Supporting Information

Heterologous Biosynthesis of Type II Polyketide Products Using *E. coli*

Xiangyang Liu[1#], Kangmin Hua[1#], Dongxu Liu[1], Zhen-long Wu[2], Ying Wang[2], Haoran Zhang[3], Zixin Deng[1], Blaine A. Pfeifer[4], Ming Jiang[1]*

Table S1 *E. coli* strains and plasmids used in this study.

Strain or plasmid	Relevant properties or genetic marker	Source or reference
Strains		
BL21(DE3)	F^-, *omp*T, *hsd*S_B(r_B^-,m_B^-), *dcm*, *gal*, λ(DE3)	Novagen
DH10B	General cloning and plasmid maintenance	GibcoBRL
BAP1	F^-*ompT hsdS*$_B$ (r_B^- m_B^-) *gal dcm* (DE3) ΔprpRBCD::T7prom-sfp, T7prom-prpE	(1)
BAP1/pGro7/pXY-1/pXY-3/pXY-6	*E.coli* strains producing dehydrorabelomycin	This study
BAP1/pGro7/pXY-2/pXY-3/pXY-6	*E.coli* strains producing dehydrorabelomycin	This study
BAP1/pGro7/pXY-1/pXY-3/pXY-5	*E.coli* strains producing prejadomycin	This study
BAP1/pGro7/pXY-1/pXY-3/pXY-	*E.coli* strains producing UWM6,	This study

4	prejadomycin and rabelomycin	
BAP1/pGro7/pXY-7/pXY-8	*E.coli* strains producing TW95c	This study
Plasmids		
pET28a	T7 promoter, Kan^r	Novagen
pET32a	T7 promoter, Amp^r	Novagen
pCDFduet	T7 promoter, Str^r	Novagen
pGro7	GroES-GroEL, Cm^r	TAKARA
pXY-1	pET28a-*alpAB*	This study
pXY-2	pET28a-*alpAB-MCAT*	This study
pXY-3	pCDF-*alpI-ravC*	This study
pXY-4	pET32a-*alpD-alpE*	This study
pXY-5	pET32a-*alpD-alpE-alpF*	This study
pXY-6	pET32a-*alpD-alpE-alpF-alpG*	This study
pXY-7	pET28a-*whiE-III-IV*	This study
pXY-8	pET32a-*whiE-V-whiE-VI-MCAT*	This study
pET28a-*alpA*	pET-28a(+) carrying *alpA*	This study
pET28a-*alpB*	pET-28a(+) carrying *alpB*	This study
pET28a-*alpC*	pET-28a(+) carrying *alpC*	This study
pET28a-*ravC*	pET-28a(+) carrying *ravC*	This study
pET32a-*alpD*	pET-32a(+) carrying *alpD*	This study

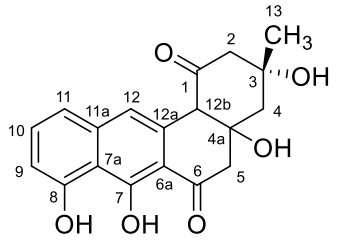
pET32a-*alpE*	pET-32a(+) carrying *alpE*	This study
pET28a-*alpF*	pET-28a(+) carrying *alpF*	This study
pET28a-*alpG*	pET-28a(+) carrying *alpG*	This study
pET32a-*alpI*	pET-32a(+) carrying *alpI*	This study
pET28a-*MCAT*	pET-28a(+) carrying *MCAT*	This study
pET32a-*whiE-V*	pET-32a(+) carrying *whiE*-ORFV	This study
pET32a-*whiE-VI*	pET-32a(+) carrying *whiE*-ORFVI	This study

Table S2. Primers used in this study

Primer	Sequence (5' to 3')
AlpA-*Nde*I-F	AAACATATGAGCGGGCGACGCGTTG
AlpA-*Hind*III-*Spe*I-R	AAAAAGCTTACTAGTCATACGAGGCTCCTTTCGG
AlpB-*Xba*I-F	AATCTAGAAGGAGATGACCGCCTCCG
AlpB-*Xba*I-8-F	AATCTAGAAGGAGATATACATATG
AlpB-*Nde*I-F	AAACATATGACCGCCTCCGTGGTGGTGACCG
AlpB-*Hind*III-*Spe*I-R	AAAAAGCTTACTAGTCAGGCGGCGCGCACGACCA
AlpC-*Nde*I-F	AAACATATGGCCAGCAAGTCCTTCACCCTCGACG
AlpC-*Hind*III-*Spe*I-R	AAAAAGCTTACTAGTCAGGCGGCCTTGGCGGGCG
AlpI-*Nde*I-F	AAACATATGCACAGCACGCTGATCGTC
AlpI-*BamH*I+1-F	AAAGGATCCGATGCACAGCACGCTGATCGTC

AlpI-*Hind*III-SpeI-R	AAAAAGCTTACTAGTCAACGGGAGGCCTCCCAGC
AlpD-*NdeI*-F	AAACATATGACCGACACCACCACCCA
AlpD-*Hind*III-*Spe*I-R	AAAAAGCTTACTAGTCAGAAGTTGCCGAGGCCGC
AlpE-*Nde*I-F	AAACATATGAGCCTCATGACCACACG
AlpE-*Hind*III-*Spe*I-R	AAAAAGCTTACTAGTCAGCCCTTCTTCTGCTCGG
AlpF-*Nde*I-F	AAACATATGGCAGCGGACGCCCTG
AlpF-*Hind*III-*Spe*I-R	AAAAAGCTTACTAGTCAGCGGGGAGGCCCGAA
AlpG-*Nde*I-F	AAACATATGGAAGGGACAGCGGCGG
AlpG-*Hind*III-*Spe*I-R	AAAAAGCTTACTAGTCAGCGGGGCGGGCCGAA
MCAT-*Nde*I-F	AAACATATGCTCGTACTCGTCGCTC
MCAT-*Hind*III-*Spe*I-R	AAAAAGCTTACTAGTCAGGCCTGGGTGTGCTC
*whiE-orfIII-Nde*I-F	ACATATGACCCGGCGCCGGGTCG
*whiE-orfIV-Hind*III-R	AAAAGCTTACTAGTCAGCGGCCCTCGGG
whiE-orfV-NdeI-F	AAACATATGACCGATCAGCAGCTGG
*whiE-orfV-Hind*III-R	AAAAAGCTTACTAGTCACACTCCCGTCTTGAG
*whiE-orfVI-Nde*I-F	AAACATATGGCAGGGCACACCGACAA
*whiE-orfVI-Hind*III-R	AAAAAGCTTACTAGTCAGTCGGCGAGCACCGA

