALVCRRSVLLRRLEGRGAMAMVALPPEEAERR
PokM1 LSALWRERGLRPDAVIGHSVGEIAAAVAAGMLTREQGARLVCRRSLLLRRVAGRGAMAMVGLTAEEAARR
ChlB1 LAAVWWHSYGAAPSAVIGHSVGEIAAAVAAGALSLTDGARLICRRSRLLRRVAGQGAMAMASISFEEAAER
AT
670 680 690 700 710 720 730
NcsB LGDRRDVEAAISAGEHSTVLSGDRSAVLRVAEEWQASEVWTRTVDSDIAFHSVHVDEVTGDIESAARLLT
Azi26 VAGRDAVVAPIAASPLSTVVSGDTAAVEALLADLEADGIQARRVNTDVAFHSPHVQEILDEVRQAAAALR
SACE5532 LHDRDDVVPAISASRSTVISGGRDAVDAIVREWTEEGLLVQRVDTEVPFHSAHMDELTPELTDRLRDIT
MdpB LAGRRDVAVAVAASPGSAVLSGDAAAMRELTERWSAQGVRIRPVDSDVAFHSPQADPLCAPLAAAAAGLR
PokM1 LGGRRDVTVAITASPGSTVVSGDVPAVVEISERWRAEGVAVRAVDSDVAFHSPHMDPLLDDLTAAAADLV
ChlB1 LAGRTDVVPAIAASPLSAVVAGDPAAINALIDEWQAQDIQMRRVASDVAFHSPHMDPLLTEIAAAAEDLT
AT
740 750 760 770 780 790 800
NcsB PRPPTVPLYTTALSDPRSRAPRDSGYWAANLRKPVRFTEAVRAAAEDGHRIFLEVSSHPVVAHSVSETLL
Azi26 AGTPRVTLYSTALADPRSDAPREGEYWATNLADPVRFHQAVRAALDDGTRVFLEVSSHPVVAHSITETAL
SACE5532 ARTPIIPFYSTSQDHPRSEVPLDDAYWVGNLRNPVRFVEAVQAAAIEDGHRTFLEVSTHPIVAHSIRETLL
MdpB PRPPRLPLYTTALDDPRDDAPRDGAYWAANLRAPVRFAAAVAAAAEDGYTVFLEVSPHPVVEHSIVETLD
PokM1 PSPATLPVYSTALDDPRSGAPRDGGYWAANLGGRVRFAQGVAAAVEDGYRIFLEVSPHPVVEHSVGETLD
ChlB1 PRQPELPVYSTAMEDPRSQATLDGSYWAANLRNPVRLQPAVTAAVEDGHRAFIEVSAHPVVTHSIGETLS
AT
810 820 830 840 850 860 870
NcsB DLGIEDAAVAGTLRRDTDEVESLLENLAELHCHGVAVDWARHHTDGELVGLPAAVWQHRPYWIFPETTAD
Azi26 DAGVPDAHVAITLRREQPEQRTVLANLARLHSLGTPVTWS...YDGDLVDVPAVRWQHKPYWIFPDTAPE
SACE5532 EASNIDAAVHTSLRRNRDEQHELLMGLAKLHCHGGLVDWSRQYPRDAFVDLPTMAWQHQHYWVN..TAIA
MdpB AHGVTDHLVAHTLRRAREPERAALLANLGLLSCHGAEVDWSACWPGGAPADLPATPWRHRPYWAA.APEGG
PokM1 HLGVEDAFVTHSLRRRRPERETLLAGLGQLYCHGADVDWSRSWPATAPAELPCVAWQRKPYWLE.EPPPR
ChlB1 ELGQEDAFTGSSLRRNQPERATLLSAVGAAHCHGIAVDWARLHPTGDLVALPLVAWQRSPHWHE.RASAA
TH
880 890 900 910 920 930 940
NcsB AGLGRGHDPASHSLLGGRMTVSGSPTRQVWQTRLDMDSRPYPQSHQLVGVEVTPAASIINTFAAAVEEDG
Azi26 QGAGLGHDPQTHTLIGARTTVASAPVQRVWQTELHMENRPYAQSHKVVGVETVPASVVLNSFITAATNEG
SACE5532 QGSGRGHAPDSHTLLGARHSVSGSSPISVWETHLDFDSRPYPDQHPVHGVEIVPAAVLLNTFIKAVSEGN
MdpB TPVTGRHDPAGHTLLGGRTRMNGTTPAHAWLTRVDRASRPYPGEHPVRENVEIVPAAVLLTTFLTAAA.AAT
PokM1 SAAGDRHDVSTHTLLGGRTTFAGTTPAQAWRTHLDRGSRPYPGDHPVRENVEIIPAAVLLNTFLNAAA.ASG
