Supporting Information

Genome Mining in *Streptomyces*. Elucidation of the Role of Baeyer-Villiger Monooxygenases and Non-Heme Iron-Dependent Dehydrogenase/Oxygenases in the Final Steps of the Biosynthesis of Pentalenolactone and Neopentalenolactone[†]

Myung-Ji Seo,[‡¶] *Dongqing Zhu,*[‡¶] *Saori Endo,*[§] *Haruo Ikeda,*[§]* *and David E. Cane*[‡]*

[‡]Department of Chemistry, Box H, Brown University, Providence, RI 02912-9108, USA

[§]Laboratory of Microbial Engineering, Kitasato Institute for Life Sciences, Kitasato University, 1-15-1 Kitasato, Sagamihara, Minami-ku, Kanagawa 252-0373, Japan

[¶]These authors contributed equally to this work.

David_Cane@brown.edu, ikeda@ls.kitasato-u.ac.jp

Table S1. Strains, plasmids and cosmids used in this study.

Strain or plasmid	Relevant phenotype and/or characteristics	Source
S. exfoliatus strains		
UC5319	Wild-type, pentalenolactone producer	Upjohn Co (Pfizer) (*1*)
ZD20	*penD* in-frame deletion mutant	This work
S. arenae strains		
TÜ469	Wild-type, pentalenolactone producer	DSM 40734, DMSZ Braunschweig, DE (*2*)
ZD17	*pntD* mutant, $Apra^{+}$	This work
ZD18	*pntD* in-frame deletion mutant	This work
S. avermitilis strain		
SUKA16::pKU462-*ermE*p-*ptl_cluster-ΔptlE*	In-frame *ptlE* mutant	(*3*)
SUKA16::pKU462-*ermE*p-*ptl_cluster-ΔptlED*	In-frame *ptlED* mutant	(*3*)
SUKA16::pKU462-*ermE*p-*ptl_cluster-ΔptlED ermEp-penE*	Complementation with *penE*	This work
SUKA16::pKU462-*ermE*p-*ptl_cluster-ΔptlED ermEp-pntE*	Complementation with *pntE*	This work
SUKA16::pKU462-*ermE*p-*ptl_cluster-ΔptlED ermEp-penEpenD*	Complementation with *penE* and *penD*	This work
E. coli strains		
DH10B	*recA*	Gibco BRL
BL21(DE3)	*dcm*, *ompT*, *hsdS* ($r_B^- m_B^-$), *gal*	Invitrogen
ET12567/pUZ8002	*dam, dcm, hsdS, cat, tet;* pUZ8002*: tra, neo*, RP4	(*4, 5*)
BW25117/pIJ790	K12 derivative: *ΔaraBAD*, *ΔrhaBAD*; pIJ790: λ-RED (*gam, bet, exo*), *cat, araC, rep101*^{ts}	(*6, 7*)

Plasmids	**Relevant phenotype and/or characteristics**	**Source**
pET-26b	Kan^{+}, *ori*^{f1}, *lacI*, *ori*^{pBR322}, T7 promoter	Novagen
pET-28a	Kan^{+}, *ori*^{f1}, *lacI*, *ori*^{pBR322}, T7 promoter	Novagen
pBluescriptII SK(+)	*bla*, *ori*^{f1}, *lacZ*, P*lac*, *ori*^{pUC}	(*8, 9*)
pHZ1357	pIJ101 derivative, *bla*, *tsr*, *cos*, *oriT, sti*	(*10*)
pSET152	*aac(3)IV, lacZ, rep*^{pUC}, *oriT*, *att*$^{\phi C31}$	(*11*)
pJTU1278	*pIJ101 derivative, bla, tsr, oriT, sti*	(*12*)
pIJ773	*aac(3)IV, rep*^{pUC}, *oriT*	(*6*)
pDQ7	ca. 6.2 kb BamHI DNA fragment harboring partial *pntM*, *pntH, pntG*, *pntF*, *pntE* and partial *pntD* from cosmid 21F7 inserted into BamHI site of pSET152	This work
pDQ9	ca. 2.2 kb BamHI DNA fragment harboring partial *orf -1* from cosmid 21F7 inserted into BamHI site of pSET152	This work
pDQ19	ca. 2.2 kb BamHI DNA fragment harboring partial *orf -1* from cosmid 1E2 inserted into BamHI site of pSET152, same as pDQ9	This work
pDQ20	ca. 3.1 kb BamHI DNA fragment harboring partial *orf -1*, *pntR*, *gapR*, and partial *pntM* from cosmid 1E2 inserted into BamHI site of pSET152	This work

Plasmids	**Relevant phenotype and/or characteristics**	**Source**
pDQ22	ca. 6.2 kb BamHI DNA fragment harboring partial *pntM*, *pntH*, *pntG*, *pntF*, *pntE* and partial *pntD* from cosmid 1E2 inserted into BamHI site of pSET152, same as pDQ7	This work
pDQ23	ca. 6.4 kb BamHI DNA fragment harboring partial *pntD*, *pntB*, *pntA*, *pntI*, *orf 1* and partial *orf 2* (*orf 2*, not shown) from cosmid 1E2 inserted into BamHI site of pSET152	This work
pDQ31	ca. 6.8 kb KpnI DNA fragment harboring partial *pntI*, *orf 1*, *orf 2*, *orf 3*, *orf 4*, and partial *orf 5* (*orf 2 – orf 5*, not shown) from cosmid 7D7 inserted into KpnI site of pBluescriptII SK(+)	This work
pDQ34	ca. 6.8 kb KpnI DNA fragment harboring partial *pntI*, *orf 1*, *orf 2*, *orf 3*, *orf 4*, and partial *orf 5* (*orf 2 – orf 5*, not shown) from cosmid 7D7 inserted into KpnI site of pBluescriptII SK(+)	This work
pDQ44	pJTU1278 digested with XbaI and SpeI, and self-ligated	This work
pDQ45	ca. 6412 bp KpnI DNA fragment harboring *pntD* from pDQ34 inserted into the KpnI site of pDQ44	This work
pDQ46	822 bp DNA fragment of *pntD* from 40 nt to 861 nt of pDQ45 replaced by ca. 1369 bp *oriT* and *aac(3)IV* using PCR targeting system	This work
pDQ47	pDQ46 digested with XbaI and self ligated to delete *oriT* and *aac(3)IV*	This work

Plasmids	**Relevant phenotype and/or characteristics**	**Source**
pDQ49	ca. 7 kb BamHI + HindIII DNA fragment harboring *penD* from p56 inserted into the corresponding site of pDQ44	This work
pDQ50	Internal 819 bp of *penD* from 40 nt to 858 nt on pDQ49 replaced by ca. 1369 bp *oriT* and *aac(3)IV* using PCR targeting system	This work
pDQ51	pDQ50 digested with XbaI and self ligated to delete *oriT* and *aac(3)IV*	This work
p56	ca. 6.9 kb PstI DNA fragment harboring partial *penG*, *penF*, *penE*, *penD*, *penB*, *penA* and partial *penI* inserted into PstI site of pBluescriptII SK(+)	This work
pKU464*aac(3)IV::ermEp::penE*	*ermEp-penE*	This work
pKU464*aac(3)IV::ermEp::pntE*	*ermEp-pntE*	This work
pKU464*aac(3)IV::ermEp::penE-penD*	*ermEp-penED*	This work
pKU464*aac(3)IV::ermEp::pntE-pntD*	*ermE-pntED*	This work
pKU464*aac(IV)::ermEp::ptlE/pntE300*	*ermEp-ptlE/pntE300*	This work
pKU464*aac(IV)::ermEp::ptlE/pntE600*	*ermEp-ptlE/pntE600*	This work
pKU464*aac(IV)::ermEp::ptlE/pntE1200*	*ermEp-ptlE/pntE1200*	This work
pJexpress401-PenE_opt_alt2	*ori*^{pUC}, Kan^{+}, *lacI*, synthetic *penE*	DNA2.0, this work
pJexpress401-PntE_opt_alt2	*ori*^{pUC}, Kan^{+}, *lacI*, synthetic *pntE*	DNA2.0, this work
pJexpress401-PenD_opt_alt2	*ori*^{pUC}, Kan^{+}, *lacI*, synthetic *penD*	DNA2.0, this work
pJexpress401-PntD_opt_alt2	*ori*^{pUC}, Kan^{+}, *lacI*, synthetic *pntD*	DNA2.0, this work
pJ201-PtlD	*ori*^{pUC}, Kan^{+}, *lacI*, synthetic *ptlD*	DNA2.0, this work