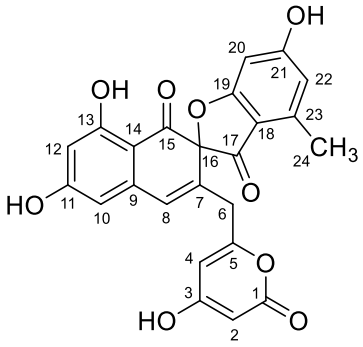
Table S3. The ^{1}H NMR and ^{13}C NMR data of UWM6[a] in CD_2Cl_2.



no.[b]	^{13}C	^{1}H[c]
1	208.1	
2	51.4	2.47 (d, J = 14.4)
		2.74 (d, J = 14.4)
3	76.2	
4	42.8	1.98 (s, 2H)
4a	76.3	
5	50.8	a 2.95 (d, J = 17.4)
		b 3.00 (d, J = 17.4)
6	202.5	
6a	108.5	
7	167.2	
7a	113.3	
8	158.5	
9	112.4	6.89 (d, J = 7.8)
10	133.8	7.53 (t, J = 7.8)
11	119.2	7.16 (d, J = 7.8)
11a	139.5	
12	117.9	6.76 (d, J = 1.8)
12a	132.7	
12b	62.2	3.93 (s)
13	30.9	1.23 (s, 3H)

[a] Spectra were obtained at 600 MHz for proton and 150 MHz for carbon and were recorded in CD_2Cl_2. Chemical shifts are reported in δ (ppm). [b] Carbons are labeled according to their number in the polyketide backbone. [c] Coupling constants are presented in hertz. Unless otherwise indicated, all proton signals integrate to 1H.

Table S4. The ^{1}H NMR and ^{13}C NMR data of TW95c[a] in CD_3OD.



no.[b]	^{13}C	^{1}H[c]
1	167.1	
2	90.2	5.22 (d, J = 2.0)
3	172.7	
4	103.9	5.67 (d, J = 2.0)
5	163.3	
6	37.3	a 3.42 (d, J = 15.6)
		b 3.32 (d, J = 15.5)
7	136.0	
8	130.8	6.81 (s)
9	141.7	
10	110.3	6.39 (d, J = 2.2)
11	169.0	
12	103.1	6.20 (d, J = 2.2)
13	167.8	
14	106.7	
15	194.1	
16	94.7	
17	192.6	
18	110.4	
19	178.2	
20	97.2	6.39 (d, J = 1.9)
21	168.9	

22	114.6	6.36 (s)
23	143.8	
24	18.0	2.32 (s, 3H)

[a] Spectra were obtained at 600 MHz for proton and 150 MHz for carbon and were recorded in CD_3OD. Chemical shifts are reported in δ (ppm). [b] Carbons are labeled according to their number in the polyketide backbone. [c] Coupling constants are presented in hertz. Unless otherwise indicated, all proton signals integrate to 1H.

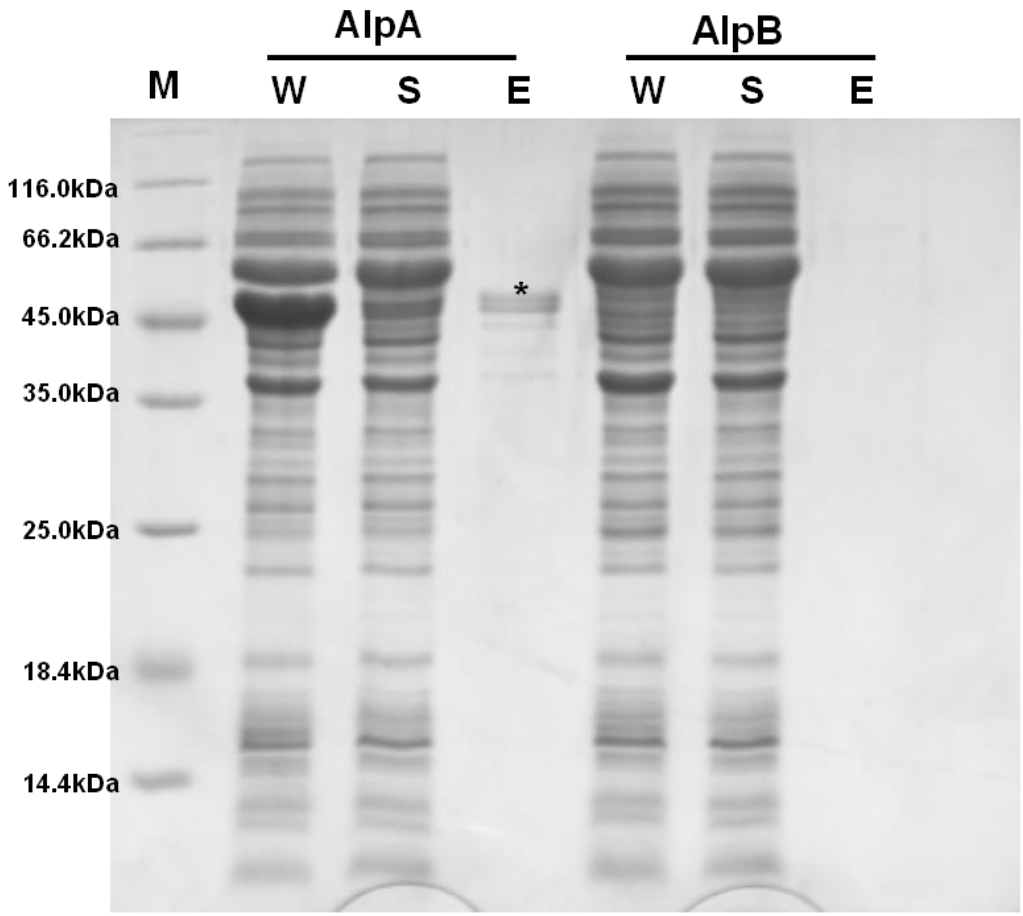


Fig. S1. Purification attempt of heterologously produced AlpA and AlpB. W: whole

cell lysate of induced cells; S: soluble fraction of induced cells; E: purified protein from affinity chromatography.