ChlB1 TGQGLQHDLDSHALLGPRVPVAG.RPLELWRTLLDDETRPYPGSHTINGTEIVPAAVLINTFLDAARAAD
TH
950 960 970 980 990 1000
NcsB ..PSALTDIVLRTPLAVEPPRVVQVVREGRSLSLATRVAEDADADGS...EWITHTTAAVTPGVRP.AGG
Azi26 ERACGLRDIVFRIPLAAHPTRVVQVVLEQDKVRIASRIKRDQESGGVRDDEWLTHTTATVVHEPEVGARP
SACE5532 Q.PAALSDVELRIPIAVEVPRTLQVVLQEGNLRLASRIDGDSSDEHV....WLTHTTAVADHHSRLSTEY
MdpB AGWSDLADVALRVPVSLTRPRDVQVVLQDGTLRLSSRMADGDPDDKG....WVTHTTAAVEPYSDPAPDG
PokM1 GTLPDLTDVALRVPVSVSQPREVQVVSQDATLRLTSRIVGEPGDDRG....LVTHTTAGVEPHGALG.EA
ChlB1 GARPVLRDMALRLPLITTERREELQVVRDDNSLRLASRSLEDG...AA....WLTHTTATAAPAGSGE.AL
TH
1010 1020 1030 1040 1050 1060 1070
NcsB RLDTEAIRSRLPE.GSLTRADEMFERMGVEGYAFPWDLEELRHDDHEQLAVLQIEPSP...AQRATSWAH
Azi26 MEDPDVIRARCPVSWTWAKVDGIFRTMGVDGYTFPWVVEELLRGEDEQFSTITVDHTP...KLHPSSWTA
SACE5532 LEEAVIRRVSSREEWTWDEFDQWFRARGVDGYGFPWKVETLHTGNGELLASLHCD.........GDSWAE
MdpB GPVAAGTPLP......PSHVVDRLGALGVAGMGFPWTVERIRAGEGVLALTVRAGDAAGPGDPAKATWAP
PokM1 SPVVAPDVLP......TGRVVDRLAELGVAAMGFAWEITELRGGEGVLVARVAADPKS...SEPPATWAS
ChlB1 QDLAAGAVLRPAD...PGDVQRHLTSVGVPTMGFEWTIEELARSEGMLAARVSVERPQ....RAQETWAP
1080 1090 1100 1110 1120 1130 1140
NcsB VIDGALTISAMVVSPGDATVLWMSRSIDQVTWSGEPPARLTVHSTRSLRSPHDTVDVRVADERGDVVCEV
Azi26 VVDAALTASGVLVMDENSNVLRTCSHLESLSFVGPPPPRIHVHTVRDPRTP.DTISMTVADESGAVVCEA
SACE5532 ALDAALTITPLLLPDDD..LTRMPAHIRKFAVSGNPPSHLLVHARASVQSP.DTIDVLIADESGRVVAET
MdpB ILDAAMSAGSVALPDPP..TLRMPARLHRLSLPPAP...PVRADLVIRVTGENTVDVDILAPEGGTLGTL
PokM1 VLDAALSTASVVFPGPP..VLRMPAHIKRVALEGAC...PATARITVRLAGEDVVDVEIGPQGTLVGRL
ChlB1 LLDAALSIAPTAIPGPP..ALRMVASFEEIVTEGAPPAGPATIQVAADPVHENTVDVRIADTDGQAVAWV
KR
1150 1160 1170 1180 1190 1200 1210
NcsB VGLRFAAVEHIG.AAVLPSELVHEIVWRPWDPEDHEGAEDAPVEQVILVGDPE.ATVPLAEQLESAGMTC
Azi26 RGLRYVKVQDIGSGAVGPRDLVHELAWEPVEVP....ADAPVPSQALVVGGAA.GGPALVEALTARGVRA
SACE5532 TGLRFGVLDGAPGTMAAPRDLAYEIIWKPLASEP...DIGSLPEWAVIIGDDDPVTTGFAELLEEQGILC
MdpB TGLRYGELDGEVGAVANPRRLVHEITWRAPDGAA...GRAAEPSGVVLVGPRSALRDALAEGGLARTGTPH
PokM1 DGLRYGLLEGDVGTTTGPRRLVHRMRWRPMEQQD...PGAHR..TVVLVGDDRSLVERLSRRLVAEAVPH
ChlB1 RGLRYDGMDQGGMTAAHPRDLVFEMAWRPFEAPA...PQDVSARRIVLIAAHD..VKPLRTALTRAGAHV