Plasmids	**Relevant phenotype and/or characteristics**	**Source**
pET-28a-penE	Synthetic *penE* gene from pJexpress401:PenE_opt_alt2 digested with NdeI and XhoI and inserted into the corresponding site of pET28a	This work
pDQ58	Synthetic *pntE* gene from pJexpress401:PntE_opt_alt2 amplified with primer DQ75F/R, digested with NdeI and XhoI, and inserted into the corresponding site of pET-26b	This work
pET-28a-penD	Synthetic *penD* gene from pJexpress401-PenD_opt_alt2 digested with NdeI and XhoI and inserted into the corresponding site of pET-28a	This work
pET-28a-pntD	Synthetic *pntD* gene from pJexpress401-PntD_opt_alt2 digested with NdeI and XhoI and inserted into the corresponding site of pET-28a	This work
pET-28a-ptlD	Synthetic *ptlD* gene from pJ201-PtlD digested with NdeI and XhoI and inserted into the corresponding site of pET-28a	This work
1E2	pHZ1357 derivative, cosmid harboring *pnt* gene cluster	This work
7D7	pHZ1357 derivative, cosmid harboring *pnt* gene cluster	This work
12B2	pHZ1357 derivative, cosmid harboring *pnt* gene cluster	This work
21A5	pHZ1357 derivative, cosmid harboring *pnt* gene cluster	This work
21F7	pHZ1357 derivative, cosmid harboring *pnt* gene cluster	This work
27H6	pHZ1357 derivative, cosmid harboring *pnt* gene cluster	This work
G21	pHZ1357 derivative, cosmid harboring *pen* gene cluster	This work

Plasmids	**Relevant phenotype and/or characteristics**	**Source**
K5	pHZ1357 derivative, cosmid harboring partial *pen* gene cluster	This work
O34	pHZ1357 derivative, cosmid harboring partial *pen* gene cluster	This work

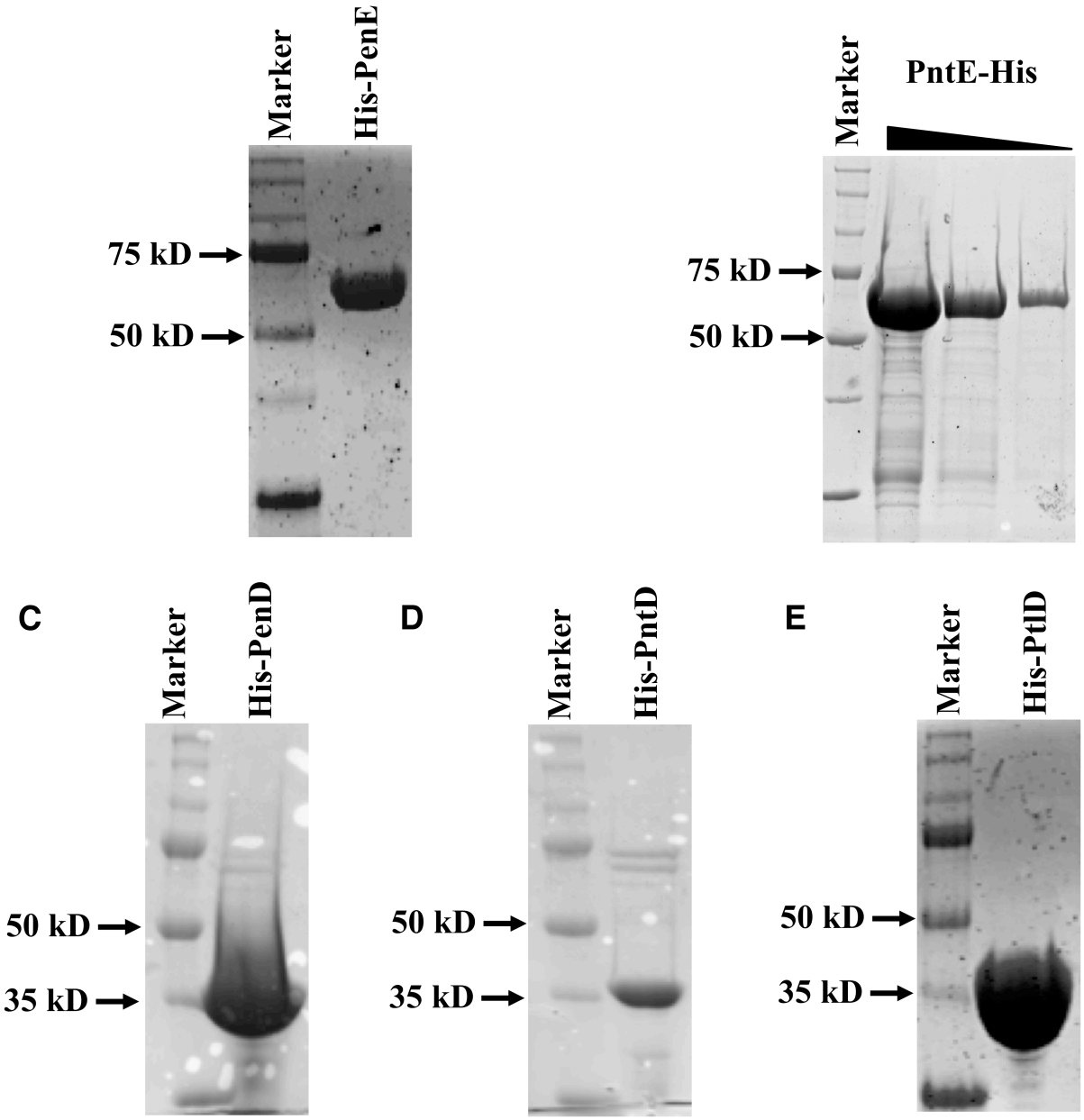

Figure S1. SDS PAGE of purified recombinant proteins. A. His_6-tag-PenE. B. PntE-His_6-tag; C. His_6-tag-PenD. D. His_6-tag-PntD. E. His_6-tag-PtlD.