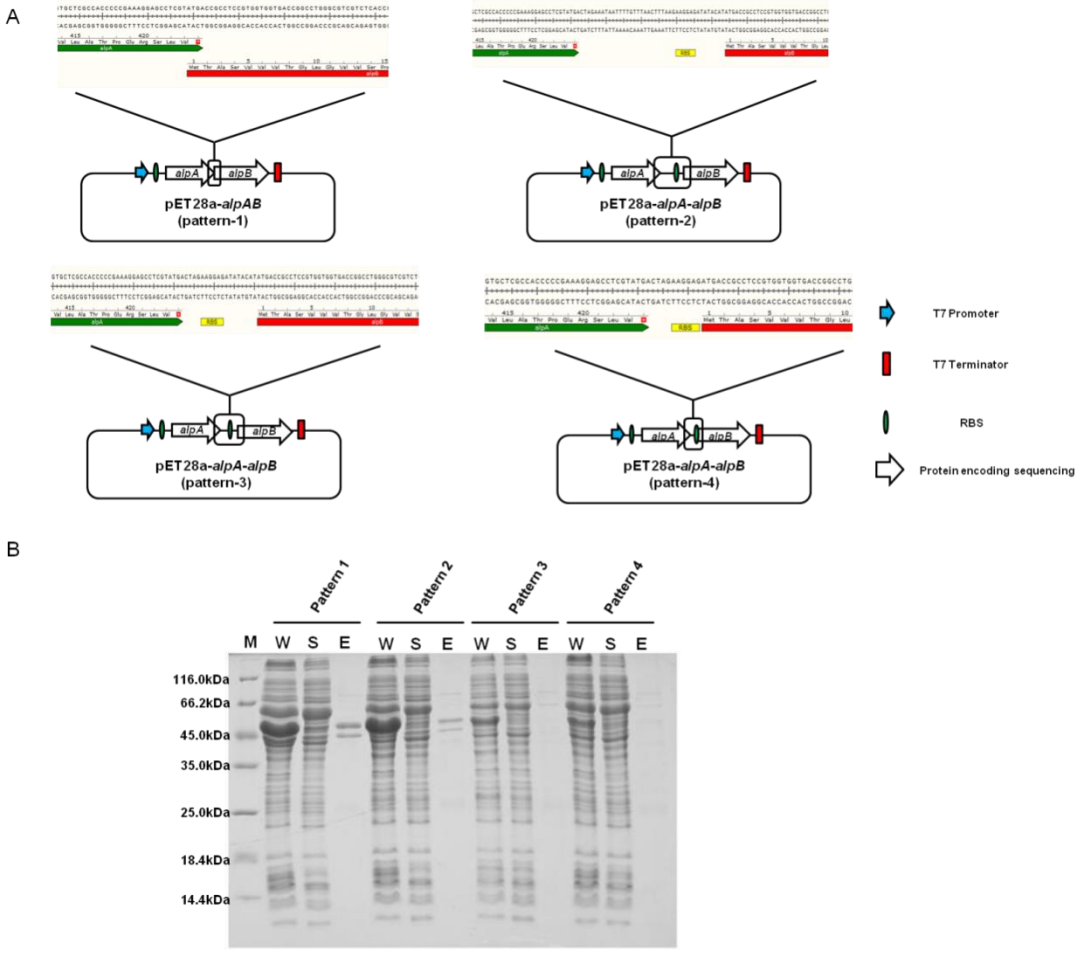


Fig. S2. (A) Plasmid designs used to test the heterologous expression of AlpA and AlpB. (B) Purification of AlpA and AlpB using different expression constructs. W: whole cell lysate of induced cells; S: soluble fraction of induced cells; E: purified protein from affinity chromatography.

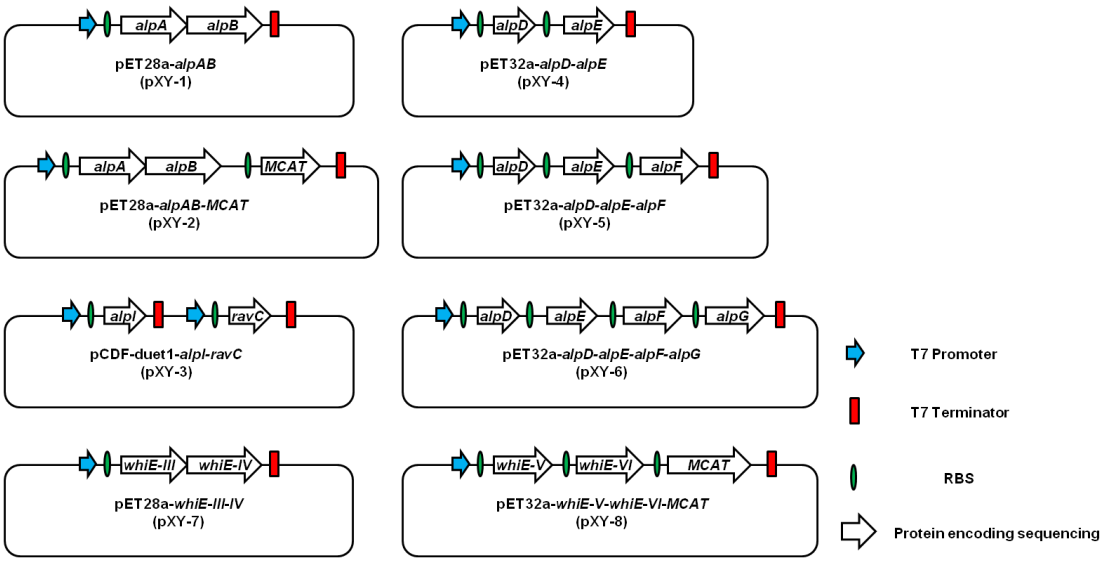


Fig. S3. Plasmids used for the heterologous production of Type II aromatic polyketides in this study