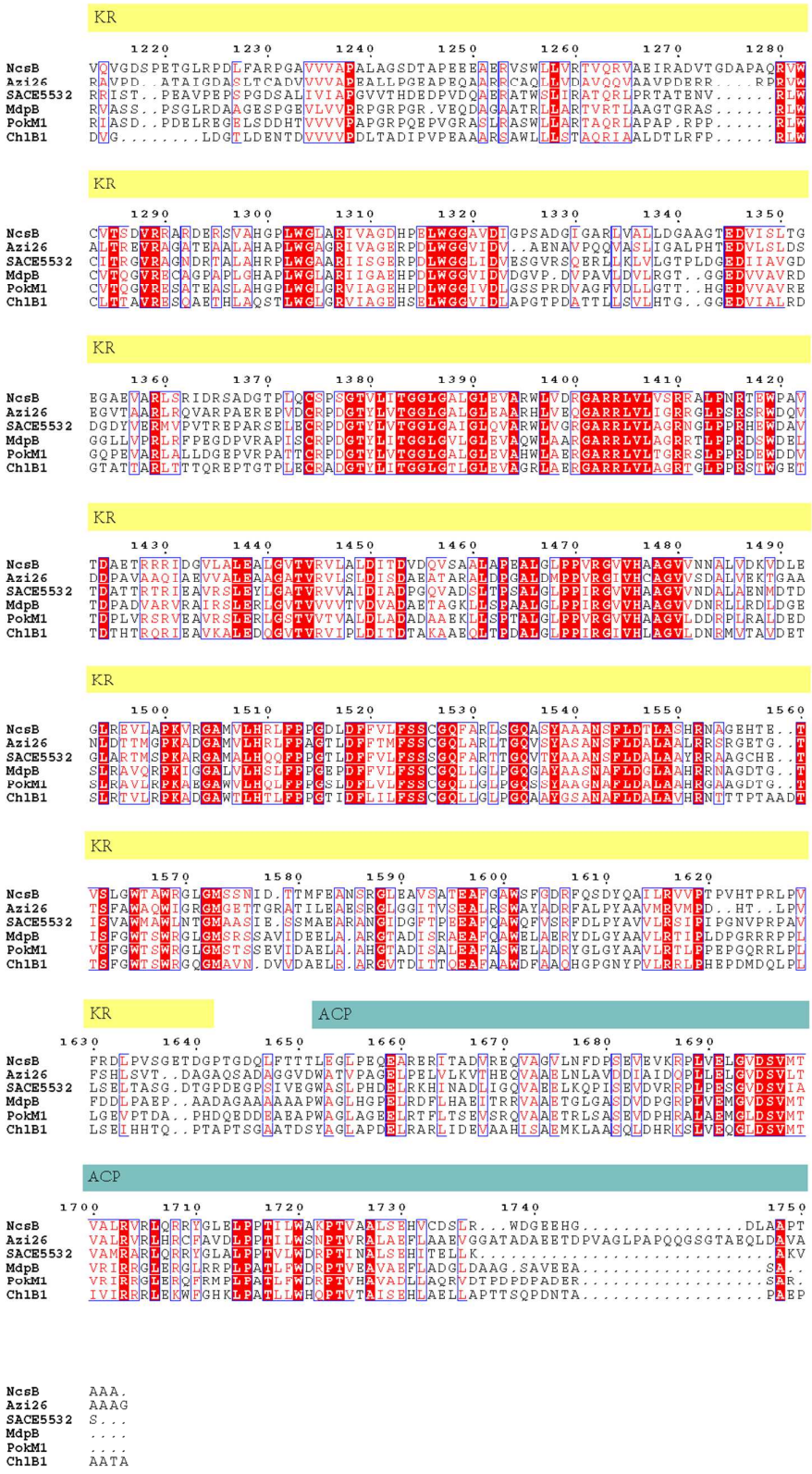


Figure S1. Alignment of the sequences of the mellein synthase and the homologous PKSs. Domains are indicated by the bars of different color.

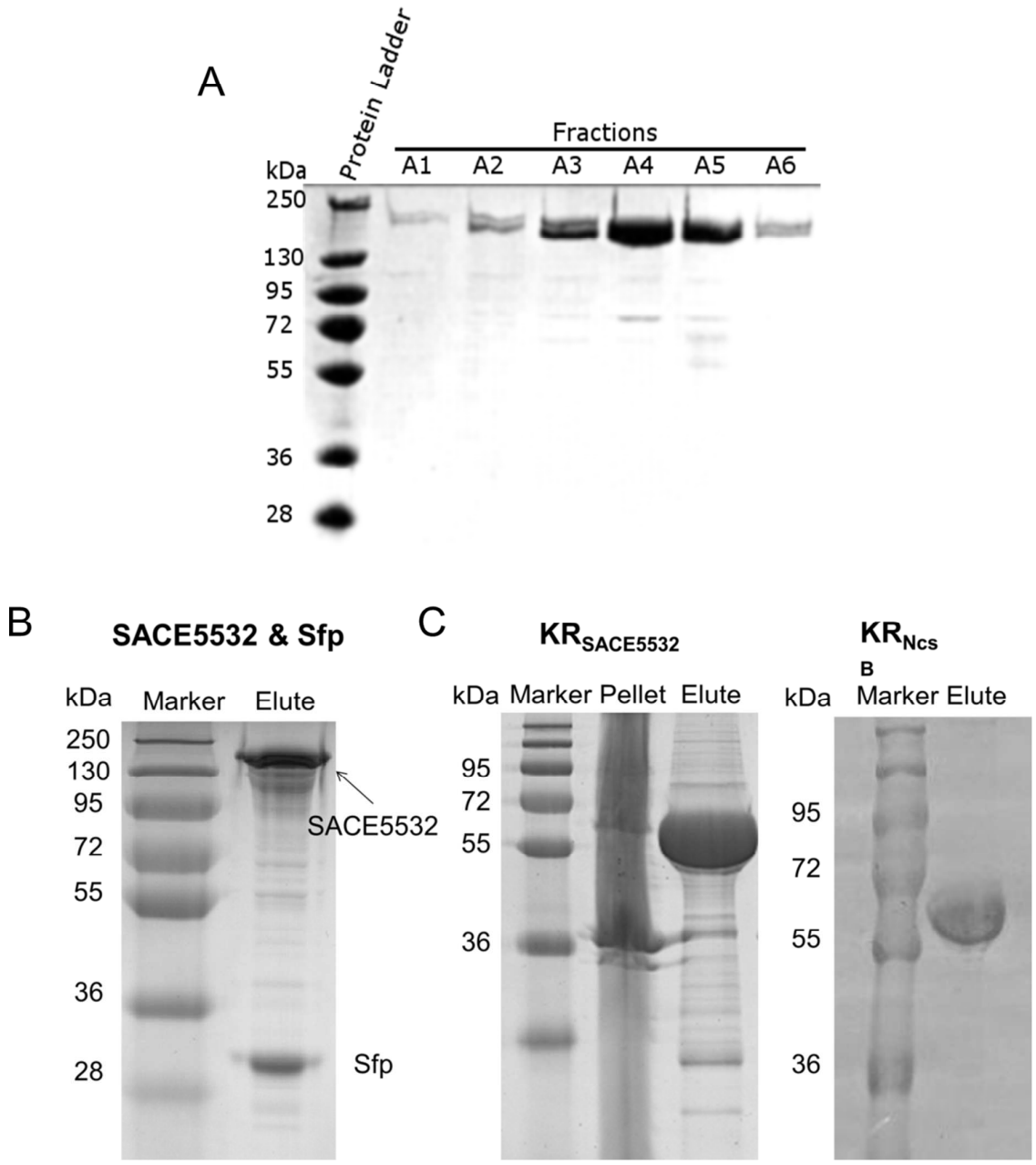


Figure S2. SDS-PAGE gels and size-exclusion chromatogram of the recombinant proteins. A. SDS-PAGE gel of the single-expressed SACE5532 after metal affinity and size-exclusion chromatography. B. SDS-PAGE gel of the co-expressed SACE5532 and Sfp eluted from Ni^{2+}-NTA column. C. SDS-PAGE gel of the stand-alone KR domains.