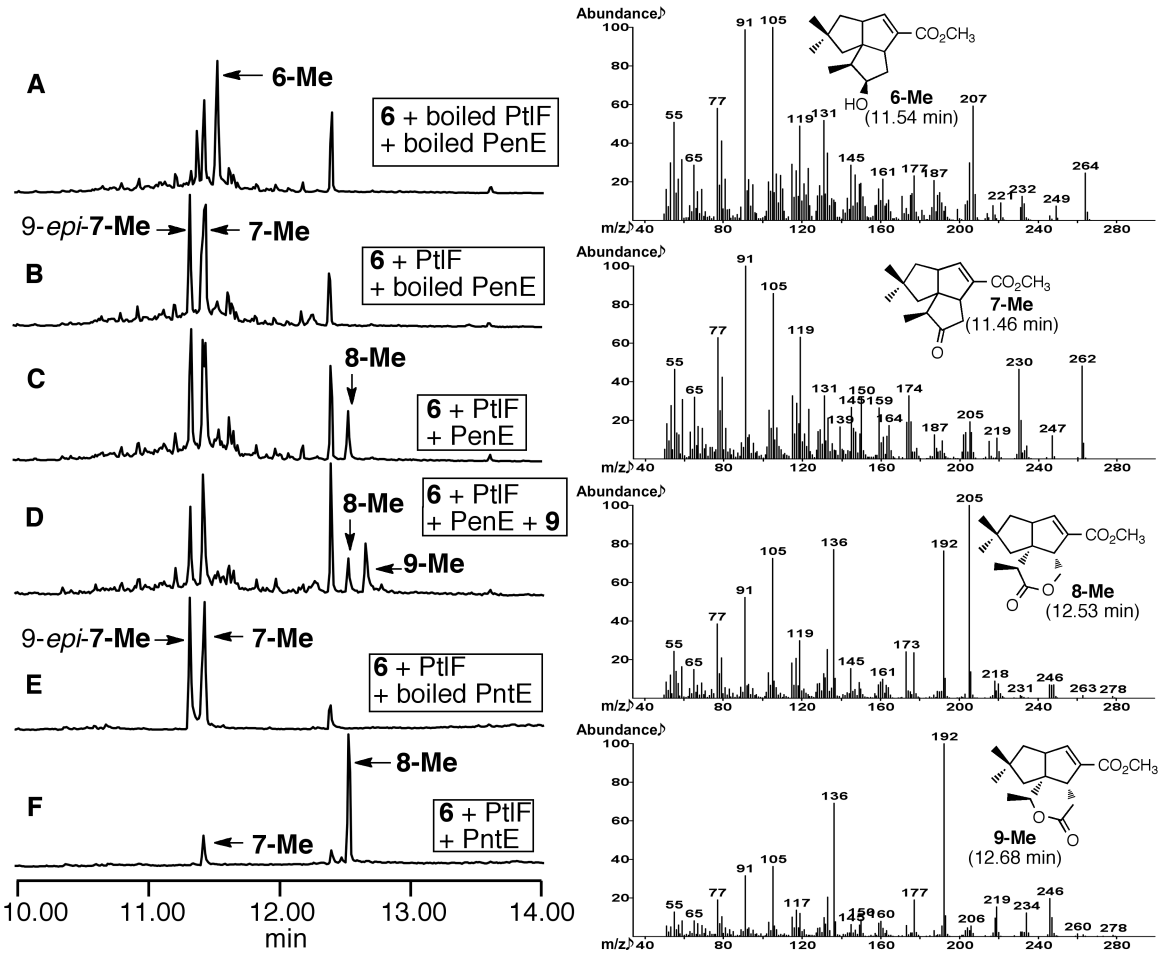

Figure S2. GC-MS analysis (Method 1) of incubations of recombinant PenE and PntE with 1-deoxy-11-oxopentalenic acid (**7**). A. Control incubation of **6** with boiled PtlF and boiled PenE. B. Control incubation of **6** with PtlF and boiled PenE. C. Incubation of **6** with PtlF and PenE. D. Incubation of **6** with PtlF and PenE with added reference **9**. E. Control incubation of **6** with PtlF and boiled PntE. F. Incubation of **6** with PtlF and PntE. Mass spectra of the individual components are given in right panel. For mass spectra of authentic **8-Me** and **9-Me**, see Figure S2 of supplemental ref. 3

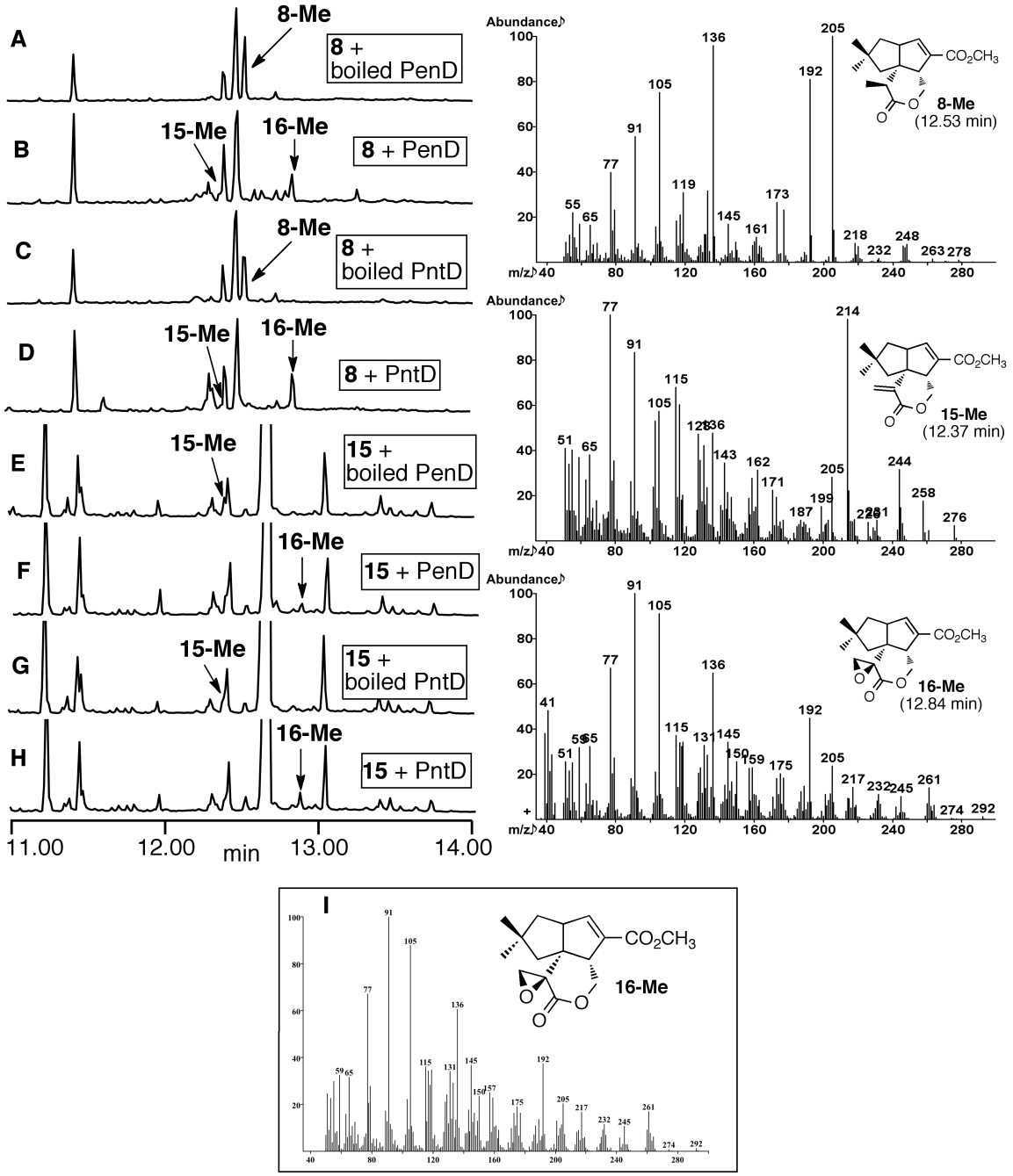

Figure S3. GC-MS analysis (Method 1) of incubations of recombinant PenD and PntD with pentalenolactone D (**8**) and with pentalenolactone E (**15**). A. Control incubation of **8** with boiled PenD. B. Incubation of **8** with PenD. C. Control incubation of **8** with boiled PntD. D. Incubation of **8** with PntD. E. Control incubation of **15** with boiled PenD. F. Incubation of **15** with PenD. G. Control incubation of **15** with boiled PntD. H. Incubation of **15** with PntD. I. MS of authentic reference **16-Me**.