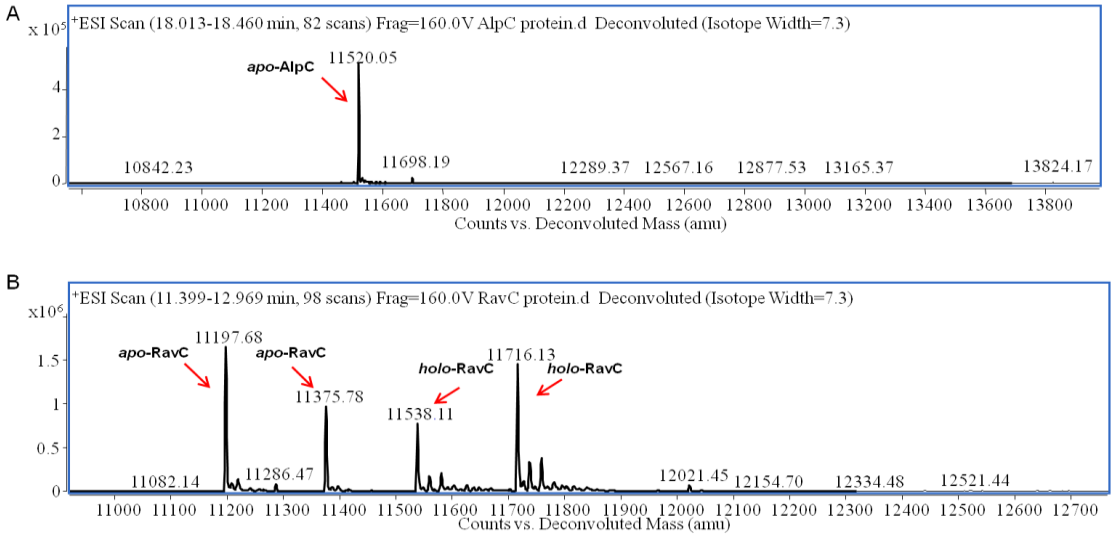


Fig. S4. Q-TOF mass analysis of the BAP1 heterologous posttranslational modification of AlpC and RavC. A) Q-TOF mass spectra of AlpC. The peak at 11520.05 Da is representative of apo-AlpC (cal. MW 111519.69 Da). B) Q-TOF mass spectra of apo- and holo-RavC. The peaks at 11197.68 and 11375.78 Da represent apo-RavC (cal. MW 11197.21 Da) and apo-RavC with fMet; whereas, the peaks at

11538.11 and 11716.13 Da represent holo-RavC and holo-RavC with fMet.

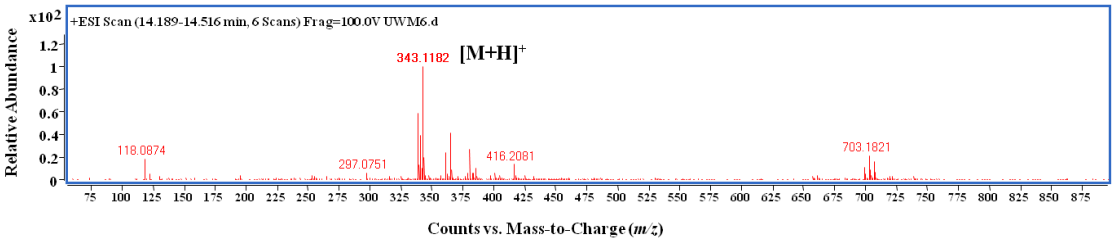


Fig. S5. Q-TOF MS spectra of UWM6.

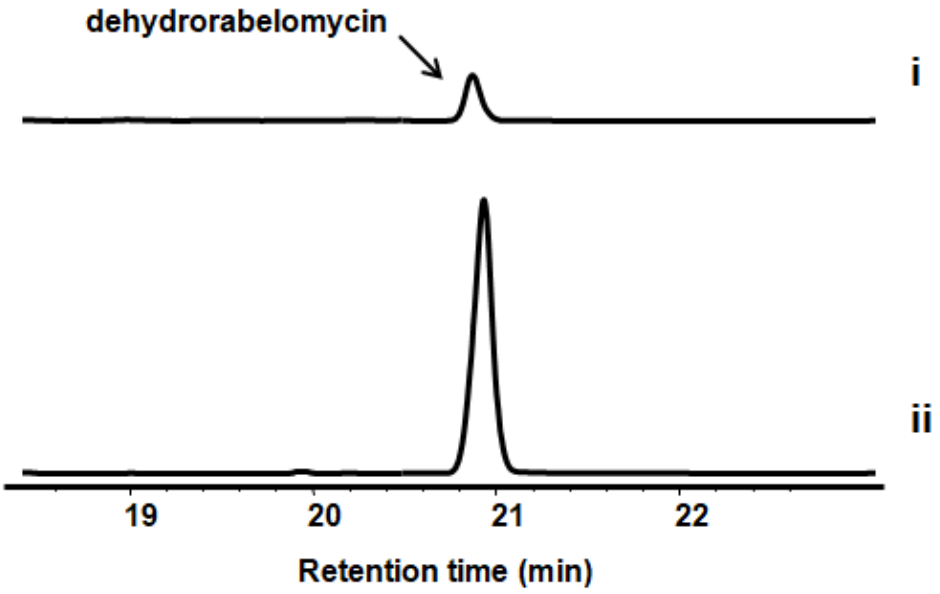


Fig. S6. HPLC analysis of dehydrorabelomycin production in *E. coli* BAP1without (i; BAP1/pGro7/pXY-1/pXY-3/pXY-6) or with (ii; BAP1/pGro7/pXY-2/pXY-3/pXY-6) MCAT.