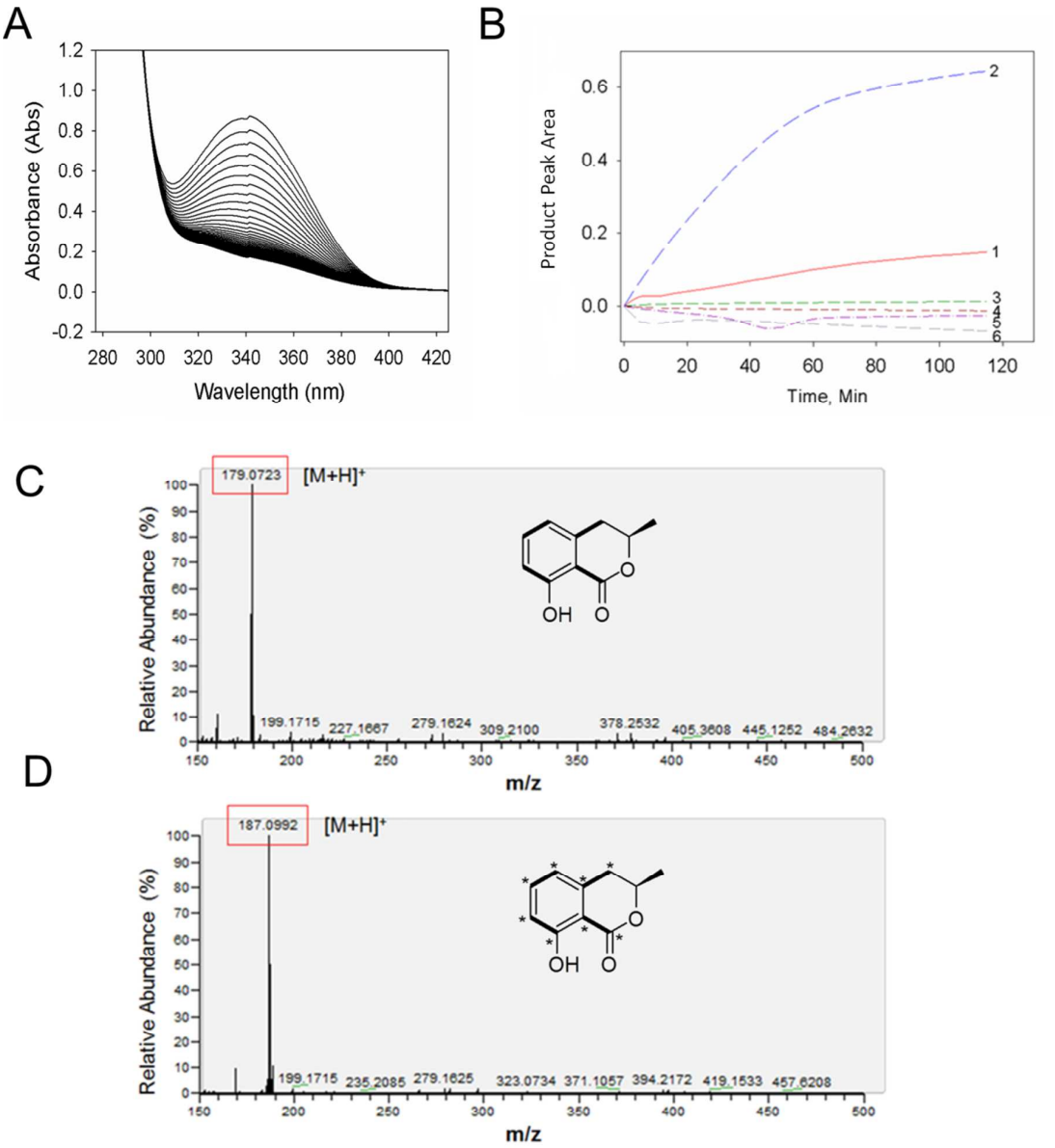


Figure S3. *In vitro* enzymatic assay and product analysis by mass spectrometry. A. Enzymatic reaction of SACE5532 as followed by the time-dependent decline of the NADPH absorbance (substrates include acetyl-CoA, malonyl-CoA and NADPH). B. Enzymatic reactions of SACE5532 as followed by the formation of the enzymatic product. (1. SACE5532 (expressed by using pET26 vector); 2. SACE5532 (expressed by using pET28 vector); 3. Cerulenin inhibited reaction 4. No Sfp modification; 5. No malonyl-CoA 6. No-acetyl-CoA). C. Mass spectrometry of the mellein produced by unlabeled acetyl-CoA and malonyl-CoA. D. Mass spectrometry of the mellein produced by unlabeled acetyl-CoA and [1, 2, 3]-^{13}C-labelled malonyl-CoA.