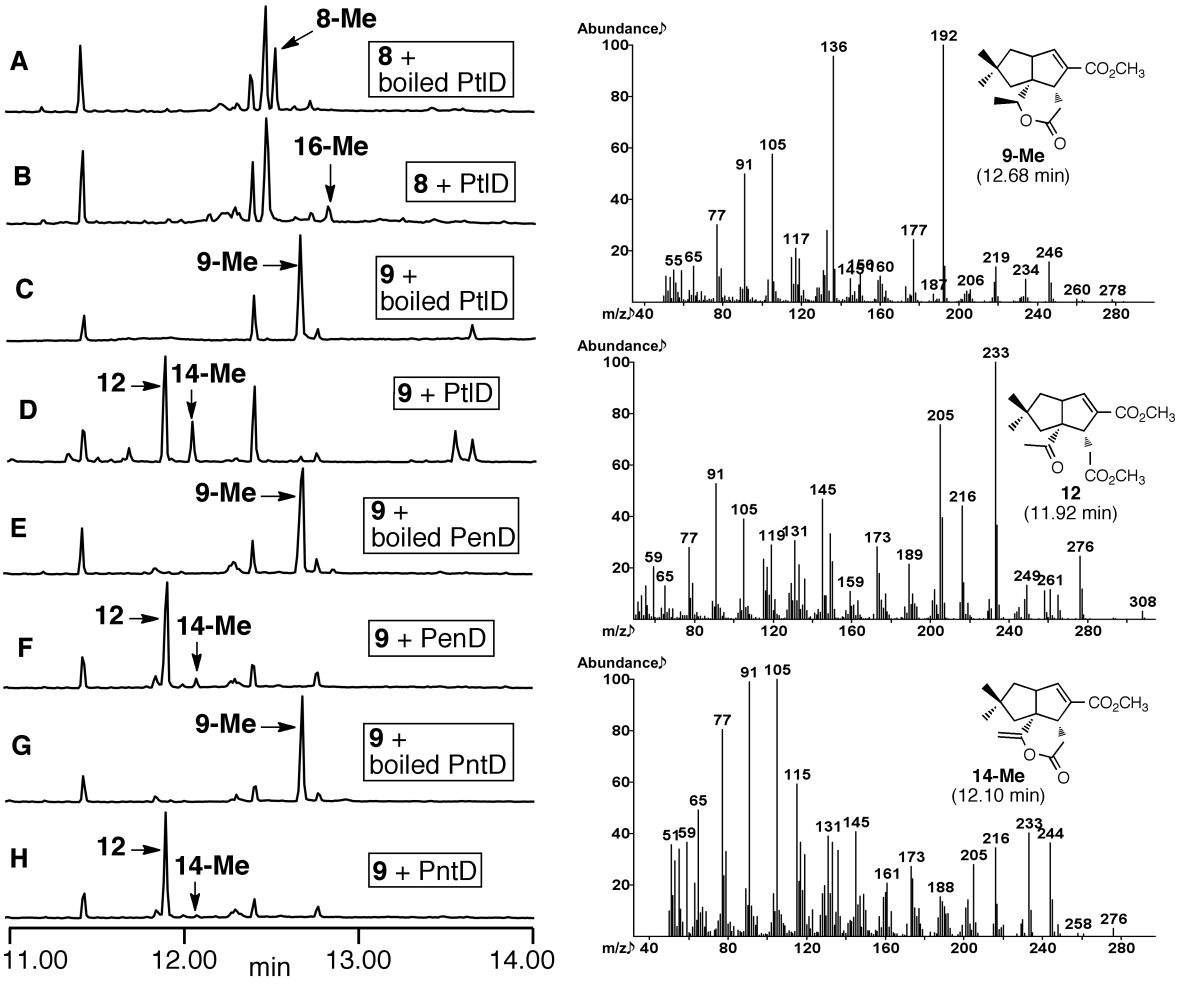

Figure S4. GC-MS analysis (Method 1) of incubations of recombinant PtlD with pentalenolactone D (**8**) and PtlD, PenD, and PntD with with neopentalenolactone D (**9**). A. Control incubation of **8** with boiled PtlD. B. Incubation of **8** with PtlD. C. Control incubation of **9** with boiled PtlD. D. Incubation of **9** with PtlD. E. Control incubation of **9** with boiled PenD. F. Incubation of **9** with PenD. G. Control incubation of **9** with boiled PntD. H. Incubation of **9** with PntD.

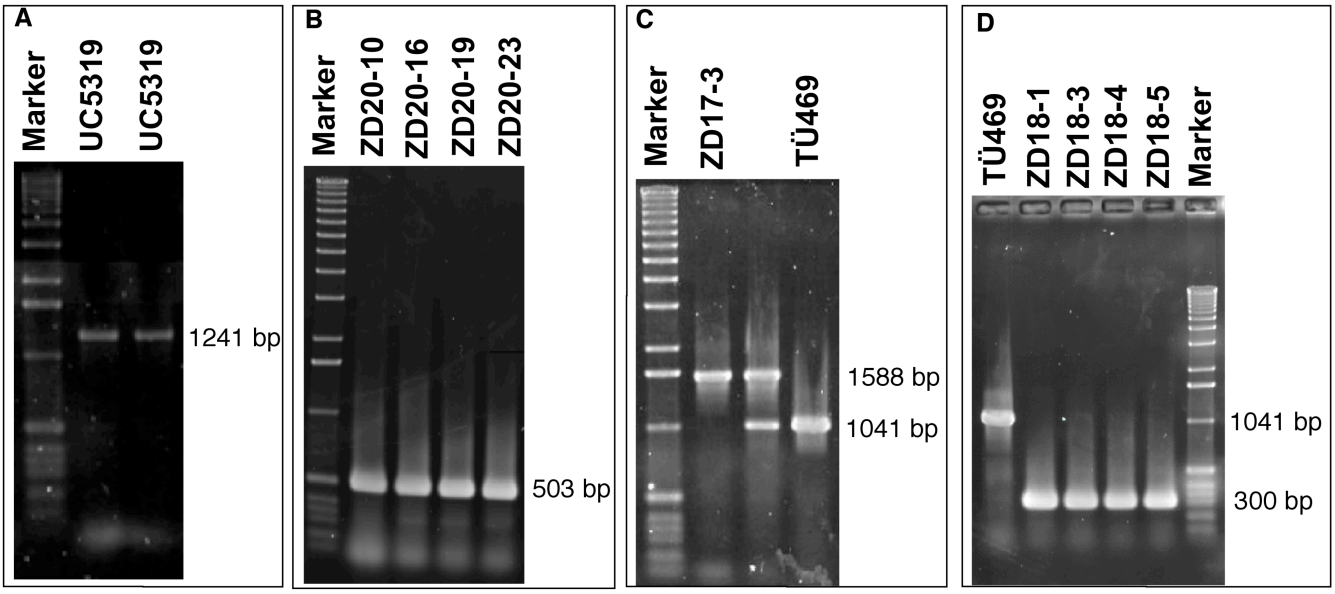

Figure S5. PCR analysis of wild-type and *penD* or *pntD* deletion mutants. A. *S. exfoliatus* UC5319, primer pair DQ81F/R. B. *S. exfoliatus* ZD20 (*ΔpenD*), primer pair DQ81F/R. C. *S. arenae* TÜ469 and ZD-17, primer pair DQ80F/R. D. *S. arenae* TÜ469 and ZD-18 (*ΔpntD*), primer pair DQ80F/R.