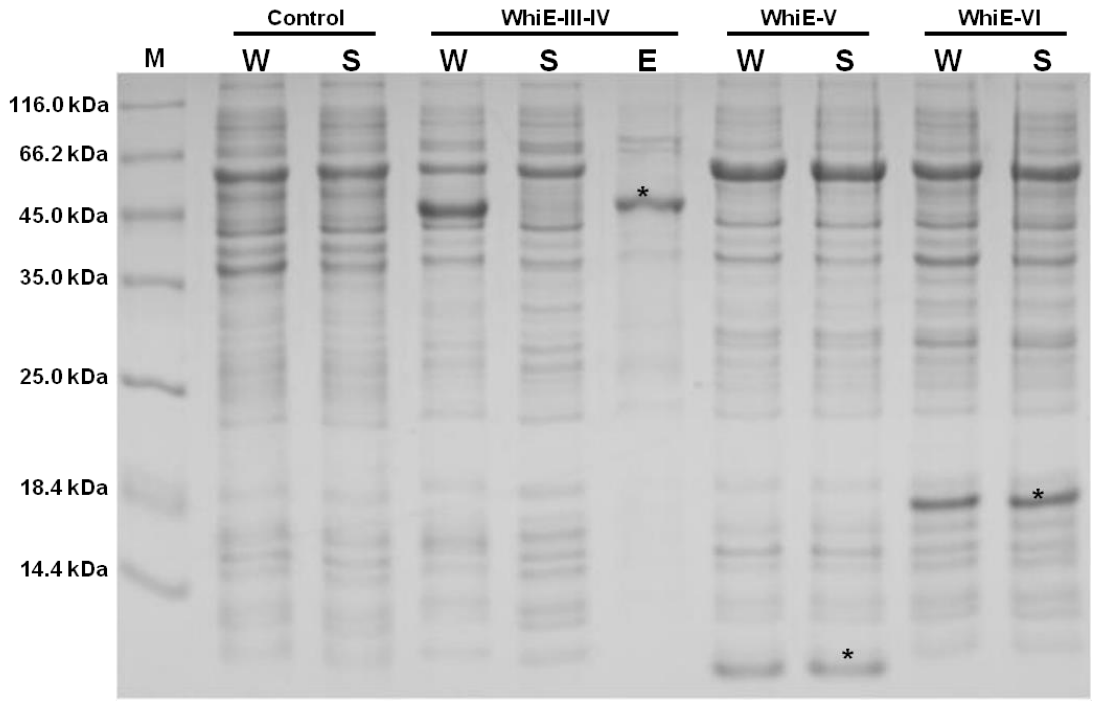
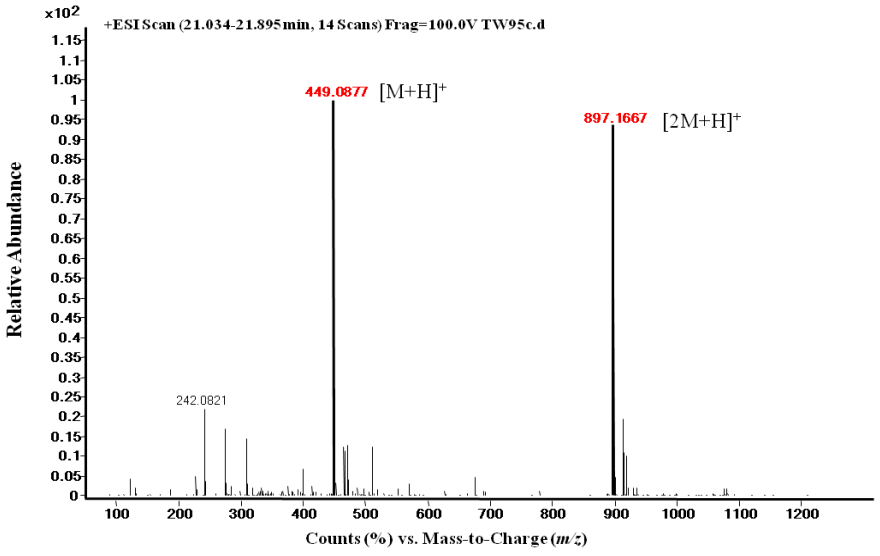


Fig. S7. SDS-PAGE analysis of proteins from the *whiE* biosynthetic pathway. M: protein marker; W: whole cell lysate of induced cells; S: soluble fraction of induced cells; E: purified protein from affinity chromatography.

A



B

Compound	Formula	$[M+H]^+$	Calculated	Found	ppm
8	$C_{24}H_{16}O_9$	$C_{24}H_{17}O_9$	449.0873	449.0877	0.89

Fig. S8. Q-TOF ESI analyses of TW95c. (A) HR-QTOF mass spectrum of TW95c

under ESI positive mode. (B) HRMS analysis data of TW95c.

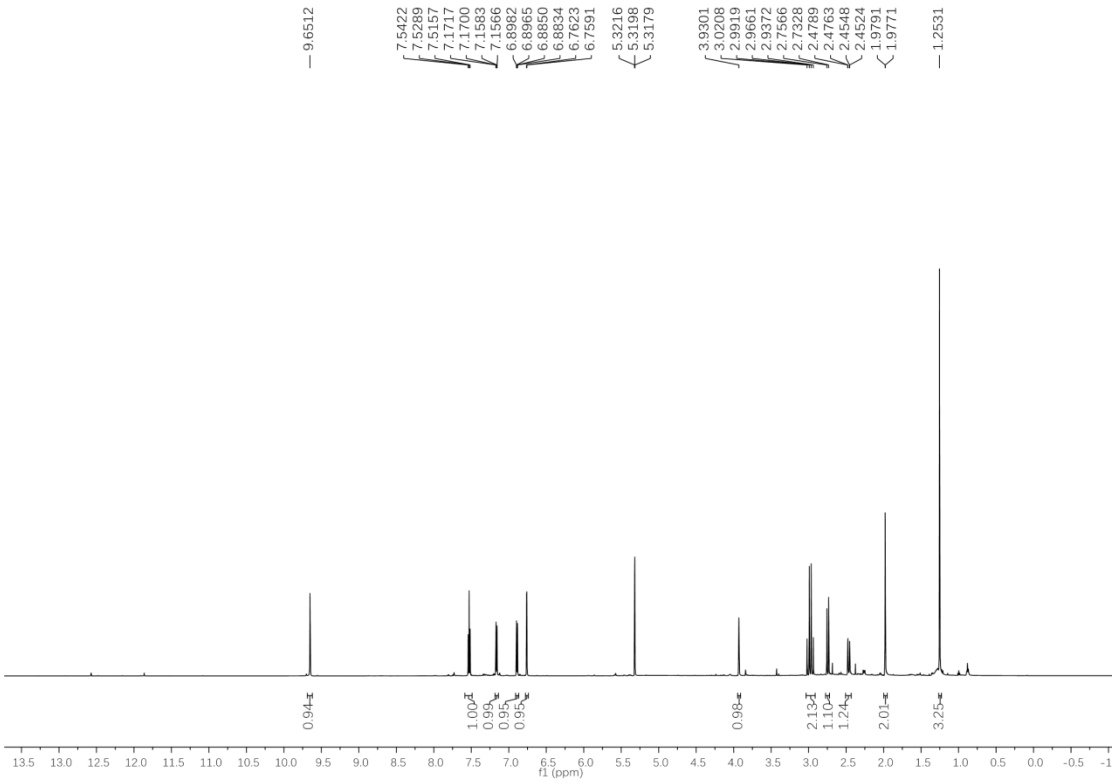


Fig. S9. ^{1}H NMR spectrum of UWM6 in CD_2Cl_2 (600 MHz)