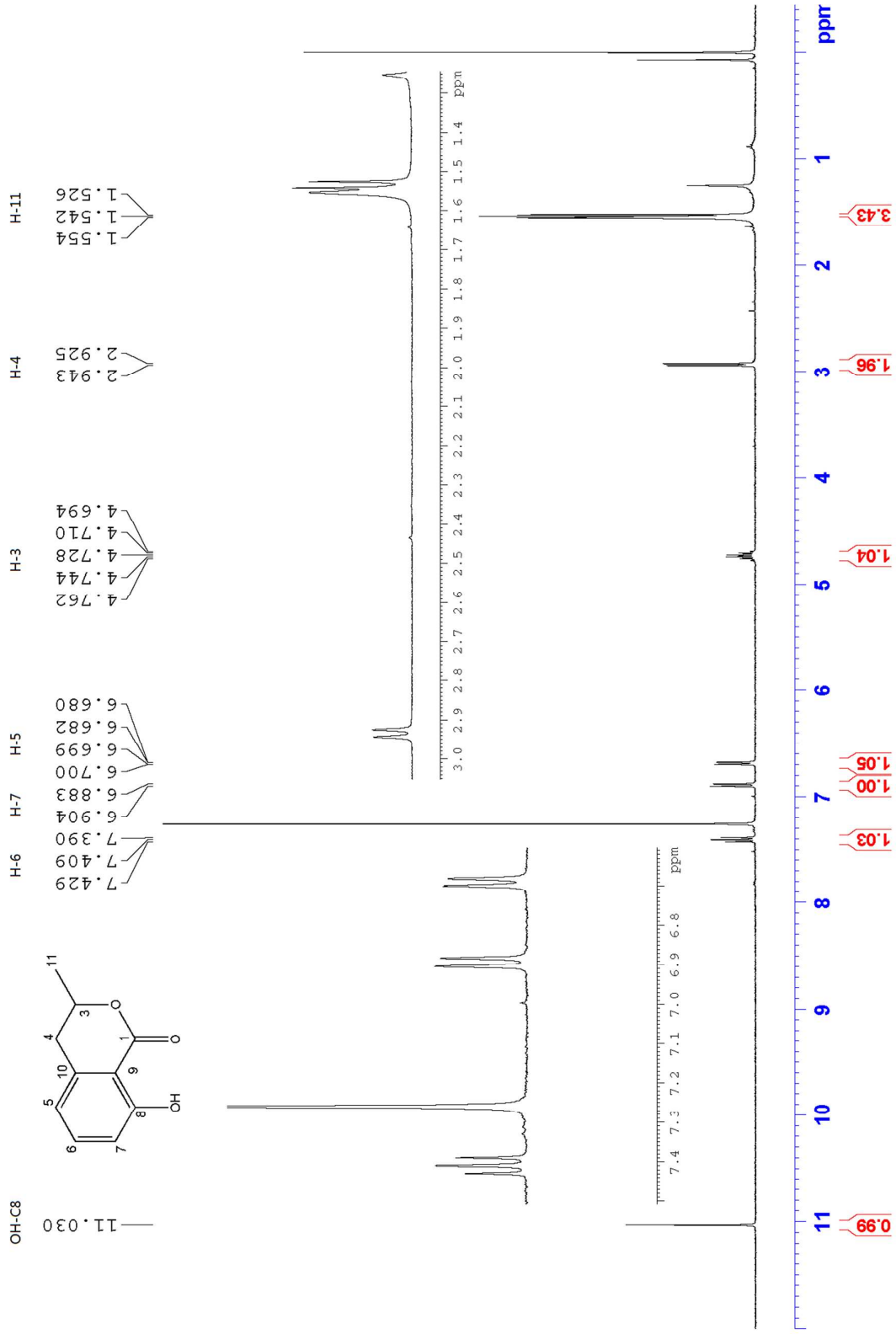


Figure S4. ^{1}H-NMR spectra of the enzymatic product mellein. ($CDCl_3$, 400 MHz).