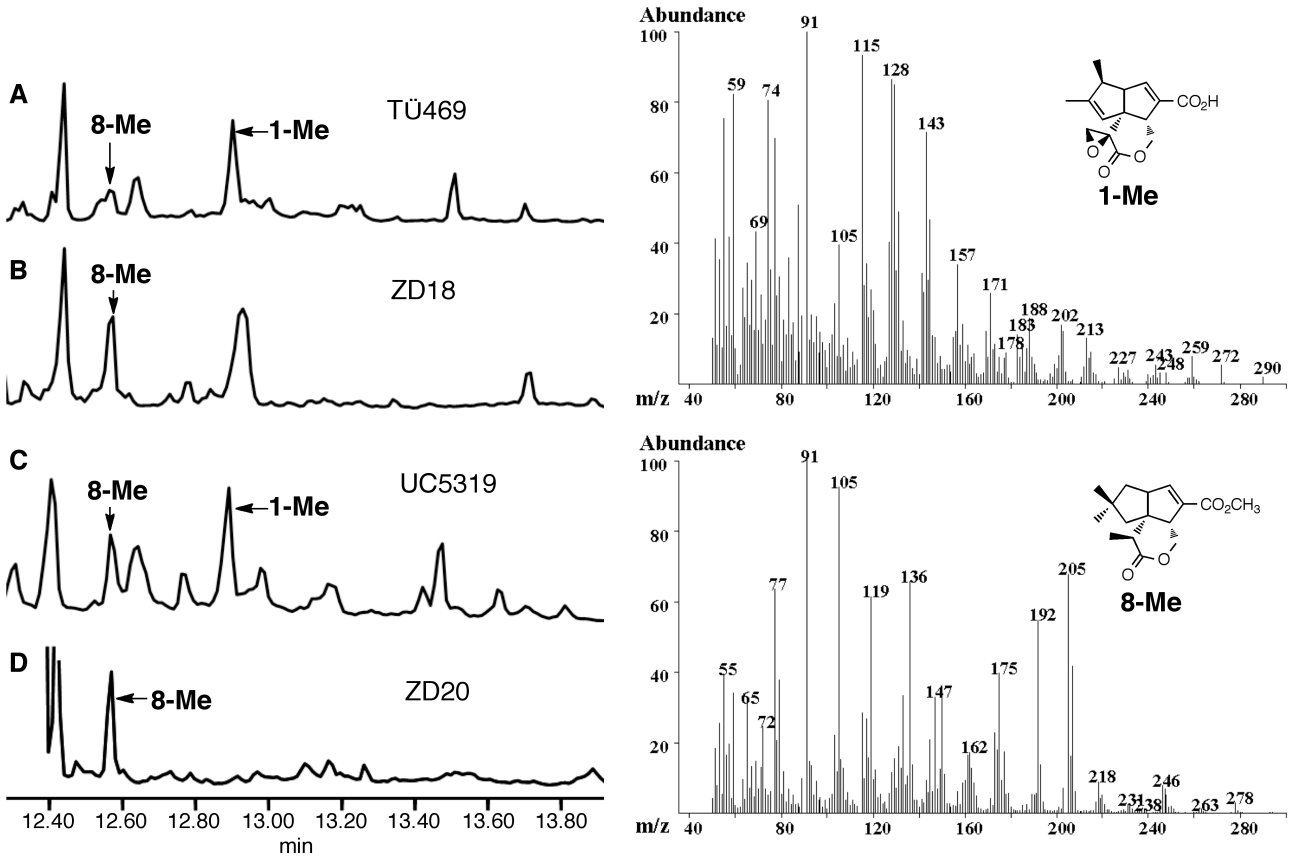

Figure S6. GC-MS analysis (Method 1) of ZD20 and ZD18 deletion mutants. A. Wild-type *S. arenae* TÜ469. B. *S. arenae* ZD18. C. Wild-type *S. exfoliatus* UC5319. D. *S. exfoliatus* ZD20. Under the growth conditions used for each strain, small quantities of the methyl esters of pentalenolactone E (**15-Me**) and pentalenolactone F (**16-Me**) can also be detected in the cultures of both wild-type strains on the shoulders of the peaks for **8-Me** (*m/z* 276, rel ret time **15-Me**, -0.16 min) and **1-Me** (*m/z* 292, rel ret time **16-Me**, -0.04 min), identical by comparison with the spectra of ret time of authentic samples.

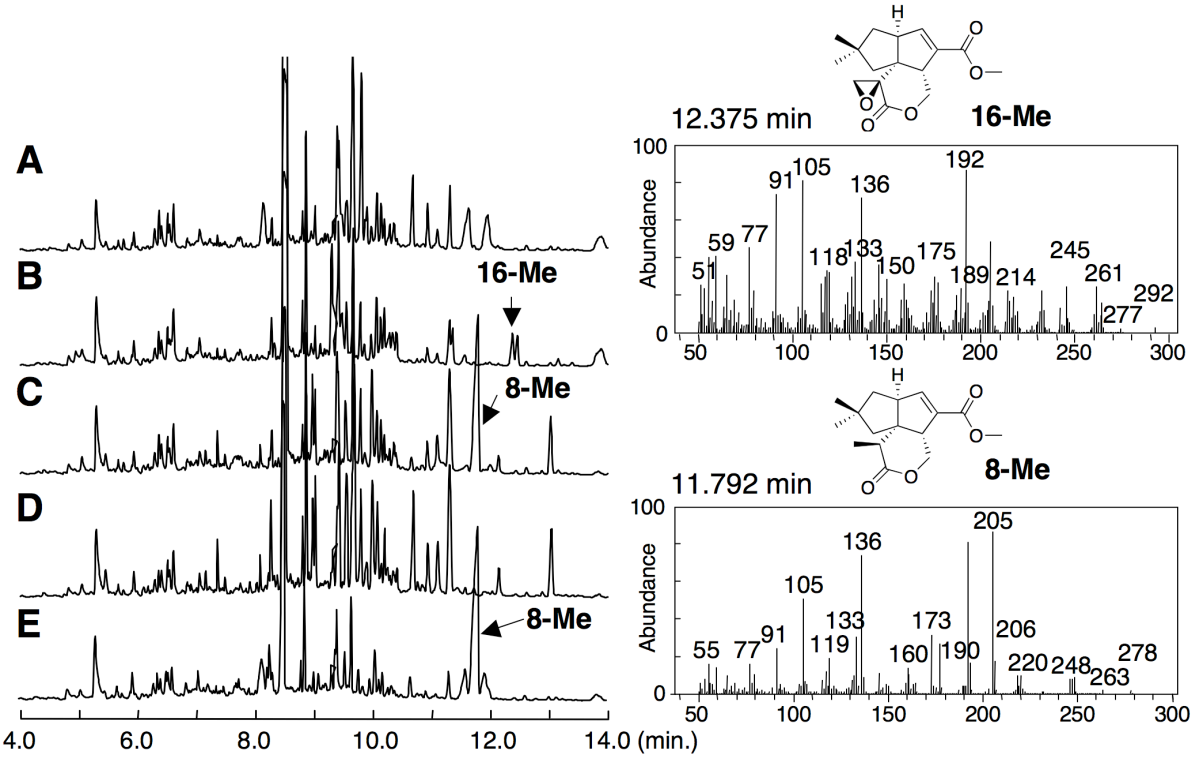

Figure S7. GC-MS analysis of *S. avermitilis* SUKA16 *ΔptlE ΔptlD* double mutant and *ΔptlE* mutant complementation experiments. A. *ΔptlE ΔptlD* double mutant. B. Complementation of *ΔptlE ΔptlD* mutant with *pntE* and *pntD*. C. Complementation of *ΔptlE ΔptlD* mutant with *pntE*. D. *ΔptlE* mutant. E. Complementation of *ΔptlE* mutant with *pntE*.