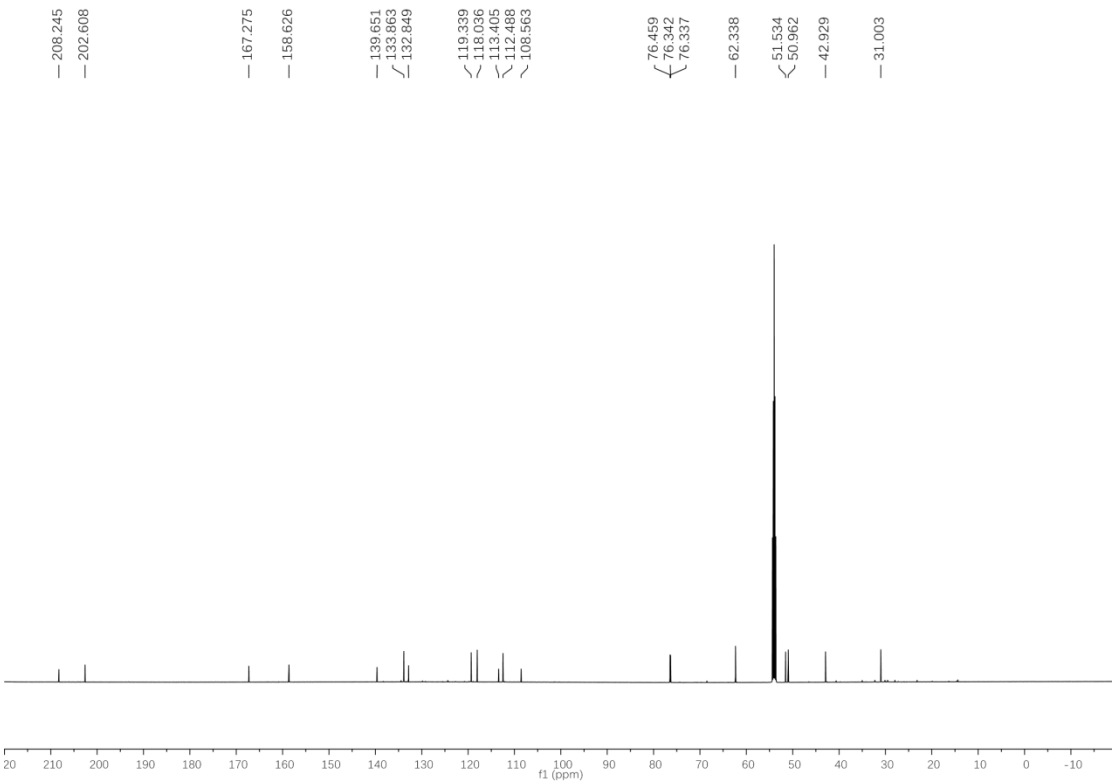


Fig. S10. ^{13}C NMR spectrum of UWM6 in CD_2Cl_2 (150 MHz)

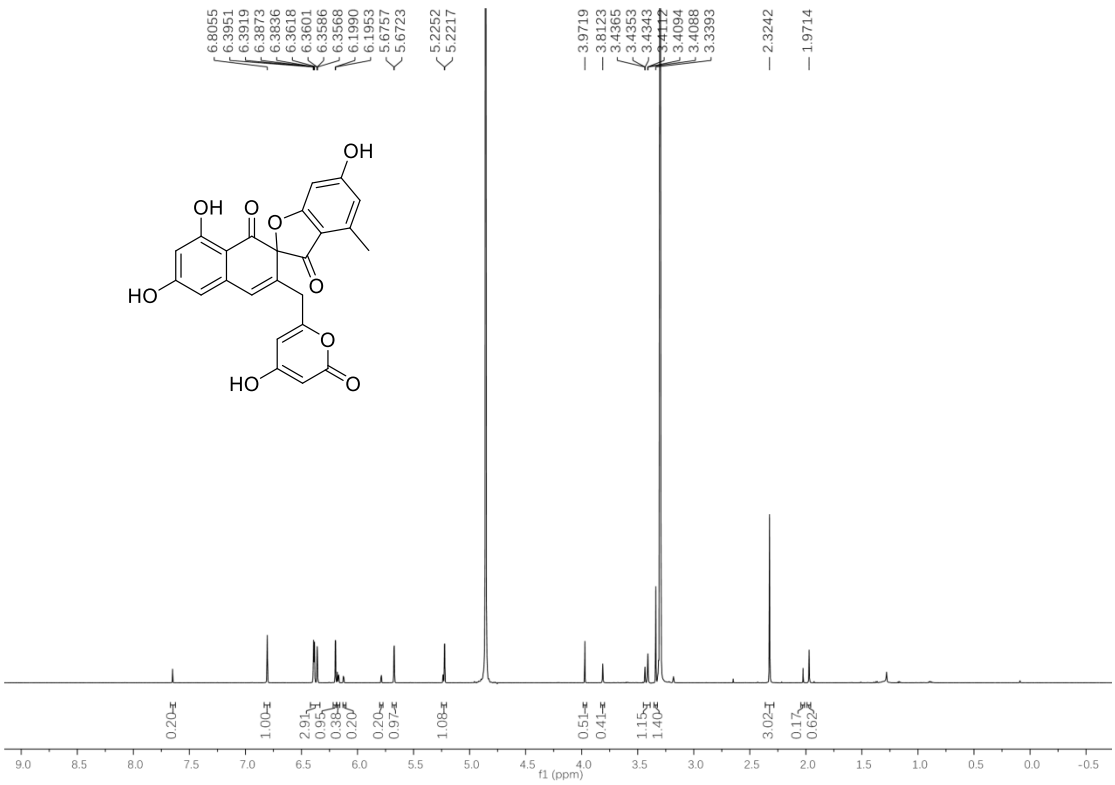


Fig. S11. ^{1}H NMR spectrum of TW95c in CD_3OD (600 MHz)