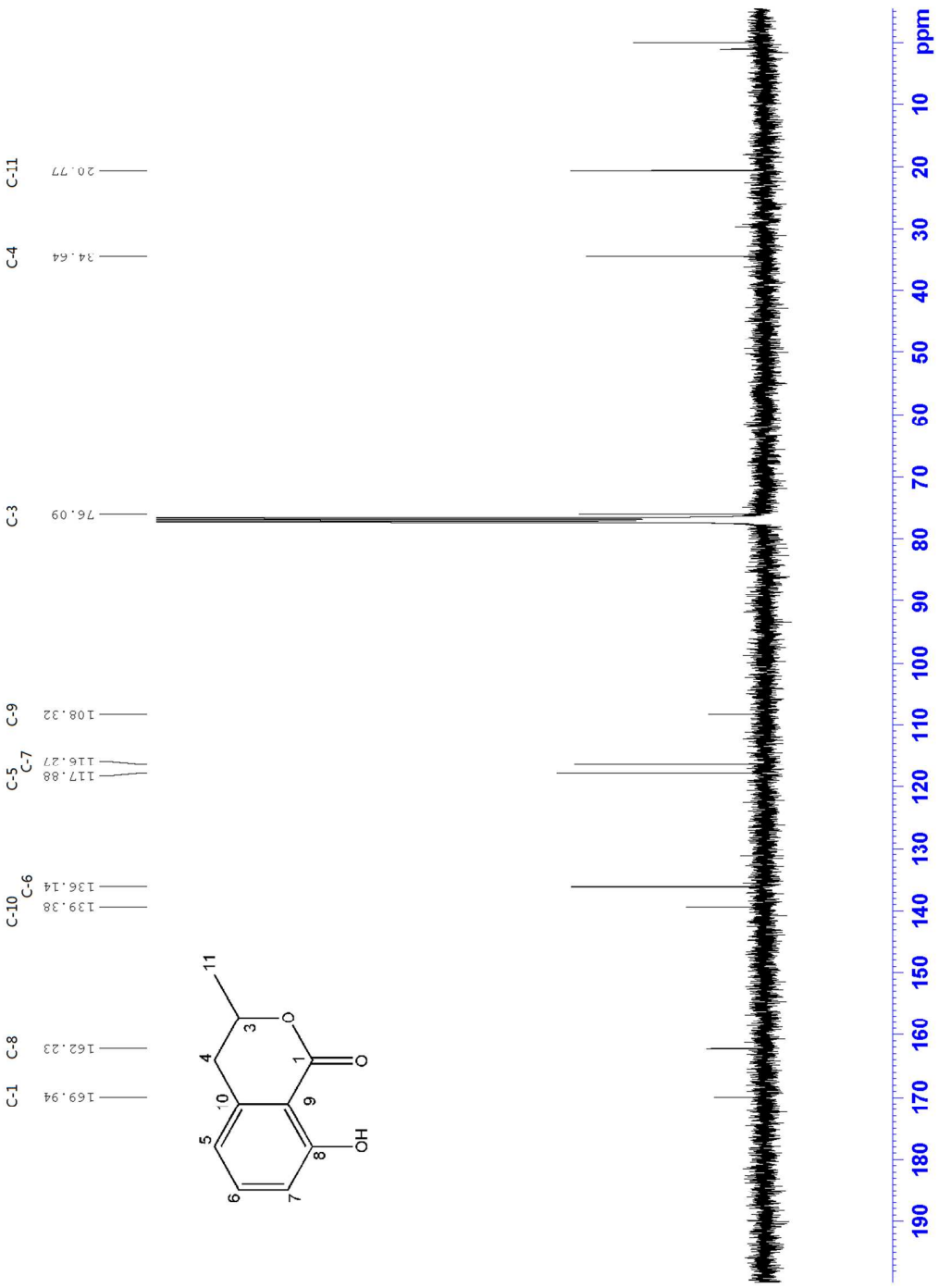


Figure S5. ^{13}C NMR spectrum of the enzymatic product mellein. ($CDCl_3$, 400 MHz).

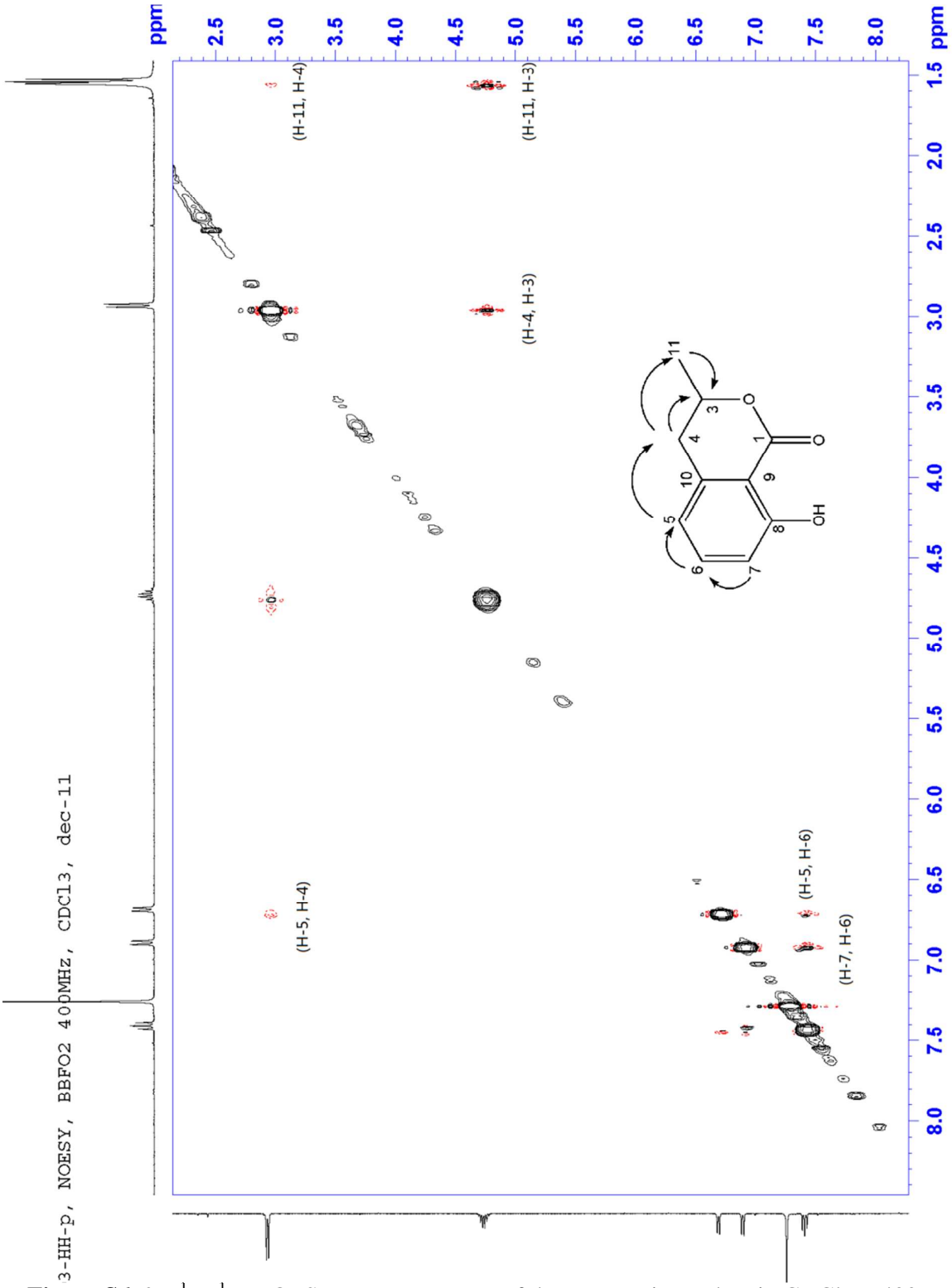


Figure S6. 2D ^{1}H, ^{1}H-NOESY NMR spectrum of the enzymatic product in $CDCl_3$ at 400 MHz.