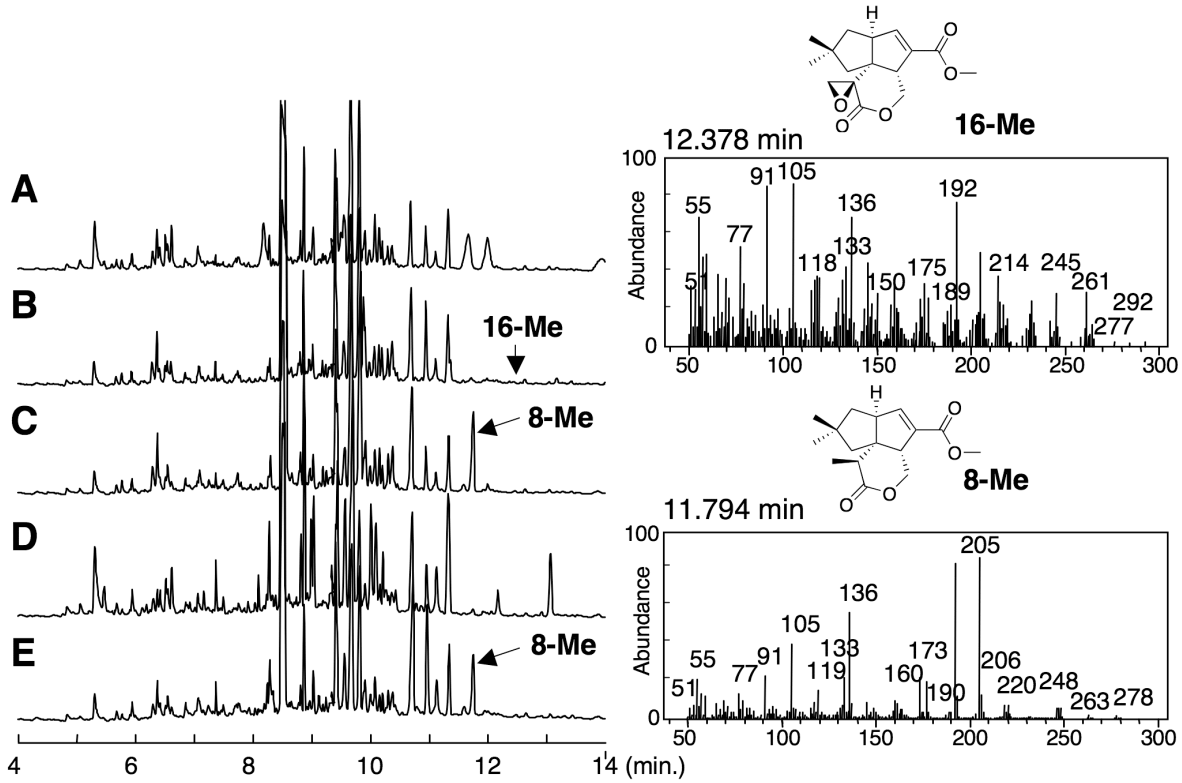

Figure S7-2. GC-MS analysis of *S. avermitilis* SUKA16 *ΔptlE ΔptlD* double mutant and *ΔptlE* mutant complementation experiments. A. *ΔptlE ΔptlD* double mutant. B. Complementation of *ΔptlEΔptlD* mutant with *penE* and *penD*. C. Complementation of *ΔptlE ΔptlD* mutant with *penE*. D. *ΔptlE* mutant. E. Complementation of *ΔptlE* mutant with *penE*.

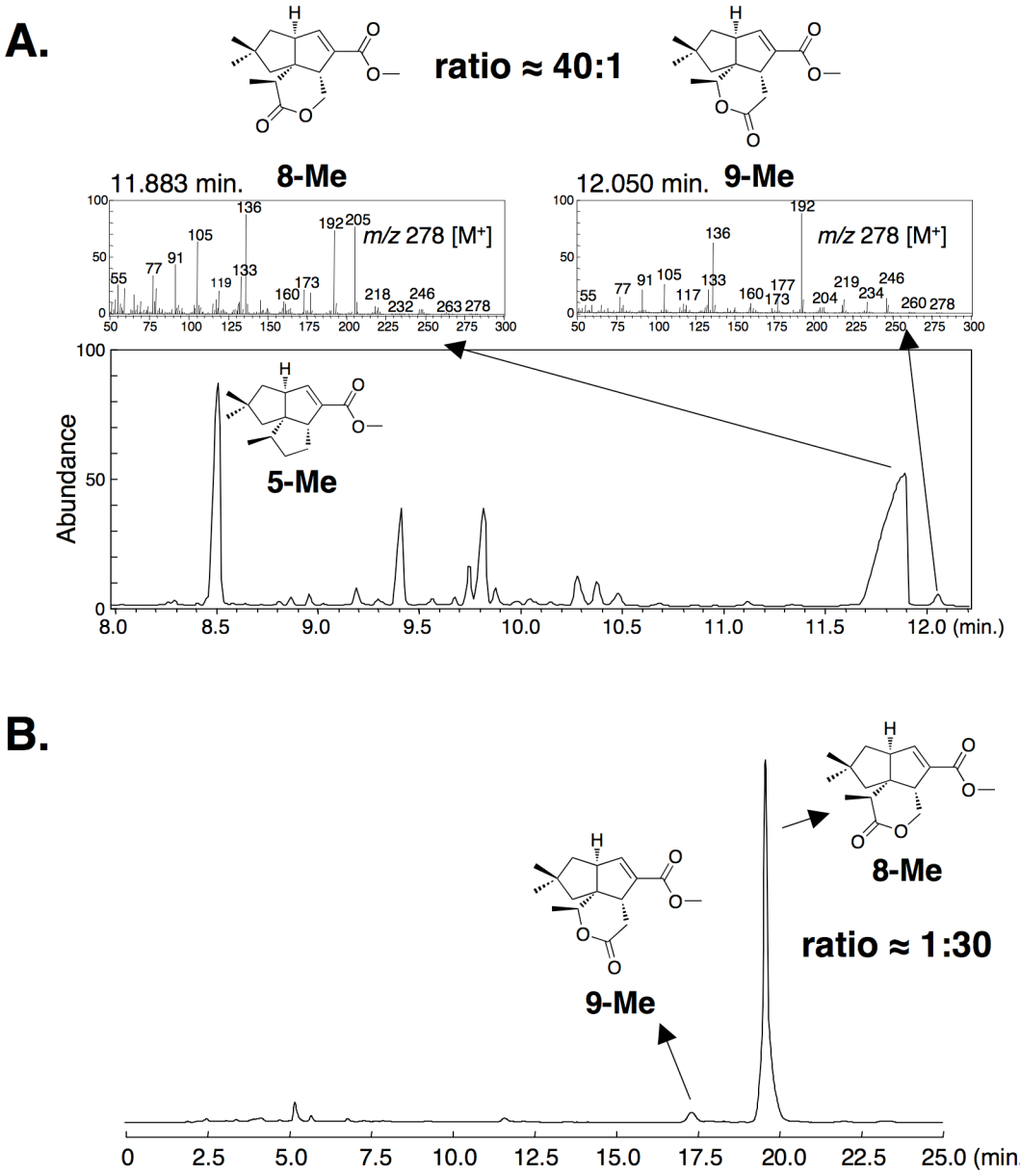

Figure S8. Analysis of methylated products of *S. avermitilis* SUKA16 *ΔptlE ΔptlD* mutant complemented with *pntE*, showing formation of both pentalenolactone D methyl ester (**8-Me**) and 2-3% neopentalenolactone D (**9-Me**). Products were identified by comparison with authentic samples. A. GC-MS analysis (GC-MS Method 2). B. LC-MS analysis, ODS-HPLC.