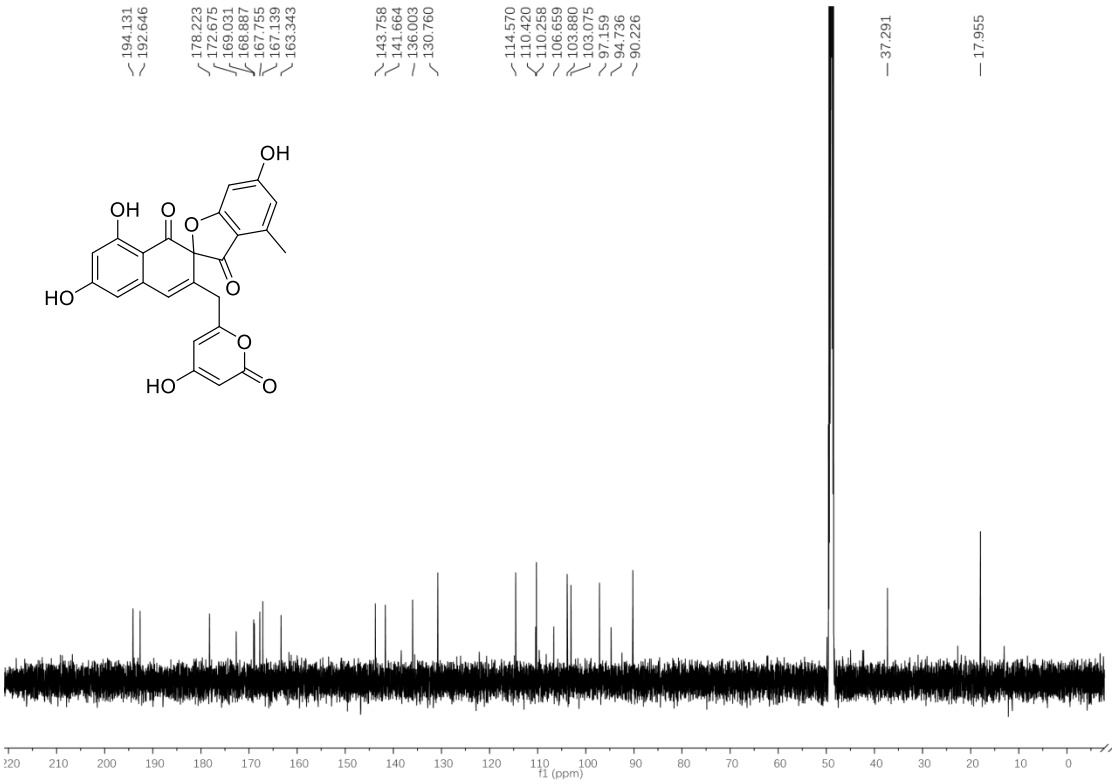


Fig. S12. ^{13}C NMR spectrum of TW95c in CD_3OD (150 MHz)