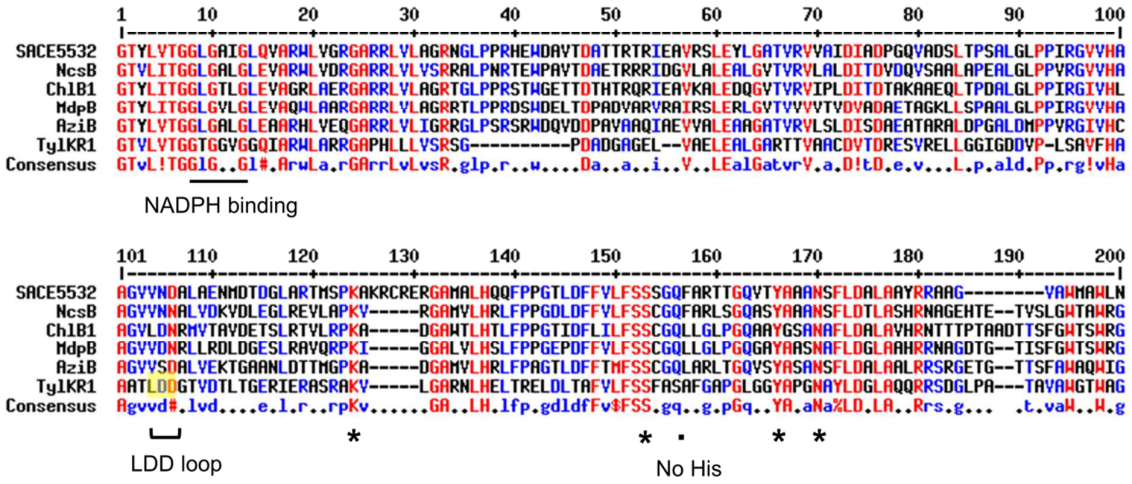


Figure. S7. Multiple sequences alignment of the catalytic KR subdomains of iterative PKSs. The PKSs included SACE5532, NcsB, ChlB1, MdpB and AziB, with TrylKR1 (PDB ID: 2FR0) from the modular type I PKS tylosin PKS. The residues with high consensus value (90%) are in red, and the residues with low consensus value (50%) are in blue. The catalytic residues K, S, Y and N are labeled with asterisks. The conserved motif (GXGXXG) for NADPH binding is underlined. The KR domains of the iterative type I PKSs lack the His residue at the active site groove (highlighted with dot) and characteristic L-D-D motif (highlighted with yellow shadow), implicating that they fall into an A1 type IR family and act on the keto groups to yield L-hydroxy configurations exclusively (Ref. Keatinge-Clay, Structure, 2007).