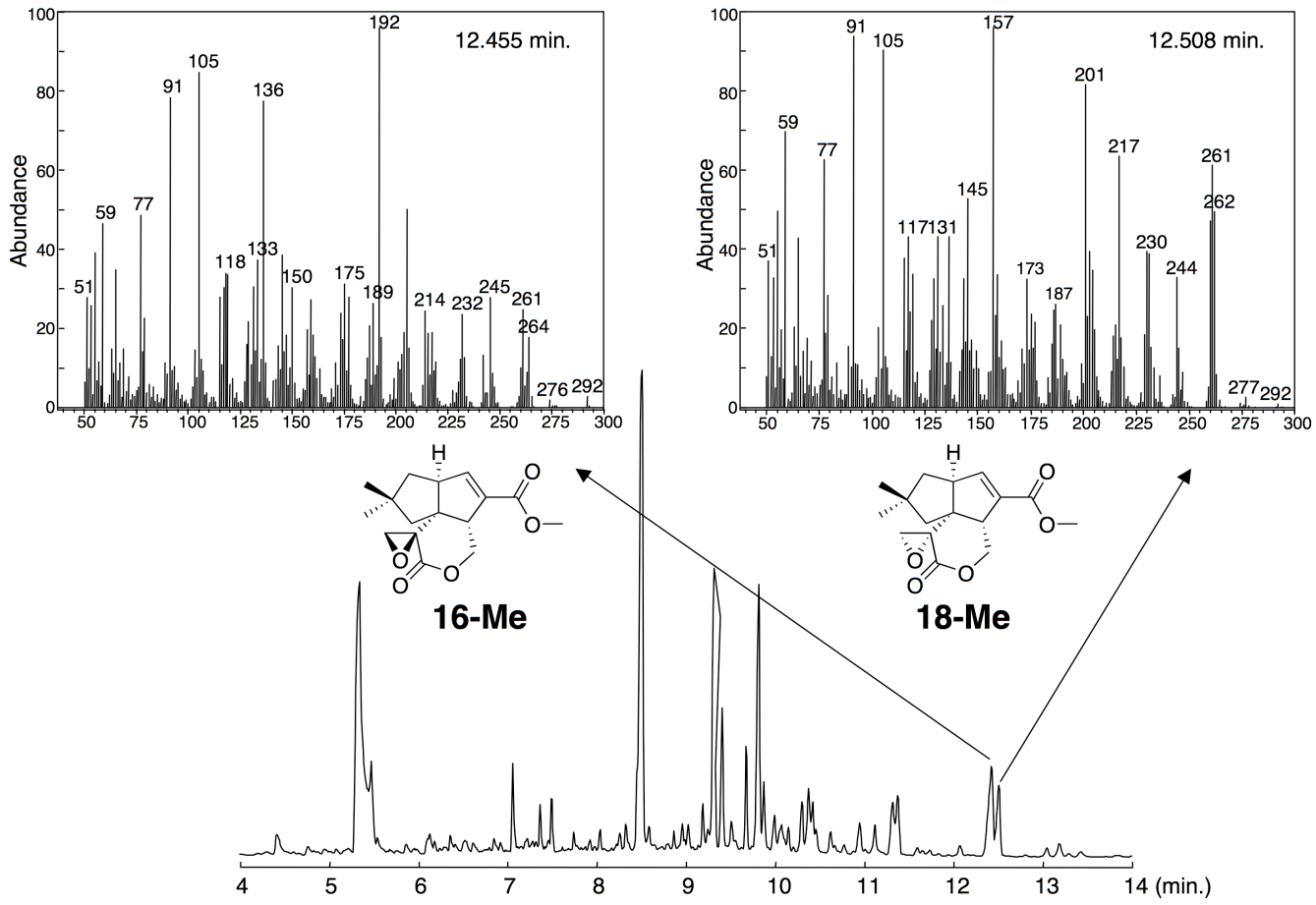

Figure S9. GC-MS analysis of methylated extract of *S. avermitilis* SUKA16 *ΔptlE ΔptlD* carrying pKU464*aac(3)IV::ermEp::pntE-pntD* showing formation of both pentalenolactone F methyl ester (**16-Me**) and 9,10-*epi*-pentalenolactone F methyl ester (**18-Me**). Cf Panel B, Figure S7.

```
PtlE  MVDIEAVRAKYREERDKRVQADGGRQYLSAQGEFAHYADDPHAKPIERAPVSDEVDVTIIGAGIGGLLLGARLREACAFDTIRLVDKAGD
PntE  -MDLEAVREKYRQERDKRGVG---RTYQFARGDFSRYARDPYTERREREPLTDEVDVAVVGAGIGGLLTGARLREETGLERIRLIDEAGD
        ****** ***:*****  .   * *  *:*:*::** **:::  ** *::*****:::******** *****:  .:: ***:* ***

      ptlE-pntE300

PtlE  VGGTWYWNRFPGLRCDVESYVYMPLLEELGRLPSEKYATGAEIFEHCQAIARTYDLYDEALLQTSVTELSWDEDSSRWLVRTDRGDLVRS
PntE  VGGTWYWNRFPGVRCDVESYVYMPLLEEIGTIPTEKYSTGPEIFAHLQRIAHRYGLYRDALFQTTVTELRWDEAAARWLVSTDRGDLFRA
      ************:*************** * :* ***:**.*** * * **: *.** :**:**:**** *** :.**** ******.*:

                                ptlE-pntE600

PtlE  RFVAMAIGSLHRPKLPSIPGTEAFQGHSFHTSRWDFAYTGGDISGGLEKLGDKRVGIVGTGATAVQCIPHLAESAAHLYVFQRTPSTVSV
PntE  RYVAMSIGLMHRPKLPGLPGLETFAGHSFHTSRWDFGYTGGDSTGGLTGLKDKRVGVIGTGSTTVQLAPHLAEWAERLYIFQRTPAAVDV
      *:***:** :******.:** *:* ********** ***** :***  * *****::***:*::*  ***** * :**:******::*.*

PtlE  RNNRPTDPGWAAGLEPGWQQRRMDNFHALTSGVDQDVDLVQDGWTEITSKLAAILPKS-AADADPKDIGTAVELADFHKMEELRKRVDAI
PntE  RGNRPTPPGWADGLDAGWQQRRMENFHALTSGIPQDEDLVQDRWTQTTAELATAILPTGDTGGDPKERALAAERADFRKMEELRARIDSV
      *.**** * ** ** .*******:********: ** ***** **: *::**: :  :  :..***: . *.* *** ****** *:*::

                                      ptlE-pntE1200

PtlE  VHDKDTADALKPYYRLFCKRPCFHDGYLDTYNRPNVTLVDTQGRGVERLTPTSVVAGGREYPVDCLIFASGYESEFGVPYTNRTGFSIVG
PntE  VTDPATAAALKPYYRVYCKRPCFHDGYLQTFNRPNVTLVDTQGQGVERLTASGVVANGREYPVDCLIFATGYEHEFAVPYTERAGYDIVG
      * *  ** *******::***********:*:***********:******.:.***.*****:******:*** **.****:*:*:.***

PtlE  RDGIRLSEKWAEGARTFHGLQVNGFPNCFILSKAQSGLHVNVPYMLNEQSKHVAYILKAVQQRGRQVVEASATGEKEWVETILRLANRNL
PntE  RGGVRLSEKWAQGAHTLHGLQVHGFPNCFILSKVQAGRHVNIAYMLGEQTRHLAHIVKCVEERGHRVVEASEAGEKEWVEEILRLASGDL
      *.*:*******:**:*:*****:**********.*:* ***:.***.**::*:*:*:*.*::**::***** :******* *****. ::

PtlE  DFTESCTPGLFNNEGNPRNVAILNSSYGGGSVGFVNILKRWREADDLADLELREG
PntE  DFLENCTPGLYNNEGDPGGLPLLNSSYGGGSVEFVNILRRWREAGDLAGLELR--
      ** *.*****:****:* .:.:********** *****:*****.***.****
```

Figure S10. Alignment of amino acid sequences of Baeyer-Villiger monooxygenases, PtlE and PntE. Underlined characters indicate characteristic motifs. Shadowed boxes correspond to the junction regions for construction of PtlE/PntE hybrid proteins.