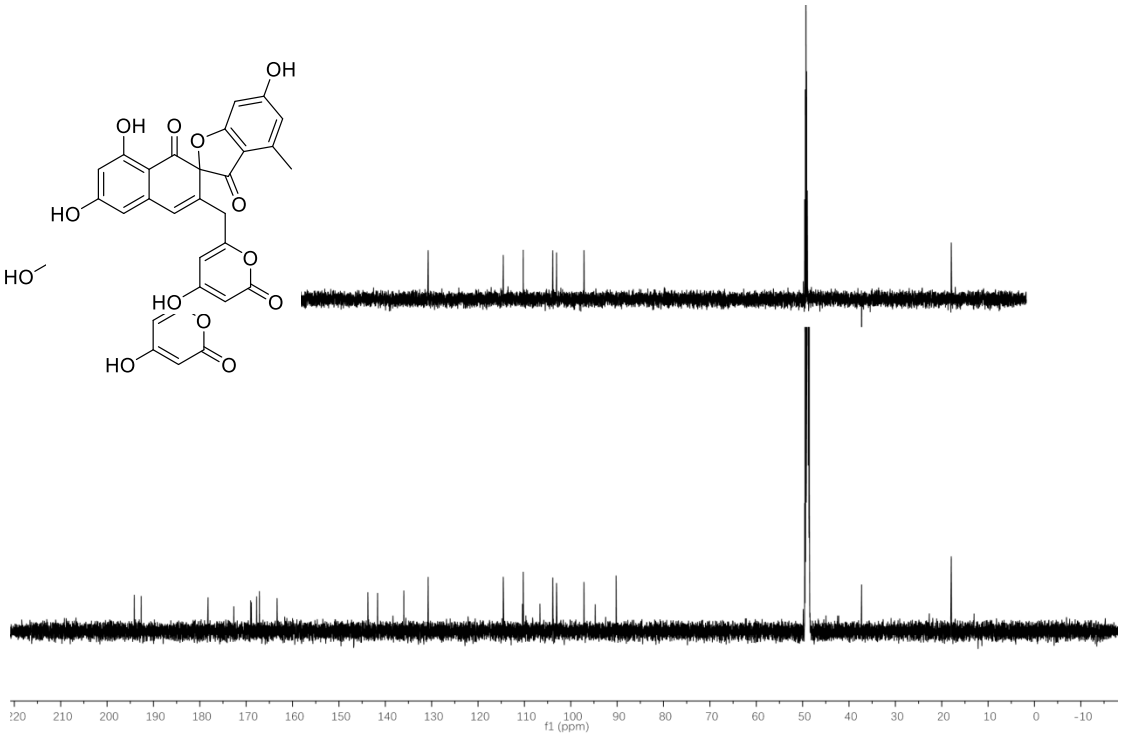


Fig. S13. DEPT-135 spectrum of TW95c in CD_3OD

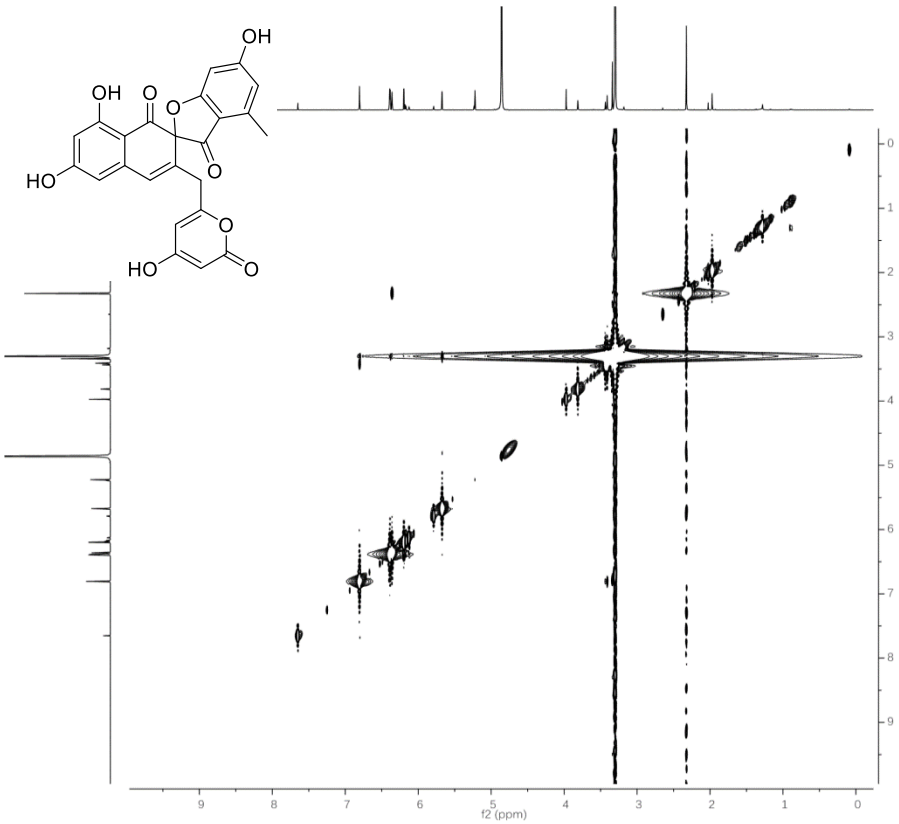


Fig. S14. ^{1}H-^{1}H COSY spectrum of TW95c in CD_3OD

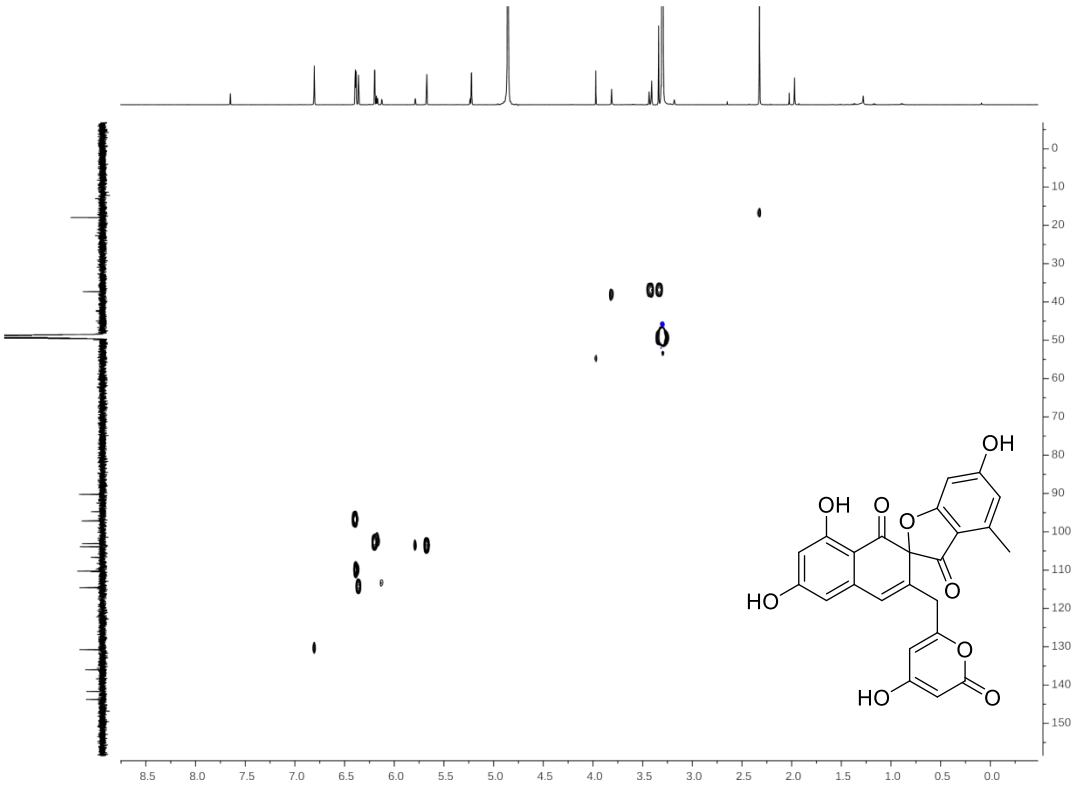


Fig. S15 HSQC spectrum of TW95c in CD_3OD

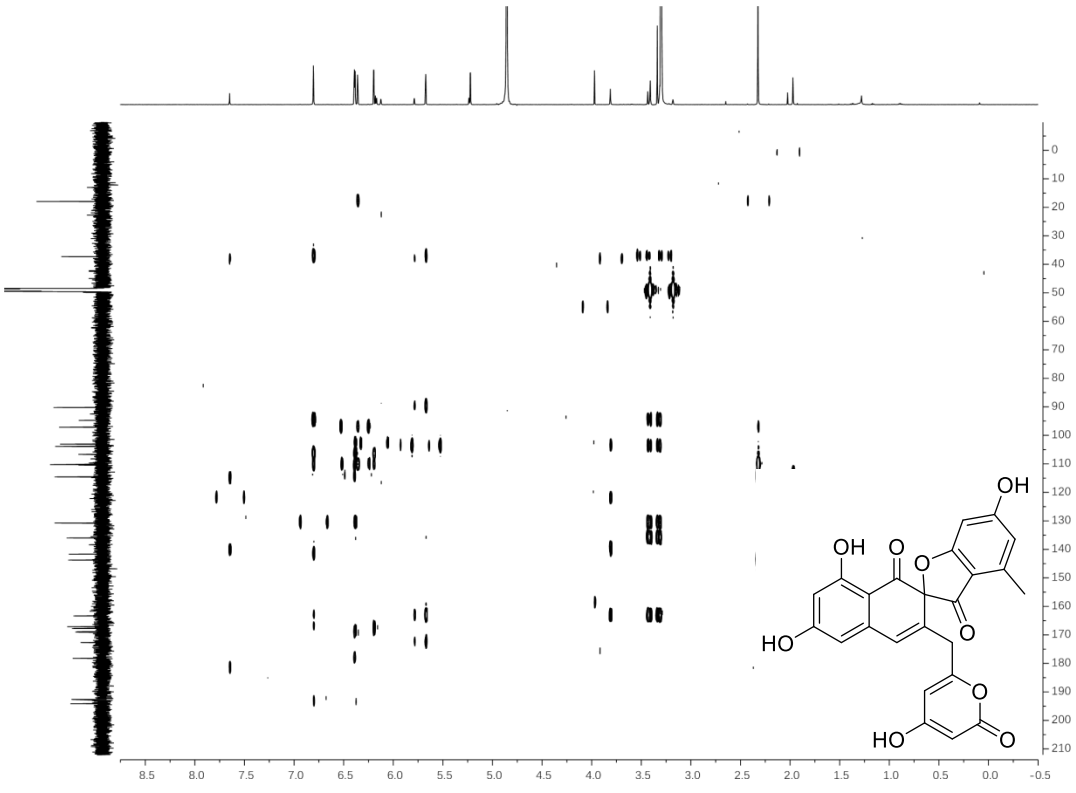


Fig. S16. HMBC spectrum of TW95c in CD_3OD

References

1. Pfeifer BA, Admiraal SJ, Gramajo H, Cane DE, & Khosla C (2001) Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. *Science* 291, 1790-1792.