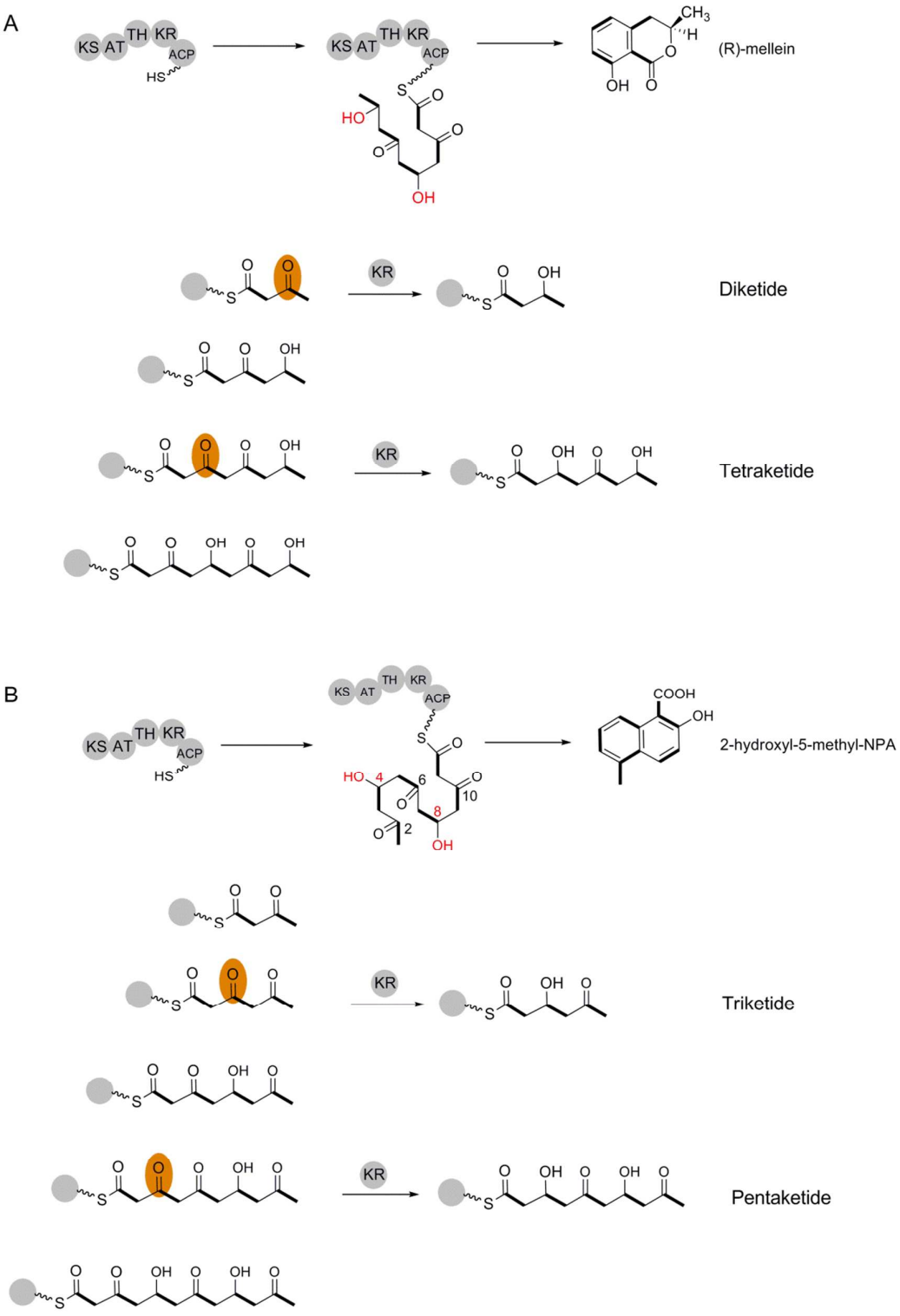


Figure. S8. Putative polyketide intermediates and keto-reduction pattern in mellein and 2-hydroxyl-5-methyl-NPA biosynthesis.

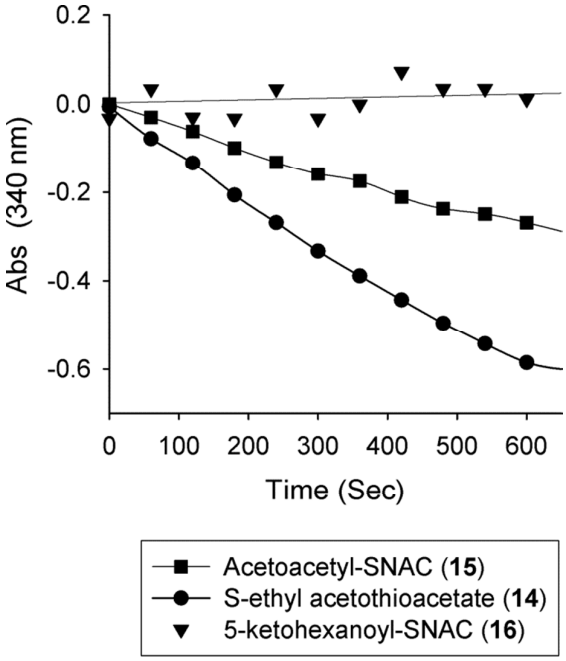


Figure S9. Ketoreducase activity of the stand-alone $KR_{SACE5532}$ domain towards di-,and triketide analogs. No reduction was observed for KR_{NcsB} for the three substrates (data not shown).

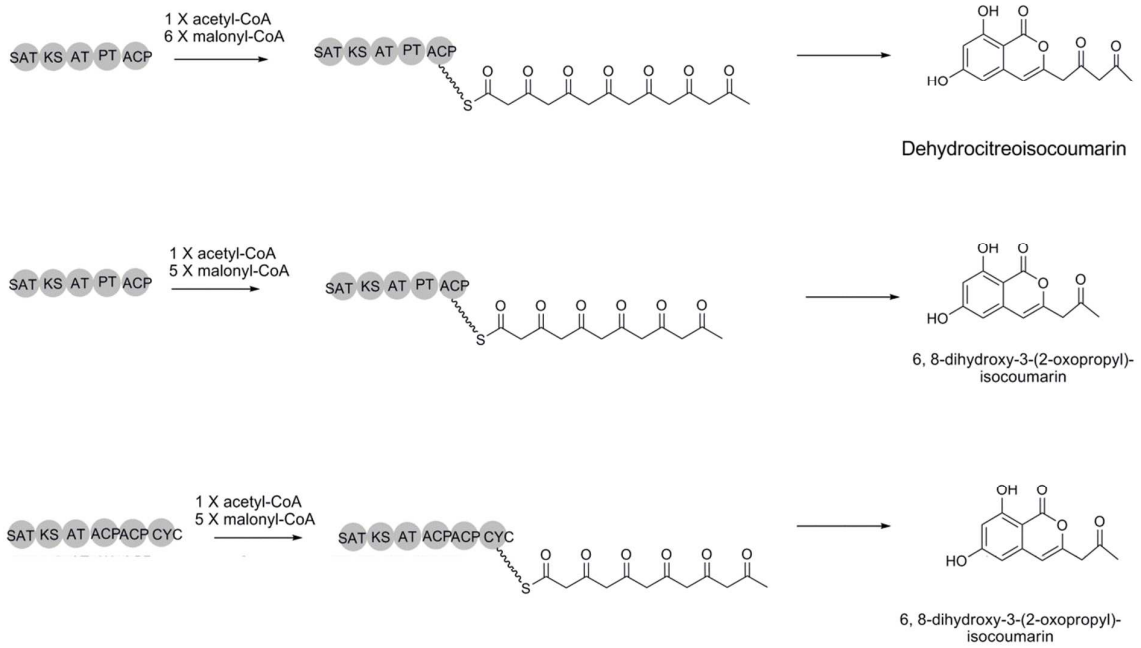


Figure S10. Synthesis of isocoumarins by non-reducing (NR) PKSs (Refs: Ahuja et al, J. Am. Chem. Soc. 2012, J. Am. Chem. Soc., Article ASAP, DOI: 10.1021/ja3016395; Nakazawa et al, ChemBioChem. 2012, 13, 855-861; Ishiuchi et al, ChemBioChem, 2012, 13, 846-854)