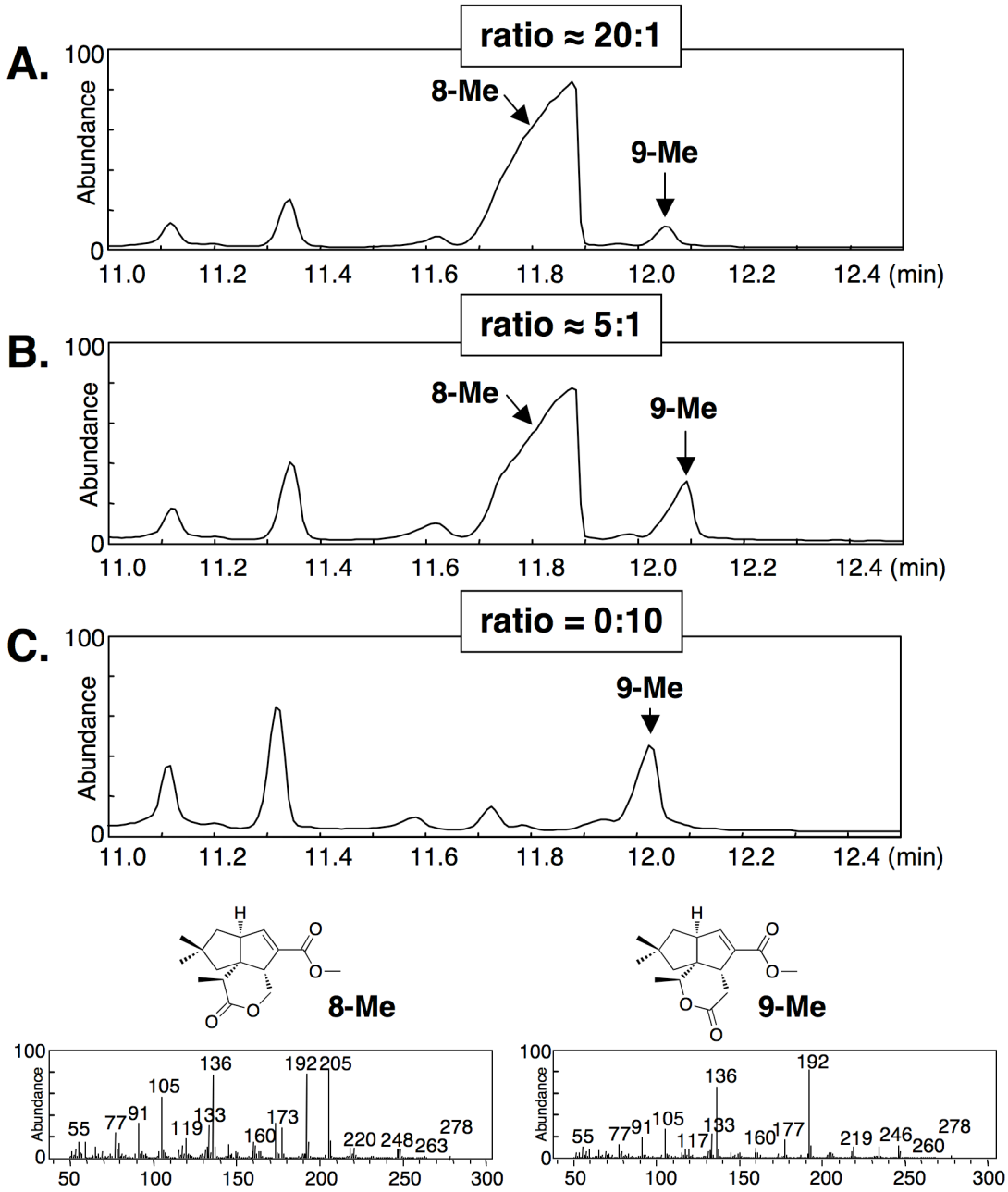

Figure S11. GC-MS analysis of *S. avermitilis* SUKA16 *ΔptlE ΔptlD* double mutant complementation experiments. A. Complementation of *ΔDptlE ΔptlD* double mutant with pKU464*aac(3)IV::ermEp::ptlE/pntE300*. B. Complementation of *ΔDptlE ΔptlD* mutant with pKU464*aac(3)IV::ermEp::ptlE/pntE600*. C. Complementation of *ΔDptlE ΔptlD* mutant with pKU464*aac(3)IV::ermEp::ptlE/pntE1200*.

References

1. Martin, D. G., Slomp, G., Mizsak, S., Duchamp, D. J., and Chidester, C. G. (1970) The structure and absolute configuration of pentalenolactone (PA 132), *Tetrahedron Lett.*, 4901-4904.

2. Keller-Schierlein, W., Lemke, J., Nyfeler, R., and Zähner, H. (1972) Stoffwechselprodukte von Mikroorganismen. 105. Arenaemycin E, D, und C., *Arch. Mikrobiol. 84*, 301-316.

3. Jiang, J., Tetzlaff, C. N., Takamatsu, S., Iwatsuki, M., Komatsu, M., Ikeda, H., and Cane, D. E. (2009) Genome mining in *Streptomyces avermitilis*: A biochemical Baeyer-Villiger reaction and discovery of a new branch of the pentalenolactone family tree, *Biochemistry 48*, 6431-6440.

4. Paget, M. S., Chamberlin, L., Atrih, A., Foster, S. J., and Buttner, M. J. (1999) Evidence that the extracytoplasmic function sigma factor *sigmaE* is required for normal cell wall structure in *Streptomyces coelicolor* A3(2), *J. Bacteriol. 181*, 204-211.

5. Macneil, D. J., Gewain, K. M., Ruby, C. L., Dezeny, G., Gibbons, P. H., and Macneil, T. (1992) Analysis of *Streptomyces-Avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector, *Gene 111*, 61-68.

6. Gust, B., Challis, G. L., Fowler, K., Kieser, T., and Chater, K. F. (2003) PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin, *Proc. Natl. Acad. Sci. U S A 100*, 1541-1546.

7. Datsenko, K. A., and Wanner, B. L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products, *Proc. Natl. Acad. Sci. U S A 97*, 6640-6645.

8. Short, J. M., Fernandez, J. M., Sorge, J. A., and Huse, W. D. (1988) Lambda ZAP: a bacteriophage lambda expression vector with *in vivo* excision properties, *Nucl. Acids Res. 16*, 7583-7600.

9. Alting-Mees, M. A., and Short, J. M. (1989) pBluescript II: gene mapping vectors, *Nucl. Acids Res. 17*, 9494.

10. Sun, Y., Zhou, X., Liu, J., Bao, K., Zhang, G., Tu, G., Kieser, T., and Deng, Z. (2002) '*Streptomyces nanchangensis*', a producer of the insecticidal polyether antibiotic nanchangmycin and the antiparasitic macrolide meilingmycin, contains multiple polyketide gene clusters, *Microbiology 148*, 361-371.

11. Bierman, M., Logan, R., O'Brien, K., Seno, E. T., Rao, R. N., and Schoner, B. E. (1992) Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp, *Gene 116*, 43-49.

12. He, Y., Wang, Z., Bai, L., Liang, J., Zhou, X., and Deng, Z. (2010) Two pHZ1358-derivative vectors for efficient gene knockout in Streptomyces, *J. Microbiol. Biotechnol. 20*, 678-682.